Probab. Theory Relat. Fields
DOI 10.1007/s00440-008-0162-x

Stein's method on Wiener chaos

Ivan Nourdin · Giovanni Peccati

Received: 18 December 2007 / Revised: 10 May 2008

Abstract We combine Malliavin calculus with Stein's method, in order to derive explicit bounds in the Gaussian and Gamma approximations of random variables in a fixed Wiener chaos of a general Gaussian process. Our approach generalizes, refines and unifies the central and non-central limit theorems for multiple Wiener–Itô integrals recently proved (in several papers, from 2005 to 2007) by Nourdin, Nualart, Ortiz-Latorre, Peccati and Tudor. We apply our techniques to prove Berry–Esséen bounds in the Breuer–Major CLT for subordinated functionals of fractional Brownian motion. By using the well-known Mehler's formula for Ornstein–Uhlenbeck semigroups, we also recover a technical result recently proved by Chatterjee, concerning the Gaussian approximation of functionals of finite-dimensional Gaussian vectors.

Keywords Berry–Esséen bounds · Breuer–Major CLT · Fractional Brownian motion · Gamma approximation · Malliavin calculus · Multiple stochastic integrals · Normal approximation · Stein's method

Mathematics Subject Classification (2000) 60F05 · 60G15 · 60H05 · 60H07

I. Nourdin (✉)
Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie,
Boîte courrier 188, 4 Place Jussieu, 75252 Paris Cedex 5, France
e-mail: ivan.nourdin@upmc.fr

G. Peccati
Laboratoire de Statistique Théorique et Appliquée, Université Pierre et Marie Curie,
8ème étage, bâtiment A, 175 rue du Chevaleret, 75013 Paris, France
e-mail: giovanni.peccati@gmail.com

1 Introduction and overview

1.1 Motivations

Let Z be a random variable whose law is absolutely continuous with respect to the Lebesgue measure (for instance, Z is a standard Gaussian random variable or a Gamma random variable). Suppose that $\{Z_n : n \geq 1\}$ is a sequence of random variables converging in distribution towards Z, that is:

$$\text{for all } z \in \mathbb{R}, \quad P(Z_n \leq z) \longrightarrow P(Z \leq z) \quad \text{as } n \to \infty. \tag{1.1}$$

It is sometimes possible to associate an explicit uniform bound with the convergence (1.1), providing a global description of the error one makes when replacing $P(Z_n \leq z)$ with $P(Z \leq z)$ for a fixed $n \geq 1$. One of the most celebrated results in this direction is the following *Berry–Esséen Theorem* (see, e.g. Feller [17] for a proof), that we record here for future reference:

Theorem 1.1 (Berry–Esséen) *Let $(U_j)_{j\geq 1}$ be a sequence of independent and identically distributed random variables, such that $E(|U_j|^3) = \rho < \infty$, $E(U_j) = 0$ and $E(U_j^2) = \sigma^2$. Then, by setting $Z_n = \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^n U_j$, $n \geq 1$, one has that $Z_n \overset{\text{Law}}{\longrightarrow} Z \sim \mathscr{N}(0, 1)$, as $n \to \infty$, and moreover*:

$$\sup_{z\in\mathbb{R}} |P(Z_n \leq z) - P(Z \leq z)| \leq \frac{3\rho}{\sigma^3\sqrt{n}}. \tag{1.2}$$

The aim of this paper is to show that one can combine *Malliavin calculus* (see, e.g. [35]) and *Stein's method* (see, e.g. [11]), in order to obtain bounds analogous to (1.2), whenever the random variables Z_n in (1.1) can be represented as functionals of a given Gaussian field. Our results are general, in the sense that (i) they do not rely on any specific assumption on the underlying Gaussian field, (ii) they do not require that the variables Z_n have the specific form of partial sums, and (iii) they allow to deal (at least in the case of Gaussian approximations) with several different notions of *distance* between probability measures. As suggested by the title, a prominent role will be played by random variables belonging to a *Wiener chaos* of order q ($q \geq 2$), that is, random variables having the form of a *multiple stochastic Wiener–Itô integral* of order q (see Sect. 2 for precise definitions). It will be shown that our results provide substantial refinements of the central and non-central limit theorems for multiple stochastic integrals, recently proved in [33,37]. Among other applications and examples, we will provide explicit Berry–Esséen bounds in the Breuer–Major CLT (see [5]) for fields subordinated to a fractional Brownian motion.

Concerning point (iii), we shall note that, as a by-product of the flexibility of Stein's method, we will indeed establish bounds for Gaussian approximations related to a number of distances of the type

$$d_{\mathscr{H}}(X, Y) = \sup\{|E(h(X)) - E(h(Y))| : h \in \mathscr{H}\}, \tag{1.3}$$

where $\mathscr{H}$ is some suitable class of functions. For instance: by taking $\mathscr{H} = \{h : \|h\|_L \leq 1\}$, where $\|\cdot\|_L$ denotes the usual Lipschitz seminorm, one obtains the *Wasserstein* (or *Kantorovich–Wasserstein*) distance; by taking $\mathscr{H} = \{h : \|h\|_{BL} \leq 1\}$, where $\|\cdot\|_{BL} = \|\cdot\|_L + \|\cdot\|_\infty$, one obtains the *Fortet–Mourier* (or *bounded Wasserstein*) distance; by taking $\mathscr{H}$ equal to the collection of all indicators $\mathbf{1}_B$ of Borel sets, one obtains the *total variation distance*; by taking $\mathscr{H}$ equal to the class of all indicators functions $\mathbf{1}_{(-\infty,z]}$, $z \in \mathbb{R}$, one has the *Kolmogorov* distance, which is the one taken into account in the Berry–Esséen bound (1.2). In what follows, we shall sometimes denote by $d_W(.,.)$, $d_{FM}(.,.)$, $d_{TV}(.,.)$ and $d_{Kol}(.,.)$, respectively, the Wasserstein, Fortet–Mourier, total variation and Kolmogorov distances. Observe that $d_W(.,.) \geq d_{FM}(.,.)$ and $d_{TV}(.,.) \geq d_{Kol}(.,.)$. Also, the topologies induced by d_W, d_{TV} and d_{Kol} are stronger than the topology of convergence in distribution, while one can show that d_{FM} metrizes the convergence in distribution (see, e.g. [16, Chap. 11] for these and further results involving distances on spaces of probability measures).

1.2 Stein's method

We shall now give a short account of Stein's method, which is basically a set of techniques allowing to evaluate distances of the type (1.3) by means of differential operators. This theory has been initiated by Stein in the path-breaking paper [48], and then further developed in the monograph [49]. The reader is referred to [11,43,44] for detailed surveys of recent results and applications. The paper by Chatterjee [8] provides further insights into the existing literature. In what follows, we will apply Stein's method to two types of approximations, namely Gaussian and (centered) Gamma. We shall denote by $\mathscr{N}(0,1)$ a standard Gaussian random variable. The centered Gamma random variables we are interested in have the form

$$F(\nu) \overset{\text{Law}}{=} 2G(\nu/2) - \nu, \quad \nu > 0, \tag{1.4}$$

where $G(\nu/2)$ has a Gamma law with parameter $\nu/2$. This means that $G(\nu/2)$ is a (a.s. strictly positive) random variable with density $g(x) = \frac{x^{\frac{\nu}{2}-1}e^{-x}}{\Gamma(\nu/2)}\mathbf{1}_{(0,\infty)}(x)$, where Γ is the usual Gamma function. We choose this parametrization in order to facilitate the connection with our previous paper [33] (observe in particular that, if $\nu \geq 1$ is an integer, then $F(\nu)$ has a centered χ^2 distribution with ν degrees of freedom).

Standard Gaussian distribution. Let $Z \sim \mathscr{N}(0,1)$. Consider a real-valued function $h : \mathbb{R} \to \mathbb{R}$ such that the expectation $E(h(Z))$ is well-defined. The *Stein equation* associated with h and Z is classically given by

$$h(x) - E(h(Z)) = f'(x) - xf(x), \quad x \in \mathbb{R}. \tag{1.5}$$

A solution to (1.5) is a function f which is Lebesgue a.e.-differentiable, and such that there exists a version of f' verifying (1.5) for every $x \in \mathbb{R}$. The following result is basically due to Stein [48,49]. The proof of point (i) (whose content is usually referred as *Stein's lemma*) involves a standard use of the Fubini theorem (see, e.g. [47] or

[11, Lemma 2.1]). Point (ii) is proved e.g. in [11, Lemma 2.2]; point (iii) can be obtained by combining e.g. the arguments in [49, p. 25] and [9, Lemma 5.1]; a proof of point (iv) is contained in [49, Lemma 3, p. 25]; point (v) is proved in [8, Lemma 4.3].

Lemma 1.2 (i) *Let W be a random variable. Then, $W \overset{\text{Law}}{=} Z \sim \mathscr{N}(0,1)$ if, and only if,*

$$E[f'(W) - Wf(W)] = 0, \tag{1.6}$$

for every continuous and piecewise continuously differentiable function f verifying the relation $E|f'(Z)| < \infty$.

(ii) *If $h(x) = \mathbf{1}_{(-\infty,z]}(x)$, $z \in \mathbb{R}$, then (1.5) admits a solution f which is bounded by $\sqrt{2\pi}/4$, piecewise continuously differentiable and such that $\|f'\|_\infty \le 1$.*

(iii) *If h is bounded by 1/2, then (1.5) admits a solution f which is bounded by $\sqrt{\pi/2}$, Lebesgue a.e. differentiable and such that $\|f'\|_\infty \le 2$.*

(iv) *If h is bounded and absolutely continuous (then, in particular, Lebesgue-a.e. differentiable), then (1.5) has a solution f which is bounded and twice differentiable, and such that $\|f\|_\infty \le \sqrt{\pi/2}\|h - E(h(Z))\|_\infty$, $\|f'\|_\infty \le 2\|h - E(h(Z))\|_\infty$ and $\|f''\|_\infty \le 2\|h'\|_\infty$.*

(v) *If h is absolutely continuous with bounded derivative, then (1.5) has a solution f which is twice differentiable and such that $\|f'\|_\infty \le \|h'\|_\infty$ and $\|f''\|_\infty \le 2\|h'\|_\infty$.*

We also recall the relation:

$$2d_{\mathrm{TV}}(X,Y) = \sup\{|E(u(X)) - E(u(Y))| : \|u\|_\infty \le 1\}. \tag{1.7}$$

Note that point (ii) and (iii) [via (1.7)] imply the following bounds on the Kolmogorov and total variation distance between Z and an arbitrary random variable Y:

$$d_{\mathrm{Kol}}(Y,Z) \le \sup_{f \in \mathscr{F}_{\mathrm{Kol}}} |E(f'(Y) - Yf(Y))| \tag{1.8}$$

$$d_{\mathrm{TV}}(Y,Z) \le \sup_{f \in \mathscr{F}_{\mathrm{TV}}} |E(f'(Y) - Yf(Y))| \tag{1.9}$$

where $\mathscr{F}_{\mathrm{Kol}}$ and $\mathscr{F}_{\mathrm{TV}}$ are, respectively, the class of piecewise continuously differentiable functions that are bounded by $\sqrt{2\pi}/4$ and such that their derivative is bounded by 1, and the class of piecewise continuously differentiable functions that are bounded by $\sqrt{\pi/2}$ and such that their derivative is bounded by 2.

Analogously, by using (iv) and (v) along with the relation $\|h\|_L = \|h'\|_\infty$, one obtains

$$d_{\mathrm{FM}}(Y,Z) \le \sup_{f \in \mathscr{F}_{\mathrm{FM}}} |E(f'(Y) - Yf(Y))|, \tag{1.10}$$

$$d_{\mathrm{W}}(Y,Z) \le \sup_{f \in \mathscr{F}_{\mathrm{W}}} |E(f'(Y) - Yf(Y))|, \tag{1.11}$$

where: $\mathscr{F}_{\mathrm{FM}}$ is the class of twice differentiable functions that are bounded by $\sqrt{2\pi}$, whose first derivative is bounded by 4, and whose second derivative is bounded by 2; $\mathscr{F}_{\mathrm{W}}$ is the class of twice differentiable functions, whose first derivative is bounded by 1 and whose second derivative is bounded by 2.

Centered Gamma distribution. Let $F(\nu)$ be as in (1.4). Consider a real-valued function $h : \mathbb{R} \to \mathbb{R}$ such that the expectation $E[h(F(\nu))]$ exists. The *Stein equation* associated with h and $F(\nu)$ is:

$$h(x) - E[h(F(\nu))] = 2(x+\nu)f'(x) - xf(x), \quad x \in (-\nu, +\infty). \tag{1.12}$$

The following statement collects some slight variations around results proved by Diaconis and Zabell [15], Luk [26], Pickett [41], Schoutens [46] and Stein [49]. It is the "Gamma counterpart" of Lemma 1.2. The proof is detailed in Sect. 7.

Lemma 1.3 (i) *Let W be a real-valued random variable* [*not necessarily with values in* $(-\nu, +\infty)$] *whose law admits a density with respect to the Lebesgue measure. Then, $W \stackrel{\text{Law}}{=} F(\nu)$ if, and only if,*

$$E[2(W+\nu)_+ f'(W) - Wf(W)] = 0, \tag{1.13}$$

where $a_+ := \max(a, 0)$, for every smooth function f such that the mapping $x \mapsto 2(x+\nu)_+ f'(x) - xf(x)$ is bounded.

(ii) *If $|h(x)| \le c\exp(ax)$ for every $x > -\nu$ and for some $c > 0$ and $a < 1/2$, and if h is twice differentiable, then* (1.12) *has a solution f which is bounded on $(-\nu, +\infty)$, differentiable and such that $\|f\|_\infty \le 2\|h'\|_\infty$ and $\|f'\|_\infty \le \|h''\|_\infty$.*

(iii) *Suppose that $\nu \ge 1$ is an integer. If $|h(x)| \le c\exp(ax)$ for every $x > -\nu$ and for some $c > 0$ and $a < 1/2$, and if h is twice differentiable with bounded derivatives, then* (1.12) *has a solution f which is bounded on $(-\nu, +\infty)$, differentiable and such that $\|f\|_\infty \le \sqrt{2\pi/\nu}\|h\|_\infty$ and $\|f'\|_\infty \le \sqrt{2\pi/\nu}\|h'\|_\infty$.*

Now define

$$\mathscr{H}_1 = \{h \in \mathscr{C}_b^2 : \|h\|_\infty \le 1,\ \|h'\|_\infty \le 1\}, \tag{1.14}$$

$$\mathscr{H}_2 = \{h \in \mathscr{C}_b^2 : \|h\|_\infty \le 1,\ \|h'\|_\infty \le 1,\ \|h''\|_\infty \le 1\}, \tag{1.15}$$

$$\mathscr{H}_{1,\nu} = \mathscr{H}_1 \cap \mathscr{C}_b^2(\nu) \tag{1.16}$$

$$\mathscr{H}_{2,\nu} = \mathscr{H}_2 \cap \mathscr{C}_b^2(\nu) \tag{1.17}$$

where $\mathscr{C}_b^2$ denotes the class of twice differentiable functions (with support in $\mathbb{R}$) and with bounded derivatives, and $\mathscr{C}_b^2(\nu)$ denotes the subset of $\mathscr{C}_b^2$ composed of functions with support in $(-\nu, +\infty)$. Note that point (ii) in the previous statement implies that, by adopting the notation (1.3) and for every $\nu > 0$ and every real random variable Y [not necessarily with support in $(-\nu, +\infty)$],

$$d_{\mathscr{H}_{2,\nu}}(Y, F(\nu)) \le \sup_{f \in \mathscr{F}_{2,\nu}} |E[2(Y+\nu)f'(Y) - Yf(Y)]| \tag{1.18}$$

where $\mathscr{F}_{2,\nu}$ is the class of differentiable functions with support in $(-\nu, +\infty)$, bounded by 2 and whose first derivatives are bounded by 1. Analogously, point (iii) implies that, for every integer $\nu \ge 1$,

$$d_{\mathscr{H}_{1,\nu}}(Y, F(\nu)) \le \sup_{f \in \mathscr{F}_{1,\nu}} |E[2(Y+\nu)f'(Y) - Yf(Y)]|, \tag{1.19}$$

where $\mathscr{F}_{1,\nu}$ is the class of differentiable functions with support in $(-\nu, +\infty)$, bounded by $\sqrt{2\pi/\nu}$ and whose first derivatives are also bounded by $\sqrt{2\pi/\nu}$. A little inspection shows that the following estimates also hold: for every $\nu > 0$ and every random variable Y,

$$d_{\mathscr{H}_2}(Y, F(\nu)) \le \sup_{f \in \mathscr{F}_2} |E[2(Y+\nu)_+ f'(Y) - Yf(Y)]| \tag{1.20}$$

where $\mathscr{F}_2$ is the class of functions (defined on $\mathbb{R}$) that are continuous and differentiable on $\mathbb{R}\backslash\{\nu\}$, bounded by $\max\{2, 2/\nu\}$, and whose first derivatives are bounded by $\max\{1, 1/\nu + 2/\nu^2\}$. Analogously, for every integer $\nu \ge 1$,

$$d_{\mathscr{H}_1}(Y, F(\nu)) \le \sup_{f \in \mathscr{F}_1} |E[2(Y+\nu)_+ f'(Y) - Yf(Y)]|, \tag{1.21}$$

where $\mathscr{F}_1$ is the class of functions (on $\mathbb{R}$) that are continuous and differentiable on $\mathbb{R}\backslash\{\nu\}$, bounded by $\max\{\sqrt{2\pi/\nu}, 2/\nu\}$, and whose first derivatives are bounded by $\max\{\sqrt{2\pi/\nu}, 1/\nu + 2/\nu^2\}$.

Now, the crucial issue is how to estimate the right-hand side of (1.8)–(1.11) and (1.18)–(1.21) for a given choice of Y. Since Stein's initial contribution [48], an impressive panoply of techniques has been developed in this direction (see again [10] or [43] for a survey; here, we shall quote e.g.: exchangeable pairs [49], diffusion generators [3, 19], size-bias transforms [20], zero-bias transforms [21], local dependency graphs [10] and graphical-geometric rules [8]). Starting from the next section, we will show that, when working within the framework of functionals of Gaussian fields, one can very effectively estimate expressions such as (1.8)–(1.11), (1.18) and (1.19) by using techniques of Malliavin calculus. Interestingly, a central role is played by an infinite dimensional version of the same *integration by parts formula* that is at the very heart of Stein's characterization of the Gaussian distribution.

1.3 The basic approach (with some examples)

Let $\mathfrak{H}$ be a real separable Hilbert space and, for $q \geq 1$, let $\mathfrak{H}^{\otimes q}$ (resp. $\mathfrak{H}^{\odot q}$) be the qth tensor product (resp. qth symmetric tensor product) of $\mathfrak{H}$. We write

$$X = \{X(h) : h \in \mathfrak{H}\} \tag{1.22}$$

to indicate a centered isonormal Gaussian process on $\mathfrak{H}$. For every $q \geq 1$, we denote by I_q the isometry between $\mathfrak{H}^{\odot q}$ (equipped with the norm $\sqrt{q!}\|\cdot\|_{\mathfrak{H}^{\otimes q}}$) and the qth Wiener chaos of X. Note that, if $\mathfrak{H}$ is a σ-finite measure space with no atoms, then each random variable $I_q(h)$, $h \in \mathfrak{H}^{\odot q}$, has the form of a multiple Wiener–Itô integral of order q. We denote by $\mathrm{L}^2(X) = \mathrm{L}^2(\Omega, \sigma(X), P)$ the space of square integrable functionals of X, and by $\mathbb{D}^{1,2}$ the domain of the Malliavin derivative operator D (see the forthcoming Sect. 2 for precise definitions). Recall that, for every $F \in \mathbb{D}^{1,2}$, the derivative DF is a random element with values in $\mathfrak{H}$.

We start by observing that, thanks to (1.6), for every $h \in \mathfrak{H}$ such that $\|h\|_{\mathfrak{H}} = 1$ and for every smooth function f, we have $E[X(h)f(X(h))] = E[f'(X(h))]$. Our point is that this last relation is a very particular case of the following corollary of the celebrated *integration by parts formula* of Malliavin calculus: for every $Y \in \mathbb{D}^{1,2}$ with zero mean,

$$E[Yf(Y)] = E[\langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}} f'(Y)], \tag{1.23}$$

where the linear operator L^{-1} is the inverse of the *generator of the Ornstein–Uhlebeck semigroup*, noted L. The reader is referred to Sects. 2 and 3 for definitions and for a full discussion of this point; here, we shall note that L is an infinite-dimensional version of the generator associated with Ornstein–Uhlenbeck diffusions (see [35, Sect. 1.4] for a proof of this fact), an object which is also crucial in the Barbour–Götze "generator approach" to Stein's method [3,19].

It follows that, for every $Y \in \mathbb{D}^{1,2}$ with zero mean, the expressions appearing on the right-hand side of (1.8)–(1.11) [or (1.18)–(1.21)] can be assessed by first replacing $Yf(Y)$ with $\langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}} f'(Y)$ inside the expectation, and then by evaluating the L^2 distance between 1 (resp. $2Y + 2\nu$) and the inner product $\langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}}$. In general, these computations are carried out by first resorting to the representation of $\langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}}$ as a (possibly infinite) series of multiple stochastic integrals. We will see that, when $Y = I_q(g)$, for $q \geq 2$ and some $g \in \mathfrak{H}^{\odot q}$, then $\langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}} = q^{-1}\|DY\|^2_{\mathfrak{H}}$. In particular, by using this last relation one can deduce bounds involving quantities that are intimately related to the central and non-central limit theorems recently proved in [33,36,37].

Remark 1.4 1. The crucial equality $E[I_q(g)f(I_q(g))] = E[q^{-1}\|DI_q(g)\|^2_{\mathfrak{H}} f'(I_q(g))]$, in the case where f is a complex exponential, has been first used in [36], in order to give refinements (as well as alternate proofs) of the main CLTs in [37,39]. The same relation has been later applied in [33], where a characterization of non-central limit theorems for multiple integrals is provided. Note that neither [33] nor [36] are concerned with Stein's method or, more generally, with bounds on distances between probability measures.

2. We will see that formula (1.23) contains as a special case a result recently proved by Chatterjee [9, Lemma 5.3], in the context of limit theorems for linear statistics of eigenvalues of random matrices. The connection between the two results can be established by means of the well-known *Mehler's formula* (see, e.g. [28, Sect. 8.5, Chap. I] or [35, Sect. 1.4]), providing a mixture-type representation of the infinite-dimensional Ornstein–Uhlenbeck semigroup. See Remarks 3.6 and 3.12 for a precise discussion of this point. See, e.g. [31] for a detailed presentation of the infinite-dimensional Ornstein–Uhlebeck semigroup.
3. We stress that the random variable $\langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}}$ appearing in (1.23) is in general *not* measurable with respect to $\sigma(Y)$. For instance, if X is taken to be the Gaussian space generated by a standard Brownian motion $\{W_t : t \geq 0\}$ and $Y = I_2(h)$ with $h \in L^2_s([0,1]^2)$, then $D_tY = 2\int_0^1 h(u,t)dW_u$, $t \in [0,1]$, and

$$\langle DY, -DL^{-1}Y\rangle_{L^2([0,1])} = 2\, I_2(h \otimes_1 h) + 2\|h\|^2_{L^2([0,1]^2)}$$

which is, in general, not measurable with respect to $\sigma(Y)$ (the symbol $h \otimes_1 h$ indicates a contraction kernel, an object that will be defined in Sect. 2).
4. Note that (1.23) also implies the relation

$$E[Yf(Y)] = E[\tau(Y)f'(Y)], \tag{1.24}$$

where $\tau(Y) = E[\langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}}|Y]$. Some general results for the existence of a real-valued function τ satisfying (1.24) are contained e.g. in [6]. Note that, in general, it is very hard to find an analytic expression for $\tau(Y)$, especially when Y is a random variable with a very complex structure, such as e.g. a multiple Wiener–Itô integral. On the other hand, we will see that, in many cases, the random variable $\langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}}$ is remarkably tractable and explicit. See the forthcoming Sect. 6, which is based on [49, Lecture VI], for a general discussions of equations of the type (1.24). See Remark 3.10 for a connection with Goldstein and Reinert's *zero bias transform* [20].
5. The reader is referred to [42] for applications of integration by parts techniques to the Stein-type estimation of drifts of Gaussian processes. See [23] for a Stein characterization of Brownian motions on manifolds by means of integration by parts formulae. See [13] for a connection between Stein's method and algebras of operators on configuration spaces.

Before proceeding to a formal discussion, and in order to motivate the reader, we shall provide two examples of the kind of results that we will obtain in the subsequent sections. The first statement involves double Wiener–Itô integrals, that is, random variables living in the second chaos of X. The proof is given in Sect. 7.

Theorem 1.5 *Let $(Z_n)_{n\geq 1}$ be a sequence belonging to the second Wiener chaos of X.*

1. *Assume that $E(Z_n^2) \to 1$ and $E(Z_n^4) \to 3$ as $n \to \infty$. Then $Z_n \xrightarrow{\text{Law}} Z \sim \mathscr{N}(0,1)$ as $n \to \infty$. Moreover, we have*:

$$d_{\mathrm{TV}}(Z_n, Z) \leq 2\sqrt{\frac{1}{6}\left|E(Z_n^4) - 3\right| + \frac{3 + E(Z_n^2)}{2}\left|E(Z_n^2) - 1\right|}.$$

2. *Fix* $\nu > 0$ *and assume that* $E(Z_n^2) \to 2\nu$ *and* $E(Z_n^4) - 12E(Z_n^3) \to 12\nu^2 - 48\nu$ *as* $n \to \infty$. *Then, as* $n \to \infty$, $Z_n \xrightarrow{\text{Law}} F(\nu)$, *where* $F(\nu)$ *has a centered Gamma distribution of parameter* ν. *Moreover, we have*:

$$d_{\mathscr{H}_2}(Z_n, F(\nu)) \leq \max\{1, 1/\nu, 2/\nu^2\}$$
$$\sqrt{\frac{1}{6}\left|E(Z_n^4) - 12E(Z_n^3) - 12\nu^2 + 48\nu\right| + \frac{|8 - 6\nu + E(Z_n^2)|}{2}\left|E(Z_n^2) - 2\nu\right|},$$

where $\mathscr{H}_2$ *is defined by* (1.15).

Note that, in the statement of Theorem 1.5, there is no mention of Malliavin operators (however, these operators will appear in the general statements presented in Sect. 3). For instance, when applied to the case where X is the isonormal process generated by a fractional Brownian motion, the first point of Theorem 1.5 can be used to derive the following bound for the Kolmogorov distance in the Breuer–Major CLT associated with quadratic transformations:

Theorem 1.6 *Let* B *be a fractional Brownian motion with Hurst index* $H \in (0, 3/4)$. *We set*

$$\sigma_H^2 = \frac{1}{2}\sum_{t \in \mathbb{Z}} \left(|t+1|^{2H} + |t-1|^{2H} - 2|t|^{2H}\right)^2 < \infty,$$

and

$$Z_n = \frac{1}{\sigma_H \sqrt{n}} \sum_{k=0}^{n-1} \left(n^{2H}(B_{(k+1)/n} - B_{k/n})^2 - 1\right), \quad n \geq 1.$$

Then, as $n \to \infty$, $Z_n \xrightarrow{\text{Law}} Z \sim \mathscr{N}(0, 1)$. *Moreover, there exists a constant* c_H (*depending uniquely on* H) *such that, for any* $n \geq 1$:

$$d_{\mathrm{Kol}}(Z_n, Z) \leq \frac{c_H}{n^{\frac{1}{2} \wedge (\frac{3}{2} - 2H)}}. \tag{1.25}$$

Note that both Theorem 1.5 and 1.6 will be significantly generalized in Sects. 3 and 4 (see, in particular, the forthcoming Theorems 3.1, 3.11 and 4.1).

Remark 1.7 1. When $H = 1/2$, then B is a standard Brownian motion (and therefore has independent increments), and we recover from the previous result the rate $n^{-1/2}$, that could be also obtained by applying the Berry–Esséen Theorem 1.1. This rate is still valid for $H < 1/2$. But, for $H > 1/2$, the rate we can prove in the Breuer–Major CLT is $n^{2H - \frac{3}{2}}$.

2. To the authors knowledge, Theorem 1.6 and its generalizations are the first Berry–Esséen bounds ever established for the Breuer–Major CLT.
3. To keep the length of this paper within limits, we do not derive the explicit expression of some of the constants [such as the quantity c_H in formula (1.25)] composing our bounds. As will become clear later on, the exact value of these quantities can be deduced by a careful bookkeeping of the bounding constants appearing at the different stages of the proofs.

1.4 Plan

The rest of the paper is organized as follows. In Sect. 2 we recall some notions of Malliavin calculus. In Sect. 3 we state and discuss our main bounds in Stein-type estimates for functionals of Gaussian fields. Section 4 contains an application to the Breuer–Major CLT. Section 5 deals with Gamma-type approximations. Section 6 provides a unified discussion of approximations by means of absolutely continuous distributions. Proofs and further refinements are collected in Sect. 7.

2 Elements of Malliavin calculus

The reader is referred to [25] or [35] for any unexplained notion discussed in this section. As in (1.22), we denote by $X = \{X(h) : h \in \mathfrak{H}\}$ an isonormal Gaussian process over $\mathfrak{H}$. By definition, X is a centered Gaussian family indexed by the elements of $\mathfrak{H}$ and such that, for every $h, g \in \mathfrak{H}$,

$$E\left[X(h)X(g)\right] = \langle h, g\rangle_{\mathfrak{H}}. \tag{2.26}$$

As before, we use the notation $\mathrm{L}^2(X) = \mathrm{L}^2(\Omega, \sigma(X), P)$. It is well-known (see again [25] or [35, Chap. 1]) that any random variable F belonging to $\mathrm{L}^2(X)$ admits the following chaotic expansion:

$$F = \sum_{q=0}^{\infty} I_q(f_q), \tag{2.27}$$

where $I_0(f_0) := E[F]$, the series converges in L^2 and the symmetric kernels $f_q \in \mathfrak{H}^{\odot q}$, $q \geq 1$, are uniquely determined by F. As already discussed, in the particular case where $\mathfrak{H} = \mathrm{L}^2(A, \mathscr{A}, \mu)$, where $(A, \mathscr{A})$ is a measurable space and μ is a σ-finite and non-atomic measure, one has that $\mathfrak{H}^{\odot q} = \mathrm{L}^2_s(A^q, \mathscr{A}^{\otimes q}, \mu^{\otimes q})$ is the space of symmetric and square integrable functions on A^q. Moreover, for every $f \in \mathfrak{H}^{\odot q}$, the random variable $I_q(f)$ coincides with the multiple Wiener-Itô integral (of order q) of f with respect to X (see [35, Chap. 1]). Observe that a random variable of the type $I_q(f)$, with $f \in \mathfrak{H}^{\odot q}$, has finite moments of all orders (see, e.g. [25, Chap. VI]). See again [35, Chap. 1] or [45] for a connection between multiple Wiener–Itô and Hermite polynomials. For every $q \geq 0$, we write J_q to indicate the orthogonal projection operator on the qth Wiener chaos associated with X, so that, if $F \in L^2(\Omega, \mathscr{F}, P)$ is as in (2.27), then $J_q F = I_q(f_q)$ for every $q \geq 0$.

Let $\{e_k,\ k \geq 1\}$ be a complete orthonormal system in $\mathfrak{H}$. Given $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$, for every $r = 0, \ldots, p \wedge q$, the rth contraction of f and g is the element of $\mathfrak{H}^{\otimes(p+q-2r)}$ defined as

$$f \otimes_r g = \sum_{i_1,\ldots,i_r=1}^{\infty} \langle f, e_{i_1} \otimes \cdots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}} \otimes \langle g, e_{i_1} \otimes \cdots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}}. \qquad (2.28)$$

Note that, in the particular case where $\mathfrak{H} = L^2(A, \mathscr{A}, \mu)$ (with μ non-atomic), one has that

$$f \otimes_r g = \int_{A^r} f(t_1, \ldots, t_{p-r}, s_1, \ldots, s_r)\, g(t_{p-r+1}, \ldots, t_{p+q-2r}, s_1, \ldots, s_r) d\mu(s_1) \ldots d\mu(s_r).$$

Moreover, $f \otimes_0 g = f \otimes g$ equals the tensor product of f and g while, for $p = q$, $f \otimes_p g = \langle f, g \rangle_{\mathfrak{H}^{\otimes p}}$. Note that, in general (and except for trivial cases), the contraction $f \otimes_r g$ is *not* a symmetric element of $\mathfrak{H}^{\otimes(p+q-2r)}$. The canonical symmetrization of $f \otimes_r g$ is written $f \widetilde{\otimes}_r g$. We also have the useful multiplication formula: if $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$, then

$$I_p(f) I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \widetilde{\otimes}_r g). \qquad (2.29)$$

Let $\mathscr{S}$ be the set of all smooth cylindrical random variables of the form

$$F = g\left(X(\phi_1), \ldots, X(\phi_n)\right)$$

where $n \geq 1$, $g : \mathbb{R}^n \to \mathbb{R}$ is a smooth function with compact support and $\phi_i \in \mathfrak{H}$. The Malliavin derivative of F with respect to X is the element of $L^2(\Omega, \mathfrak{H})$ defined as

$$DF = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i} \left(X(\phi_1), \ldots, X(\phi_n)\right) \phi_i.$$

In particular, $DX(h) = h$ for every $h \in \mathfrak{H}$. By iteration, one can define the mth derivative $D^m F$ (which is an element of $L^2(\Omega, \mathfrak{H}^{\otimes m})$) for every $m \geq 2$. As usual, for $m \geq 1$, $\mathbb{D}^{m,2}$ denotes the closure of $\mathscr{S}$ with respect to the norm $\| \cdot \|_{m,2}$, defined by the relation

$$\|F\|_{m,2}^2 = E\left[F^2\right] + \sum_{i=1}^{m} E\left[\|D^i F\|_{\mathfrak{H}^{\otimes i}}^2\right].$$

Note that, if $F \neq 0$ and F is equal to a finite sum of multiple Wiener-Itô integrals, then $F \in \mathbb{D}^{m,2}$ for every $m \geq 1$ and the law of F admits a density with respect to the Lebesgue measure. The Malliavin derivative D satisfies the following *chain rule*: if

$\varphi : \mathbb{R}^n \to \mathbb{R}$ is in $\mathscr{C}_b^1$ (that is, the collection of continuously differentiable functions with a bounded derivative) and if $\{F_i\}_{i=1,\dots,n}$ is a vector of elements of $\mathbb{D}^{1,2}$, then $\varphi(F_1, \dots, F_n) \in \mathbb{D}^{1,2}$ and

$$D\varphi(F_1, \dots, F_n) = \sum_{i=1}^{n} \frac{\partial \varphi}{\partial x_i}(F_1, \dots, F_n) DF_i.$$

Observe that the previous formula still holds when φ is a Lipschitz function and the law of $(F_1, \dots, F_n)$ has a density with respect to the Lebesgue measure on $\mathbb{R}^n$ (see, e.g. Proposition 1.2.3 in [35]).

We denote by δ the adjoint of the operator D, also called the *divergence operator*. A random element $u \in \mathrm{L}^2(\Omega, \mathfrak{H})$ belongs to the domain of δ, noted $\mathrm{Dom}\delta$, if, and only if, it satisfies

$$|E\langle DF, u\rangle_{\mathfrak{H}}| \le c_u \, \|F\|_{\mathrm{L}^2} \quad \text{for any } F \in \mathscr{S},$$

where c_u is a constant depending uniquely on u. If $u \in \mathrm{Dom}\delta$, then the random variable $\delta(u)$ is defined by the duality relationship (customarily called "integration by parts formula"):

$$E(F\delta(u)) = E\langle DF, u\rangle_{\mathfrak{H}}, \tag{2.30}$$

which holds for every $F \in \mathbb{D}^{1,2}$. One sometimes needs the following property: for every $F \in \mathbb{D}^{1,2}$ and every $u \in \mathrm{Dom}\delta$ such that Fu and $F\delta(u) + \langle DF, u\rangle_{\mathfrak{H}}$ are square integrable, one has that $Fu \in \mathrm{Dom}\delta$ and

$$\delta(Fu) = F\delta(u) - \langle DF, u\rangle_{\mathfrak{H}}. \tag{2.31}$$

The operator L, acting on square integrable random variables of the type (2.27), is defined through the projection operators $\{J_q\}_{q\ge 0}$ as $L = \sum_{q=0}^{\infty} -qJ_q$, and is called the *infinitesimal generator of the Ornstein–Uhlenbeck semigroup*. It verifies the following crucial property: a random variable F is an element of $\mathrm{Dom}L$ $(= \mathbb{D}^{2,2})$ if, and only if, $F \in \mathrm{Dom}\delta D$ (i.e. $F \in \mathbb{D}^{1,2}$ and $DF \in \mathrm{Dom}\delta$), and in this case: $\delta DF = -LF$. Note that a random variable F as in (2.27) is in $\mathbb{D}^{1,2}$ (resp. $\mathbb{D}^{2,2}$) if, and only if,

$$\sum_{q=1}^{\infty} q\|f_q\|^2_{\mathfrak{H}^{\odot q}} < \infty \quad \left(\text{resp. } \sum_{q=1}^{\infty} q^2\|f_q\|^2_{\mathfrak{H}^{\odot q}} < \infty\right),$$

and also $E[\|DF\|^2_{\mathfrak{H}}] = \sum_{q\ge 1} q\|f_q\|^2_{\mathfrak{H}^{\odot q}}$. If $\mathfrak{H} = \mathrm{L}^2(A, \mathscr{A}, \mu)$ (with μ non-atomic), then the derivative of a random variable F as in (2.27) can be identified with the element of $L^2(A \times \Omega)$ given by

$$D_a F = \sum_{q=1}^{\infty} q I_{q-1}\left(f_q(\cdot, a)\right), \quad a \in A. \tag{2.32}$$

We also define the operator L^{-1}, which is the *inverse* of L, as follows: for every $F \in \mathrm{L}^2(X)$ with zero mean, we set $L^{-1}F = \sum_{q\geq 1} -\frac{1}{q} J_q(F)$. Note that L^{-1} is an operator with values in $\mathbb{D}^{2,2}$. The following Lemma contains two statements: the first one [formula (2.33)] is an immediate consequence of the definition of L and of the relation $\delta D = -L$, whereas the second [formula (2.34)] corresponds to Lemma 2.1 in [33].

Lemma 2.1 *Fix an integer $q \geq 2$ and set $F = I_q(f)$, with $f \in \mathfrak{H}^{\odot q}$. Then,*

$$\delta D F = q F. \tag{2.33}$$

Moreover, for every integer $s \geq 0$,

$$E\left(F^s \|DF\|^2_{\mathfrak{H}}\right) = \frac{q}{s+1} E\left(F^{s+2}\right). \tag{2.34}$$

3 Stein's method and integration by parts on Wiener space

3.1 Gaussian approximations

Our first result provides explicit bounds for the normal approximation of random variables that are Malliavin derivable. Although its proof is quite easy to obtain, the following statement will be central for the rest of the paper.

Theorem 3.1 *Let $Z \sim \mathscr{N}(0,1)$, and let $F \in \mathbb{D}^{1,2}$ be such that $E(F) = 0$. Then, the following bounds are in order:*

$$d_{\mathrm{W}}(F, Z) \leq E[(1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})^2]^{1/2}, \tag{3.35}$$
$$d_{\mathrm{FM}}(F, Z) \leq 4E[(1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})^2]^{1/2}. \tag{3.36}$$

If, in addition, the law of F is absolutely continuous with respect to the Lebesgue measure, one has that

$$d_{\mathrm{Kol}}(F, Z) \leq E[(1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})^2]^{1/2}, \tag{3.37}$$
$$d_{\mathrm{TV}}(F, Z) \leq 2E[(1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})^2]^{1/2}. \tag{3.38}$$

Proof Start by observing that one can write $F = LL^{-1}F = -\delta DL^{-1}F$. Now let f be a real differentiable function. By using the integration by parts formula and the fact that $Df(F) = f'(F)DF$ (note that, for this formula to hold when f is only a.e.

differentiable, one needs F to have an absolutely continuous law, see Proposition 1.2.3 in [35]), we deduce

$$E(Ff(F)) = E[f'(F)\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}].$$

It follows that $E[f'(F) - Ff(F)] = E(f'(F)(1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}))$ so that relations (3.35)–(3.38) can be deduced from (1.8)–(1.11) and the Cauchy–Schwarz inequality. □

We shall now prove that the bounds appearing in the statement of Theorem 3.1 can be *explicitly* computed, whenever F belongs to a fixed Wiener chaos.

Proposition 3.2 *Let $q \geq 2$ be an integer, and let $F = I_q(f)$, where $f \in \mathfrak{H}^{\odot q}$. Then, $\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}} = q^{-1}\|DF\|^2_{\mathfrak{H}}$, and*

$$E[(1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})^2] = E[(1 - q^{-1}\|DF\|^2_{\mathfrak{H}})^2] \tag{3.39}$$

$$-(1 - q!\,\|f\|^2_{\mathfrak{H}^{\otimes q}})^2 + q^2 \sum_{r=1}^{q-1} (2q-2r)!(r-1)!^2 \binom{q-1}{r-1}^4 \times \|f\widetilde{\otimes}_r f\|^2_{\mathfrak{H}^{\otimes 2(q-r)}} \tag{3.40}$$

$$\leq (1 - q!\,\|f\|^2_{\mathfrak{H}^{\otimes q}})^2 + q^2 \sum_{r=1}^{q-1} (2q-2r)!(r-1)!^2 \binom{q-1}{r-1}^4 \times \|f \otimes_r f\|^2_{\mathfrak{H}^{\otimes 2(q-r)}}. \tag{3.41}$$

Proof The equality $\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}} = q^{-1}\|DF\|^2_{\mathfrak{H}}$ is an immediate consequence of the relation $L^{-1}I_q(f) = -q^{-1}I_q(f)$. From the multiplication formulae between multiple stochastic integrals, see (2.29), one deduces that

$$\|D[I_q(f)]\|^2_{\mathfrak{H}} = qq!\,\|f\|^2_{\mathfrak{H}^{\otimes q}} + q^2 \sum_{r=1}^{q-1} (r-1)! \binom{q-1}{r-1}^2 I_{2(q-r)}\left(f\widetilde{\otimes}_r f\right) \tag{3.42}$$

(see also [36, Lemma 2]). We therefore obtain (3.40) by using the orthogonality and isometric properties of multiple stochastic integrals. The inequality in (3.41) is just a consequence of the relation $\|f\widetilde{\otimes}_r f\|_{\mathfrak{H}^{\otimes 2(q-r)}} \leq \|f \otimes_r f\|_{\mathfrak{H}^{\otimes 2(q-r)}}$. □

The previous result should be compared with the forthcoming Theorem 3.3, where we collect the main findings of [36,37]. In particular, the combination of Proposition 3.2 and Theorem 3.3 shows that, for every (normalized) sequence $\{F_n : n \geq 1\}$ living in a fixed Wiener chaos, the bounds given in (3.35) and (3.36) are "tight" with respect to the convergence in distribution towards $Z \sim \mathscr{N}(0, 1)$, in the sense that these bounds converge to zero if, and only if, F_n converges in distribution to Z.

Theorem 3.3 [36,37] *Fix $q \geq 2$, and consider a sequence $\{F_n : n \geq 1\}$ such that $F_n = I_q(f_n)$, $n \geq 1$, where $f_n \in \mathfrak{H}^{\odot q}$. Assume moreover that $E[F_n^2] = q!\|f_n\|^2_{\mathfrak{H}^{\otimes q}} \to 1$. Then, the following four conditions are equivalent, as $n \to \infty$*:

(i) *F_n converges in distribution to $Z \sim \mathscr{N}(0, 1)$*;
(ii) $E[F_n^4] \to 3$;
(iii) *for every $r = 1, \ldots, q-1$, $\|f_n \otimes_r f_n\|_{\mathfrak{H}^{\otimes 2(q-r)}} \to 0$*;
(iv) *$\|DF_n\|^2_{\mathfrak{H}} \to q$ in L^2.*

The implications (i) $\leftrightarrow$ (ii) $\leftrightarrow$ (iii) have been first proved in [37] by means of stochastic calculus techniques. The fact that (iv) is equivalent to either one of conditions (i)–(iii) is proved in [36]. Note that Theorem 3.1 and Proposition 3.2 above provide an alternate proof of the implications (iii) $\to$ (iv) $\to$ (i). The implication (ii) $\to$ (i) can be seen as a drastic simplification of the "method of moments and cumulants", that is a customary tool in order to prove limit theorems for functionals of Gaussian fields (see, e.g. [5,7,18,27,50]). In [39] one can find a multidimensional version of Theorem 3.3.

Remark 3.4 Theorem 3.3 and its generalizations have been applied to a variety of frameworks, such as: p-variations of stochastic integrals with respect to Gaussian processes [2,12], quadratic functionals of bivariate Gaussian processes [14], self-intersection local times of fractional Brownian motion [24], approximation schemes for scalar fractional differential equations [32], high-frequency CLTs for random fields on homogeneous spaces [29,30,38], needlets analysis on the sphere [1], estimation of self-similarity orders [55], power variations of iterated Brownian motions [34]. We expect that the new bounds proved in Theorem 3.1 and Proposition 3.2 will lead to further refinements of these results. See Sects. 4 and 5 for applications and examples.

As shown in the following statement, the combination of Proposition 3.2 and Theorem 3.3 implies that, on any fixed Wiener chaos, the Kolmogorov, total variation and Wasserstein distances *metrize the convergence in distribution towards Gaussian random variables*. Other topological characterizations of the set of laws of random variables belonging to a fixed sum of Wiener chaoses are discussed in [25, Chap. VI].

Corollary 3.5 *Let the assumptions and notation of Theorem 3.3 prevail. Then, the fact that F_n converges in distribution to $Z \sim \mathscr{N}(0, 1)$ is equivalent to either one of the following conditions*:

(a) $d_{\mathrm{Kol}}(F_n, Z) \to 0$;
(b) $d_{\mathrm{TV}}(F_n, Z) \to 0$;
(c) $d_{\mathrm{W}}(F_n, Z) \to 0$.

Proof If $F_n \overset{\mathrm{Law}}{\to} Z$ then, by Theorem 3.3, we have necessarily that $\|DF_n\|^2_{\mathfrak{H}} \to q$ in L^2. The desired conclusion follows immediately from relations (3.35), (3.37)–(3.39). □

Note that the previous result is not trivial, since the topologies induced by d_{Kol}, d_{TV} and d_{W} are stronger than convergence in distribution.

Remark 3.6 (*Mehler's formula and Stein's method, I*). In [9, Lemma 5.3], Chatterjee has proved the following result (we use a notation which is slightly different from the original statement). Let $Y = g(V)$, where $V = (V_1, \dots, V_n)$ is a vector of centered i.i.d. standard Gaussian random variables, and $g : \mathbb{R}^n \to \mathbb{R}$ is a smooth function such that: (i) g and its derivatives have subexponential growth at infinity, (ii) $E(g(V)) = 0$, and (iii) $E(g(V)^2) = 1$. Then, for any Lipschitz function f, one has that

$$E[Yf(Y)] = E[S(V)f'(Y)], \tag{3.43}$$

where, for every $v = (v_1, \dots, v_n) \in \mathbb{R}^n$,

$$S(v) = \int_0^1 \frac{1}{2\sqrt{t}} E\left[\sum_{i=1}^n \frac{\partial g}{\partial v_i}(v) \frac{\partial g}{\partial v_i}(\sqrt{t}v + \sqrt{1-t}V)\right] dt, \tag{3.44}$$

so that, for instance, for $Z \sim \mathscr{N}(0, 1)$ and by using (1.9), Lemma 1.2 (iii), (1.7) and Cauchy–Schwarz inequality,

$$d_{\mathrm{TV}}(Y, Z) \le 2E[(S(V) - 1)^2]^{1/2}. \tag{3.45}$$

We shall prove that (3.43) is a very special case of (1.23). Observe first that, without loss of generality, we can assume that $V_i = X(h_i)$, where X is an isonormal process over some Hilbert space of the type $\mathfrak{H} = \mathrm{L}^2(A, \mathscr{A}, \mu)$ and $\{h_1, \dots, h_n\}$ is an orthonormal system in $\mathfrak{H}$. Since $Y = g(V_1, \dots, V_n)$, we have $D_a Y = \sum_{i=1}^n \frac{\partial g}{\partial x_i}(V) h_i(a)$. On the other hand, since Y is centered and square integrable, it admits a chaotic representation of the form $Y = \sum_{q\ge1} I_q(\psi_q)$. This implies in particular that $D_a Y = \sum_{q=1}^\infty q I_{q-1}(\psi_q(a, \cdot))$. Moreover, one has that $-L^{-1}Y = \sum_{q\ge1} \frac{1}{q} I_q(\psi_q)$, so that $-D_a L^{-1} Y = \sum_{q\ge1} I_{q-1}(\psi_q(a, \cdot))$. Now, let T_z, $z \ge 0$, denote the (infinite dimensional) *Ornstein–Uhlenbeck semigroup*, whose action on random variables $F \in \mathrm{L}^2(X)$ is given by $T_z(F) = \sum_{q\ge0} \mathrm{e}^{-qz} J_q(F)$. We can write

$$\begin{aligned}\int_0^1 \frac{1}{2\sqrt{t}} T_{\ln(1/\sqrt{t})}(D_a Y) dt &= \int_0^\infty e^{-z} T_z(D_a Y) dz = \sum_{q\ge1} \frac{1}{q} J_{q-1}(D_a Y) \\ &= \sum_{q\ge1} I_{q-1}(\psi_q(a, \cdot)) = -D_a L^{-1} Y.\end{aligned} \tag{3.46}$$

Now recall that *Mehler's formula* (see, e.g. [35, formula (1.54)]) implies that, for every function f with subexponential growth,

$$T_z(f(V)) = E\left[f(e^{-z}v + \sqrt{1 - e^{-2z}}V)\right]\Big|_{v=V}, \quad z \ge 0.$$

In particular, by applying this last relation to the partial derivatives $\frac{\partial g}{\partial v_i}$, $i = 1, \ldots, n$, we deduce from (3.46) that

$$\int_0^1 \frac{1}{2\sqrt{t}} T_{\ln(1/\sqrt{t})}(D_a Y)dt = \sum_{i=1}^n h_i(a) \int_0^1 \frac{1}{2\sqrt{t}} E\left[\frac{\partial g}{\partial v_i}(\sqrt{t}\, v + \sqrt{1-t}\, V)\right] dt \;\Big|_{v=V}.$$

Consequently, (3.43) follows, since

$$\begin{aligned}
&\langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}} \\
&= \left\langle \sum_{i=1}^n \frac{\partial g}{\partial v_i}(V)h_i, \sum_{i=1}^n \int_0^1 \frac{1}{2\sqrt{t}} E\left[\frac{\partial g}{\partial v_i}(\sqrt{t}\, v + \sqrt{1-t}\, V)\right] dt \;\Big|_{v=V} \; h_i \right\rangle_{\mathfrak{H}} \\
&= S(V).
\end{aligned}$$

See also Houdré and Pérez–Abreu [22] for related computations in an infinite-dimensional setting.

The following result concerns finite sums of multiple integrals.

Proposition 3.7 *For $s \geq 2$, fix s integers $2 \leq q_1 < \cdots < q_s$. Consider a sequence of the form*

$$Z_n = \sum_{i=1}^s I_{q_i}(f_n^i), \quad n \geq 1,$$

where $f_n^i \in \mathfrak{H}^{\odot q_i}$. Set

$$\mathscr{I} = \left\{(i, j, r) \in \{1, \ldots, s\}^2 \times \mathbb{N} : \; 1 \leq r \leq q_i \wedge q_j \quad \text{and} \quad (r, q_i, q_j) \neq (q_i, q_i, q_i)\right\}.$$

Then,

$$\begin{aligned}
E[(1 - \langle DZ_n, -DL^{-1}Z_n\rangle_{\mathfrak{H}})^2] &\leq 2\left(1 - \sum_{i=1}^s q_i! \|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}}\right)^2 \\
&+ 2s^2 \sum_{(i,j,r)\in\mathscr{I}} q_i^2 (r-1)!^2 \binom{q_i - 1}{r-1}^2 \binom{q_j - 1}{r-1}^2 (q_i + q_j - 2r)! \\
&\times \|f_n^i \otimes_{q_i - r} f_n^i\|_{\mathfrak{H}^{\otimes 2r}} \|f_n^j \otimes_{q_j - r} f_n^j\|_{\mathfrak{H}^{\otimes 2r}}.
\end{aligned}$$

In particular, if (as $n \to \infty$) $E[Z_n^2] = \sum_{i=1}^s q_i! \|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}} \longrightarrow 1$ and if, for any $i = 1, \ldots, s$ and $r = 1, \ldots, q_i - 1$, one has that $\|f_n^i \otimes_r f_n^i\|_{\mathfrak{H}^{\otimes 2(q_i - r)}} \longrightarrow 0$, then $Z_n \xrightarrow{\text{Law}} Z \sim \mathscr{N}(0, 1)$ as $n \to \infty$, and the inequalities in Theorem 3.1 allow to associate bounds with this convergence.

Remark 3.8 1. In principle, by using Proposition 3.7 it is possible to prove bounds for limit theorems involving the Gaussian approximation of *infinite* sums of multiple integrals, such as for instance the CLT proved in [24, Theorem 4].

2. Note that, to obtain the convergence result stated in Proposition 3.7, one does not need to suppose that the quantity $E[I_q(f_i)^2] = q_i!\|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}}$ is convergent for every i. One should compare this finding with the CLTs proved in [39], as well as the Gaussian approximations established in [38].

Proof of Proposition 3.7 Observe first that, without loss of generality, we can assume that X is an isonormal process over some Hilbert space of the type $\mathfrak{H} = \mathrm{L}^2(A, \mathscr{A}, \mu)$. For every $a \in A$, it is immediately checked that

$$D_a Z_n = \sum_{i=1}^{s} q_i I_{q_i-1}\left(f_n^i(\cdot, a)\right)$$

and

$$-D_a(L^{-1}Z_n) = D_a\left(\sum_{i=1}^{s} \frac{1}{q_i} I_{q_i}(f_n^i)\right) = \sum_{i=1}^{s} I_{q_i-1}\left(f_n^i(\cdot, a)\right).$$

This yields, using in particular the multiplication formula (2.29):

$$\begin{aligned}
&\langle DZ_n, -DL^{-1}Z_n\rangle_{\mathfrak{H}} \\
&\quad= \sum_{i,j=1}^{s} q_i \int_A I_{q_i-1}\left(f_n^i(\cdot, a)\right) I_{q_j-1}\left(f_n^j(\cdot, a)\right)\mu(da) \\
&\quad= \sum_{i,j=1}^{s} q_i \sum_{r=0}^{q_i\wedge q_j-1} r!\binom{q_i-1}{r}\binom{q_j-1}{r} I_{q_i+q_j-2-2r} \\
&\qquad\times\left(\int_A f_n^i(\cdot, a)\otimes_r f_n^j(\cdot, a)\mu(da)\right) \\
&\quad= \sum_{i,j=1}^{s} q_i \sum_{r=0}^{q_i\wedge q_j-1} r!\binom{q_i-1}{r}\binom{q_j-1}{r} I_{q_i+q_j-2-2r}\left(f_n^i\otimes_{r+1} f_n^j\right) \\
&\quad= \sum_{i,j=1}^{s} q_i \sum_{r=1}^{q_i\wedge q_j} (r-1)!\binom{q_i-1}{r-1}\binom{q_j-1}{r-1} I_{q_i+q_j-2r}\left(f_n^i\otimes_r f_n^j\right) \\
&\quad= \sum_{i=1}^{s} q_i!\|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}} + \sum_{(i,j,r)\in\mathscr{I}} q_i(r-1)!\binom{q_i-1}{r-1}\binom{q_j-1}{r-1} I_{q_i+q_j-2r}\left(f_n^i\otimes_r f_n^j\right).
\end{aligned}$$

Thus, by using (among others) inequalities of the type $(a_1 + \cdots + a_v)^2 \leq v(a_1^2 + \cdots + a_v^2)$, the isometric properties of multiple integrals as well $\|f\widetilde{\otimes}_r g\| \leq \|f \otimes_r g\|$, we obtain

$$
\begin{aligned}
&E\left([\langle DZ_n, -DL^{-1}Z_n\rangle_{\mathfrak{H}} - 1]^2\right)\\
&\quad\le 2\left(1-\sum_{i=1}^{s} q_i!\|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}}\right)^2\\
&\qquad +2E\left(\sum_{(i,j,r)\in\mathscr{I}} q_i(r-1)!\binom{q_i-1}{r-1}\binom{q_j-1}{r-1} I_{q_i+q_j-2r}\left(f_n^i\otimes_r f_n^j\right)\right)^2\\
&\quad\le 2\left(1-\sum_{i=1}^{s} q_i!\|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}}\right)^2\\
&\qquad +2s^2\sum_{(i,j,r)\in\mathscr{I}} q_i^2(r-1)!^2\binom{q_i-1}{r-1}^2\binom{q_j-1}{r-1}^2\\
&\qquad \times(q_i+q_j-2r)!\|f_n^i\otimes_r f_n^j\|^2_{\mathfrak{H}^{\otimes q_i+q_j-2r}}\\
&\quad\le 2\left(1-\sum_{i=1}^{s} q_i!\|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}}\right)^2\\
&\qquad +2s^2\sum_{(i,j,r)\in\mathscr{I}} q_i^2(r-1)!^2\binom{q_i-1}{r-1}^2\binom{q_j-1}{r-1}^2(q_i+q_j-2r)!\\
&\qquad \times\|f_n^i\otimes_{q_i-r} f_n^i\|_{\mathfrak{H}^{\otimes 2r}}\|f_n^j\otimes_{q_j-r} f_n^j\|_{\mathfrak{H}^{\otimes 2r}},
\end{aligned}
$$

the last inequality being a consequence of the (easily verified) relation

$$
\|f_n^i\otimes_r f_n^j\|^2_{\mathfrak{H}^{\otimes q_i+q_j-2r}} = \langle f_n^i\otimes_{q_i-r} f_n^i, f_n^j\otimes_{q_j-r} f_n^j\rangle_{\mathfrak{H}^{\otimes 2r}}.
$$

□

3.2 A property of $\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}$

Before dealing with Gamma approximations, we shall prove the a.s. positivity of a specific projection of the random variable $\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}$ appearing in Theorem 3.1. This fact will be used in the proof of the main result of the next section.

Proposition 3.9 *Let $F\in\mathbb{D}^{1,2}$. Then, P-a.s.,*

$$
E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}|F]\ge 0. \tag{3.47}
$$

Proof Let g be a non-negative real function, and set $G(x)=\int_0^x g(t)dt$, with the usual convention $\int_0^x = -\int_x^0$ for $x<0$. Since G is increasing and vanishing at zero, we have $xG(x)\ge 0$ for all $x\in\mathbb{R}$. In particular, $E(FG(F))\ge 0$. Moreover,

$$
E[F\,G(F)] = E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\, g(F)] = E[E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}|F]g(F)].
$$

We therefore deduce that

$$E[E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}|F]\mathbf{1}_A] \geq 0$$

for any $\sigma(F)$-measurable set A. This implies the desired conclusion. □

Remark 3.10 According to Goldstein and Reinert [20], for F as in the previous statement, there exists a random variable F^* having the F-*zero biased* distribution, that is, F^* is such that, for every absolutely continuous function f,

$$E[f'(F^*)] = E[Ff(F)].$$

By the computations made in the previous proof, one also has that

$$E[g(F^*)] = E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\, g(F)],$$

for any real-valued and smooth function g. This implies, in particular, that the conditional expectation $E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}|F]$ is a version of the Radon–Nikodym derivative of the law of F^* with respect to the law of F, whenever the two laws are equivalent.

3.3 Gamma approximations

We now combine Malliavin calculus with the Gamma approximations discussed in the second part of Sect. 1.2.

Theorem 3.11 *Fix $\nu > 0$ and let $F(\nu)$ have a centered Gamma distribution with parameter ν. Let $G \in \mathbb{D}^{1,2}$ be such that $E(G) = 0$ and the law of G is absolutely continuous with respect to the Lebesgue measure. Then:*

$$d_{\mathscr{H}_2}(G, F(\nu)) \leq K_2 E[(2\nu + 2G - \langle DG, -DL^{-1}G\rangle_{\mathfrak{H}})^2]^{1/2}, \tag{3.48}$$

and, if $\nu \geq 1$ is an integer,

$$d_{\mathscr{H}_1}(G, F(\nu)) \leq K_1 E[(2\nu + 2G - \langle DG, -DL^{-1}G\rangle_{\mathfrak{H}})^2]^{1/2}, \tag{3.49}$$

where $\mathscr{H}_1$ and $\mathscr{H}_2$ are defined in (1.14)–(1.15), $K_1 := \max\left\{\sqrt{2\pi/\nu}, 1/\nu + 2/\nu^2\right\}$ and $K_2 := \max\{1, 1/\nu + 2/\nu^2\}$.

Proof We will only prove (3.48), the proof of (3.49) being analogous. Fix $\nu > 0$. Thanks to (1.20) and (1.23) (in the case $Y = G$) and by applying Cauchy–Schwarz, we deduce that

$$\begin{aligned} d_{\mathscr{H}_2}(G, F(\nu)) &\leq \sup_{\mathscr{F}_2} |E[f'(G)(2(\nu+G)_+ - \langle DG, -DL^{-1}G\rangle_{\mathfrak{H}})]| \\ &\leq K_2 \times E[(2(\nu+G)_+ - \langle DG, -DL^{-1}G\rangle_{\mathfrak{H}})^2]^{1/2} \\ &\leq K_2 \times E[(2(\nu+G) - \langle DG, -DL^{-1}G\rangle_{\mathfrak{H}})^2]^{1/2}, \end{aligned}$$

where the last inequality is a consequence of the fact that $E[\langle DG, -DL^{-1}G\rangle_{\mathfrak{H}}|G] \geq 0$ (thanks to Proposition 3.9). □

Remark 3.12 (*Mehler's formula and Stein's method, II*). Define $Y = g(V)$ as in Remark 3.6. Then, since (3.44) and (3.45) are in order, one deduces from Theorem 3.11 that, for every $\nu > 0$,

$$d_{\mathscr{H}_2}(Y, F(\nu)) \leq K_2 E[(2\nu + 2Y - S(V))^2]^{1/2}.$$

An analogous estimate holds for $d_{\mathscr{H}_1}$, when applied to the case where $\nu \geq 1$ is an integer.

We will now connect the previous results to the main findings of [33]. To do this, we shall provide explicit estimates of the bounds appearing in Theorem 3.11, in the case where G belongs to a fixed Wiener chaos of *even order* q.

Proposition 3.13 *Let $q \geq 2$ be an even integer, and let $G = I_q(g)$, where $g \in \mathfrak{H}^{\odot q}$. Then,*

$$E[(2\nu + 2G - \langle DG, -DL^{-1}G\rangle_{\mathfrak{H}})^2] = E[(2\nu + 2G - q^{-1}\|DG\|^2_{\mathfrak{H}})^2] \tag{3.50}$$
$$\leq (2\nu - q!\,\|g\|^2_{\mathfrak{H}^{\otimes q}})^2 + +q^2 \sum_{\substack{r\in\{1,\ldots,q-1\}\\ r\neq q/2}} (2q-2r)!(r-1)!^2 \binom{q-1}{r-1}^4 \|$$
$$g \otimes_r g\|^2_{\mathfrak{H}^{\otimes 2(q-r)}} + +4q! \left\| c_q^{-1} \times g\widetilde{\otimes}_{q/2} g - g \right\|^2_{\mathfrak{H}^{\otimes q}},$$

where

$$c_q := \frac{1}{(q/2)!\binom{q-1}{q/2-1}^2} = \frac{4}{(q/2)!\binom{q}{q/2}^2}. \tag{3.51}$$

Proof By using (3.42) we deduce that

$$\begin{aligned} q^{-1}\|DG\|^2_{\mathfrak{H}} - 2\nu - 2G = (q!\,\|g\|^2_{\mathfrak{H}^{\otimes q}} - 2\nu) \\ + q \sum_{\substack{r\in\{1,\ldots,q-1\}\\ r\neq q/2}} (r-1)!\binom{q-1}{r-1}^2 I_{2(q-r)}\left(g\widetilde{\otimes}_r g\right) \\ + q(q/2-1)!\binom{q-1}{q/2-1} I_q(g\widetilde{\otimes}_{q/2} g - 2g). \end{aligned}$$

The conclusion is obtained by using the isometric properties of multiple Wiener-Itô integrals, as well as the relation $\|g\widetilde{\otimes}_r g\|_{\mathfrak{H}^{\otimes 2(q-r)}} \leq \|g \otimes_r g\|_{\mathfrak{H}^{\otimes 2(q-r)}}$, for every $r \in \{1, \ldots, q-1\}$ such that $r \neq q/2$. □

By using Proposition 3.13, we immediately recover the implications (iv) → (iii) → (i) in the statement of the following result, recently proved in [33, Theorem 1.2].

Theorem 3.14 [33] *Let $\nu > 0$ and let $F(\nu)$ have a centered Gamma distribution with parameter ν. Fix an* even *integer $q \geq 2$, and define c_q according to* (3.51). *Consider a sequence of the type $G_n = I_q(g_n)$, where $n \geq 1$ and $g_n \in \mathfrak{H}^{\odot q}$, and suppose that*

$$\lim_{n\to\infty} E\left[G_n^2\right] = \lim_{n\to\infty} q!\|g_n\|^2_{\mathfrak{H}^{\otimes q}} = 2\nu.$$

Then, the following four conditions are equivalent:

(i) *as $n \to \infty$, the sequence $(G_n)_{n\geq 1}$ converges in distribution to $F(\nu)$*;
(ii) $\lim_{n\to\infty} E[G_n^4] - 12E[G_n^3] = 12\nu^2 - 48\nu$;
(iii) *as $n \to \infty$, $\|DG_n\|^2_{\mathfrak{H}} - 2qG_n \longrightarrow 2q\nu$ in* L^2.
(iv) $\lim_{n\to\infty} \|g_n \widetilde{\otimes}_{q/2} g_n - c_q \times g_n\|_{\mathfrak{H}^{\otimes q}} = 0$, *where c_q is given by* (3.51), *and* $\lim_{n\to\infty} \|g_n \otimes_r g_n\|_{\mathfrak{H}^{\otimes 2(q-r)}} = 0$, *for every $r = 1, \ldots, q-1$ such that $r \neq q/2$.*

Observe that $E(F(\nu)^2) = 2\nu$, $E(F(\nu)^3) = 8\nu$ and $E(F(\nu)^4) = 48\nu + 12\nu^2$, so that the implication (ii) $\to$ (i) in the previous statement can be seen as a further simplification of the *method of moments and cumulants*, as applied to non-central limit theorems (see, e.g. [51], and the references therein, for a survey of classic non-central limit theorems). Also, the combination of Proposition 3.13 and Theorem 3.14 shows that, inside a fixed Wiener chaos of even order, one has that: (i) $d_{\mathscr{H}_2}$ metrizes the weak convergence towards centered Gamma distributions, and (ii) $d_{\mathscr{H}_1}$ metrizes the weak convergence towards centered χ^2 distributions with arbitrary degrees of freedom.

The following result concerns the Gamma approximation of a sum of two multiple integrals. Note, at the cost of a quite heavy notation, one could easily establish analogous estimates for sums of three or more integrals. The reader should compare this result with Proposition 3.7.

Proposition 3.15 *Fix two real numbers $\nu_1, \nu_2 > 0$, as well as two even integers $2 \leq q_1 < q_2$. Set $\nu = \nu_1 + \nu_2$ and suppose (for the sake of simplicity) that $q_2 > 2q_1$. Consider a sequence of the form*

$$Z_n = I_{q_1}(f_n^1) + I_{q_2}(f_n^2), \quad n \geq 1,$$

where $f_n^i \in \mathfrak{H}^{\odot q_i}$. Set

$$\mathscr{I} = \Big\{(i,j,r) \in \{1,2\}^2 \times \mathbb{N} : 1 \leq r \leq q_i \wedge q_j \quad \text{and, whenever } i = j, \\ r \neq q_i \quad \text{and} \quad r \neq \tfrac{q_i}{2}\Big\}.$$

Then

$$E[(2Z_n + 2\nu - \langle DZ_n, -DL^{-1}Z_n\rangle_{\mathfrak{H}})^2] \\ \leq 3\left(2\nu - \sum_{i=1,2} q_i!\|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}}\right)^2 + 24\sum_{i=1,2} c_{q_i}^{-2} q_i!\,\|f_n^i \widetilde{\otimes}_{q_i/2} f_n^i - c_{q_i} \times f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}}$$

$$+ 12 \sum_{(i,j,r)\in\mathscr{J}} q_i^2 (r-1)!^2 \binom{q_i-1}{r-1}^2 \binom{q_j-1}{r-1}^2 (q_i+q_j-2r)! \tag{3.52}$$
$$\times \|f_n^i \otimes_{q_i-r} f_n^i\|_{\mathfrak{H}^{\otimes 2r}} \|f_n^j \otimes_{q_j-r} f_n^j\|_{\mathfrak{H}^{\otimes 2r}}.$$

In particular, if

(i) $E[Z_n^2] = \sum_{i=1,2} q_i! \|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}} \longrightarrow 2\nu$ *as* $n \to \infty$*,*
(ii) *for* $i = 1, 2$, $\|f_n^i \widetilde{\otimes}_{q_i/2} f_n^i - c_{q_i} \times f_n^i\|_{\mathfrak{H}^{\otimes q_i}} \longrightarrow 0$ *as* $n \to \infty$*, where* c_{q_i} *is defined in Theorem 3.14,*
(iii) *for any* $i = 1, 2$ *and* $r = 1, \ldots, q_i - 1$ *such that* $r \neq \frac{q_i}{2}$, $\|f_n^i \otimes_r f_n^i\|_{\mathfrak{H}^{\otimes 2(q_i-r)}} \longrightarrow 0$ *as* $n \to \infty$*,*

then $Z_n \overset{\text{Law}}{\longrightarrow} F(\nu)$ *as* $n \to \infty$*, and the combination of Theorem 3.1 and (3.52) allows to associate explicit bounds with this convergence.*

Proof of Proposition 3.15 We have (see the proof of Proposition 3.7)

$$\begin{aligned}
&\langle DZ_n, -DL^{-1}Z_n\rangle_{\mathfrak{H}} - 2Z_n - 2\nu \\
&\quad = \left(\sum_{i=1,2} q_i! \|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}} - 2\nu\right) + \sum_{i=1,2} 2\, c_{q_i}^{-1} I_{q_i}(f_n^i \widetilde{\otimes}_{q_i/2} f_n^i - c_{q_i} \times f_n^i) \\
&\qquad + \sum_{(i,j,r)\in\mathscr{J}} q_i (r-1)! \binom{q_i-1}{r-1}\binom{q_j-1}{r-1} I_{q_i+q_j-2r}\left(f_n^i \otimes_r f_n^j\right).
\end{aligned}$$

Thus

$$\begin{aligned}
&E\left([\langle DZ_n, -DL^{-1}Z_n\rangle_{\mathfrak{H}} - 2Z_n - 2\nu]^2\right) \\
&\leq 3\left(2\nu - \sum_{i=1,2} q_i! \|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}}\right)^2 + 24 \sum_{i=1,2} c_{q_i}^{-2} q_i! \,\|f_n^i \widetilde{\otimes}_{q_i/2} f_n^i - c_{q_i} \times f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}} \\
&\quad + 3\, E\left(\sum_{(i,j,r)\in\mathscr{J}} q_i (r-1)! \binom{q_i-1}{r-1}\binom{q_j-1}{r-1} I_{q_i+q_j-2r}\left(f_n^i \otimes_r f_n^j\right)\right)^2 \\
&\leq 3\left(2\nu - \sum_{i=1,2} q_i! \|f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}}\right)^2 + 24 \sum_{i=1,2} c_{q_i}^{-2} q_i! \,\|f_n^i \widetilde{\otimes}_{q_i/2} f_n^i - c_{q_i} \times f_n^i\|^2_{\mathfrak{H}^{\otimes q_i}} \\
&\quad + 12 \sum_{(i,j,r)\in\mathscr{J}} q_i^2 (r-1)!^2 \binom{q_i-1}{r-1}^2 \\
&\quad \binom{q_j-1}{r-1}^2 (q_i+q_j-2r)! \|f_n^i \otimes_r f_n^j\|^2_{\mathfrak{H}^{\otimes q_i+q_j-2r}}
\end{aligned}$$

$$\leq 3 \left(2\nu - \sum_{i=1,2} q_i! \| f_n^i \|^2_{\mathfrak{H}^{\otimes q_i}} \right)^2 + 24 \sum_{i=1,2} c_{q_i}^{-2} q_i! \, \| f_n^i \widetilde{\otimes}_{q_i/2} f_n^i - c_{q_i} \times f_n^i \|^2_{\mathfrak{H}^{\otimes q_i}}$$

$$+ 12 \sum_{(i,j,r) \in \mathscr{J}} q_i^2 (r-1)!^2 \binom{q_i - 1}{r-1}^2 \binom{q_j - 1}{r-1}^2 (q_i + q_j - 2r)!$$

$$\times \| f_n^i \otimes_{q_i - r} f_n^i \|_{\mathfrak{H}^{\otimes 2r}} \| f_n^j \otimes_{q_j - r} f_n^j \|_{\mathfrak{H}^{\otimes 2r}}.$$

□

4 Berry–Esséen bounds in the Breuer–Major CLT

In this section, we use our main results in order to derive an explicit Berry–Esséen bound for the celebrated *Breuer–Major CLT* for Gaussian-subordinated random sequences. For simplicity, we focus on sequences that can be represented as Hermite-type functions of the (normalized) increments of a fractional Brownian motion. Our framework include examples of Gaussian sequences whose autocovariance functions display long dependence. Plainly, the techniques developed in this paper can also accommodate the analysis of more general transformations (for instance, obtained from functions with an arbitrary *Hermite rank*—see [52]), as well as alternative covariance structures.

4.1 General setup

We recall that a *fractional Brownian motion* (fBm) $B = \{B_t : t \in [0,1]\}$, of Hurst index $H \in (0,1)$, is a centered Gaussian process, started from zero and with covariance function $E(B_s B_t) = R_H(s,t)$, where

$$R_H(s,t) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H} \right); \quad s,t \in [0,1].$$

If $H = 1/2$, then $R_H(s,t) = \min(s,t)$ and B is a standard Brownian motion. For any choice of the Hurst parameter $H \in (0,1)$, the Gaussian space generated by B can be identified with an isonormal Gaussian process of the type $X = \{X(h) : h \in \mathfrak{H}\}$, where the real and separable Hilbert space $\mathfrak{H}$ is defined as follows: (i) denote by $\mathscr{E}$ the set of all $\mathbb{R}$-valued step functions on $[0,1]$, (ii) define $\mathfrak{H}$ as the Hilbert space obtained by closing $\mathscr{E}$ with respect to the scalar product

$$\left\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \right\rangle_{\mathfrak{H}} = R_H(t,s).$$

In particular, with such a notation one has that $B_t = X(\mathbf{1}_{[0,t]})$. Note that, if $H = 1/2$, then $\mathfrak{H} = L^2[0,1]$; when $H > 1/2$, the space $\mathfrak{H}$ coincides with the space of distributions f such that $s^{\frac{1}{2}-H} \mathscr{I}_{0+}^{H-\frac{1}{2}} (f(u) u^{H-\frac{1}{2}})(s)$ belongs to $L^2[0,1]$; when $H < 1/2$ one has that $\mathfrak{H}$ is $\mathscr{I}_{0+}^{H-\frac{1}{2}}(L^2[0,1])$. Here, $\mathscr{I}_{0+}^{H-\frac{1}{2}}$ denotes the action of the *fractional Riemann-Liouville operator*, defined as

$$\mathscr{I}_{0+}^{H-\frac{1}{2}} f(x) = \frac{1}{\Gamma(H-\frac{1}{2})} \int_0^x (x-y)^{H-\frac{3}{2}} f(y) dy.$$

The reader is referred e.g. to [35] for more details on fBm and fractional operators.

4.2 A Berry–Esséen bound

In what follows, we will be interested in the asymptotic behaviour (as $n \to \infty$) of random vectors that are subordinated to the array

$$V_{n,H} = \left\{ n^H (B_{(k+1)/n} - B_{k/n}) : k = 0, \ldots, n-1 \right\}, \quad n \geq 1. \tag{4.53}$$

Note that, for every $n \geq 1$, the law of $V_{n,H}$ in (4.53) coincides with the law of the first n instants of a centered stationary Gaussian sequence indexed by $\{0, 1, 2, \ldots\}$ and with autocovariance function given by

$$\rho_H(k) = \frac{1}{2} \left(|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H} \right), \quad k \in \mathbb{Z}$$

(in particular, $\rho_H(0) = 1$ and $\rho_H(k) = \rho_H(-k)$). From this last expression, one deduces that the components of the vector $V_{n,H}$ are: (a) i.i.d. for $H = 1/2$, (b) negatively correlated for $H \in (0, 1/2)$ and (c) positively correlated for $H \in (1/2, 1)$. In particular, if $H \in (1/2, 1)$, then $\sum_k \rho_H(k) = +\infty$: in this case, one customarily says that ρ_H exhibits *long-range dependence* (or, equivalently, *long memory*—see, e.g. [53] for a general discussion of this point).

Now denote by H_q, $q \geq 2$, the qth Hermite polynomial, defined as

$$H_q(x) = \frac{(-1)^q}{q!} e^{\frac{x^2}{2}} \frac{d^q}{dx^q} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

For instance, $H_2(x) = (x^2 - 1)/2$, $H_3(x) = (x^3 - 3x)/6$, and so on. Finally, set

$$\sigma = \sqrt{\frac{1}{q!} \sum_{t \in \mathbb{Z}} \rho_H(t)^q},$$

and define

$$Z_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=0}^{n-1} H_q\left(n^H (B_{(k+1)/n} - B_{k/n}) \right) = \frac{n^{qH-\frac{1}{2}}}{q!\sigma} \sum_{k=0}^{n-1} I_q(\delta_{k/n}^{\otimes q}), \tag{4.54}$$

where I_q denotes the qth multiple integral with respect to the isonormal process associated with B (see Sect. 2). For simplicity, here (and for the rest of this section) we write $\delta_{k/n}$ instead of $\mathbf{1}_{[k/n,(k+1)/n]}$, and also $\delta_{k/n}^{\otimes q} = \delta_{k/n} \otimes \cdots \otimes \delta_{k/n}$ (q times). Note that in (4.54) we have used the standard relation: $q!H_q\left(B(h)\right) = I_q(h^{\otimes q})$ for every $h \in \mathfrak{H}$ such that $\|h\|_{\mathfrak{H}} = 1$ (see, e.g. [35, Chap. 1]).

Now observe that, for every $q \geq 2$, one has that $\sum_t |\rho_H(t)|^q < \infty$ if, and only if, $H \in \left(0, \frac{2q-1}{2q}\right)$. Moreover, in this case, $E(Z_n^2) \to 1$ as $n \to \infty$. As a consequence, according to Breuer and Major's well-known result [5, Theorem 1], as $n \to \infty$

$$Z_n \to Z \sim \mathscr{N}(0,1) \quad \text{in distribution.}$$

To the authors' knowledge, the following statement contains the first Berry–Esséen bound ever proved for the Breuer–Major CLT:

Theorem 4.1 *As $n \to \infty$, Z_n converges in law towards $Z \sim \mathscr{N}(0,1)$. Moreover, there exists a constant c_H, depending uniquely on H, such that, for any $n \geq 1$:*

$$\sup_{z\in\mathbb{R}} |P(Z_n \leq z) - P(Z \leq z)| \leq c_H \times \begin{cases} n^{-\frac{1}{2}} & \text{if } H \in (0, \frac{1}{2}] \\ n^{H-1} & \text{if } H \in [\frac{1}{2}, \frac{2q-3}{2q-2}] \\ n^{qH-q+\frac{1}{2}} & \text{if } H \in [\frac{2q-3}{2q-2}, \frac{2q-1}{2q}) \end{cases}$$

Remark 4.2
1. Theorem 1.6 (see the Introduction) can be proved by simply setting $q = 2$ in Theorem 4.1. Observe that in this case one has $\frac{2q-3}{2q-2} = \frac{1}{2}$, so that the middle line in the previous display becomes immaterial.
2. When $H > \frac{2q-1}{2q}$, the sequence Z_n does not converge in law towards a Gaussian random variable. Indeed, in this case a non-central limit theorem takes place. See Breton and Nourdin [4] for bounds associated with this convergence.
3. As discussed in [5, p. 429], it is in general not possible to derive CLTs such as the one in Theorem 4.1 from mixing-type conditions. In particular, it seems unfeasible to deduce Theorem 4.1 from any mixing characterization of the increments of fractional Brownian motion (as the one proved e.g. by Picard in [40, Theorem A.1]). See, e.g. Tikhomirov [54] for general derivations of Berry–Esséen bounds from strong mixing conditions.

4.3 Proof of Theorem 4.1

We have

$$DZ_n = \frac{n^{qH-\frac{1}{2}}}{(q-1)!\sigma} \sum_{k=0}^{n-1} I_{q-1}(\delta_{k/n}^{\otimes q-1})\delta_{k/n},$$

hence

$$\|DZ_n\|_{\mathfrak{H}}^2 = \frac{n^{2qH-1}}{(q-1)!^2\sigma^2} \sum_{k,\ell=0}^{n-1} I_{q-1}(\delta_{k/n}^{\otimes q-1}) I_{q-1}(\delta_{\ell/n}^{\otimes q-1}) \langle \delta_{k/n}, \delta_{\ell/n}\rangle_{\mathfrak{H}}.$$

By the multiplication formula (2.29):

$$I_{q-1}(\delta_{k/n}^{\otimes q-1}) I_{q-1}(\delta_{\ell/n}^{\otimes q-1}) = \sum_{r=0}^{q-1} r! \binom{q-1}{r}^2 I_{2q-2-2r}\left(\delta_{k/n}^{\otimes q-1-r} \widetilde{\otimes} \delta_{\ell/n}^{q-1-r}\right) \langle \delta_{k/n}, \delta_{\ell/n}\rangle_{\mathfrak{H}}^r.$$

Consequently,

$$\|DZ_n\|_{\mathfrak{H}}^2 = \frac{n^{2qH-1}}{(q-1)!^2\sigma^2} \sum_{r=0}^{q-1} r! \binom{q-1}{r}^2 \sum_{k,\ell=0}^{n-1} I_{2q-2-2r} \left(\delta_{k/n}^{\otimes q-1-r} \widetilde{\otimes} \delta_{\ell/n}^{q-1-r}\right) \langle \delta_{k/n}, \delta_{\ell/n}\rangle_{\mathfrak{H}}^{r+1}.$$

Thus, we can write

$$\frac{1}{q}\|DZ_n\|_{\mathfrak{H}}^2 - 1 = \sum_{r=0}^{q-1} A_r(n) - 1$$

where

$$A_r(n) = \frac{r!\binom{q-1}{r}^2}{q(q-1)!^2\sigma^2} n^{2qH-1} \sum_{k,\ell=0}^{n-1} I_{2q-2-2r}\left(\delta_{k/n}^{\otimes q-1-r} \widetilde{\otimes} \delta_{\ell/n}^{q-1-r}\right) \langle \delta_{k/n}, \delta_{\ell/n}\rangle_{\mathfrak{H}}^{r+1}.$$

We will need the following easy Lemma (the proof is omitted). Here and for the rest of the proof of Theorem 4.1, the notation $a_n \trianglelefteq b_n$ means that $\sup_{n\geq 1} |a_n|/|b_n| < \infty$.

Lemma 4.3 1. *We have* $\rho_H(n) \trianglelefteq |n|^{2H-2}$.

2. *For any* $\alpha \in \mathbb{R}$, *we have*

$$\sum_{k=1}^{n-1} k^\alpha \trianglelefteq 1 + n^{\alpha+1}.$$

3. *If* $\alpha \in (-\infty, -1)$, *we have*

$$\sum_{k=n}^{\infty} k^\alpha \trianglelefteq n^{\alpha+1}.$$

By using elementary computations (in particular, observe that $n^{2H}\langle\delta_{k/n},\delta_{\ell/n}\rangle_{\mathfrak{H}}=\rho_H(k-\ell)$) and then Lemma 4.3, it is easy to check that

$$\begin{aligned}
A_{q-1}(n)-1&=\frac{1}{q!\sigma^2}n^{2qH-1}\sum_{k,\ell=0}^{n-1}\langle\delta_{k/n},\delta_{\ell/n}\rangle_{\mathfrak{H}}^q-1\\
&=\frac{1}{q!\sigma^2}\left(\frac{1}{n}\sum_{k,\ell=0}^{n-1}\rho_H(k-\ell)^q-\sum_{t\in\mathbb{Z}}\rho_H(t)^q\right)\\
&=\frac{1}{q!\sigma^2}\left(\frac{1}{n}\sum_{|t|<n}(n-|t|)\rho_H(t)^q-\sum_{t\in\mathbb{Z}}\rho_H(t)^q\right)\\
&=\frac{1}{q!\sigma^2}\left(-\frac{1}{n}\sum_{|t|<n}|t|\,\rho_H(t)^q-\sum_{|t|\geq n}\rho_H(t)^q\right)\\
&\trianglelefteq\frac{1}{n}\sum_{t=1}^{n-1}t^{2qH-2q+1}+\sum_{t=n}^{\infty}t^{2qH-2q}\trianglelefteq n^{-1}+n^{2qH-2q+1}.
\end{aligned}$$

Now, we assume that $r\leq q-2$ is fixed. We have

$$\begin{aligned}
E|A_r(n)|^2&=c(H,r,q)n^{4qH-2}\sum_{i,j,k,\ell=0}^{n-1}\langle\delta_{k/n},\delta_{\ell/n}\rangle_{\mathfrak{H}}^{r+1}\langle\delta_{i/n},\delta_{j/n}\rangle_{\mathfrak{H}}^{r+1}\\
&\quad\times\langle\delta_{k/n}^{\otimes q-1-r}\widetilde{\otimes}\delta_{\ell/n}^{q-1-r},\delta_{i/n}^{\otimes q-1-r}\widetilde{\otimes}\delta_{j/n}^{q-1-r}\rangle_{\mathfrak{H}^{\otimes 2q-2-2r}}\\
&=\sum_{\substack{\alpha,\beta\geq 0\\ \alpha+\beta=q-r-1}}\ \sum_{\substack{\gamma,\delta\geq 0\\ \gamma+\delta=q-r-1}}c(H,r,q,\alpha,\beta,\gamma,\delta)\,B_{r,\alpha,\beta,\gamma,\delta}(n)
\end{aligned}$$

where $c(\cdot)$ denotes a generic constant depending only on the objects inside its argument (and which can be equal to zero), and

$$\begin{aligned}
B_{r,\alpha,\beta,\gamma,\delta}(n)&=n^{4qH-2}\sum_{i,j,k,\ell=0}^{n-1}\langle\delta_{k/n},\delta_{\ell/n}\rangle_{\mathfrak{H}}^{r+1}\langle\delta_{i/n},\delta_{j/n}\rangle_{\mathfrak{H}}^{r+1}\langle\delta_{k/n},\delta_{i/n}\rangle_{\mathfrak{H}}^{\alpha}\\
&\quad\times\langle\delta_{k/n},\delta_{j/n}\rangle_{\mathfrak{H}}^{\beta}\langle\delta_{\ell/n},\delta_{i/n}\rangle_{\mathfrak{H}}^{\gamma}\langle\delta_{\ell/n},\delta_{j/n}\rangle_{\mathfrak{H}}^{\delta}\\
&=n^{-2}\sum_{i,j,k,\ell=0}^{n-1}\rho_H(k-\ell)^{r+1}\rho_H(i-j)^{r+1}\rho_H(k-i)^{\alpha}\\
&\quad\times\rho_H(k-j)^{\beta}\rho_H(\ell-i)^{\gamma}\rho_H(\ell-j)^{\delta}.
\end{aligned}$$

When $\alpha,\beta,\gamma,\delta$ are fixed, we can decompose the sum $\sum_{i,j,k,\ell}$ appearing in $B_{r,\alpha,\beta,\gamma,\delta}(n)$ just above, as follows:

$$\sum_{i=j=k=\ell} + \left(\sum_{\substack{i=j=k\\ \ell\neq i}} + \sum_{\substack{i=j=\ell\\ k\neq i}} + \sum_{\substack{i=k=\ell\\ j\neq i}} + \sum_{\substack{j=k=\ell\\ i\neq j}} \right) + \left(\sum_{\substack{i=j,k=\ell\\ k\neq i}} + \sum_{\substack{i=k,j=\ell\\ j\neq i}} + \sum_{\substack{i=\ell,j=k\\ j\neq i}} \right)$$
$$+ \left(\sum_{\substack{i=j,k\neq i\\ k\neq \ell,\ell\neq i}} + \sum_{\substack{i=k,j\neq i\\ j\neq \ell,k\neq \ell}} + \sum_{\substack{i=\ell,k\neq i\\ k\neq j,j\neq i}} + \sum_{\substack{j=k,k\neq i\\ k\neq \ell,\ell\neq i}} + \sum_{\substack{j=\ell,k\neq i\\ k\neq \ell,\ell\neq i}} + \sum_{\substack{k=\ell,k\neq i\\ k\neq j,j\neq i}} \right) + \sum_{\substack{i,j,k,\ell\\ \text{are all different}}}$$

(all these sums must be understood as being defined over indices $\{i, j, k, \ell\} \in \{0, \dots, n-1\}^4$). Now, we will deal with each of these fifteen sums separately.

The first sum is particularly easy to handle: indeed, it is immediately checked that

$$n^{-2} \sum_{i=j=k=\ell} \rho_H(k-\ell)^{r+1} \rho_H(i-j)^{r+1} \rho_H(k-i)^{\alpha} \rho_H(k-j)^{\beta}$$
$$\times \rho_H(\ell-i)^{\gamma} \rho_H(\ell-j)^{\delta} \trianglelefteq n^{-1}.$$

For the second sum, one can write

$$n^{-2} \sum_{\substack{i=j=k\\ \ell\neq i}} \rho_H(k-\ell)^{r+1} \rho_H(i-j)^{r+1} \rho_H(k-i)^{\alpha} \rho_H(k-j)^{\beta}$$
$$\times \rho_H(\ell-i)^{\gamma} \rho_H(\ell-j)^{\delta}$$
$$\trianglelefteq n^{-2} \sum_{i\neq \ell} \rho_H(\ell-i)^q \trianglelefteq n^{-1} \sum_{\ell=1}^{n-1} \ell^{2qH-2q} = n^{-1} + n^{2qH-2q} \quad \text{by Lemma 4.3.}$$

For the third sum, we can proceed analogously and we again obtain the bound $n^{-1} + n^{2qH-2q}$.

For the fourth sum, we write

$$n^{-2} \sum_{\substack{i=k=\ell\\ j\neq i}} \rho_H(k-\ell)^{r+1} \rho_H(i-j)^{r+1} \rho_H(k-i)^{\alpha} \rho_H(k-j)^{\beta}$$
$$\times \rho_H(\ell-i)^{\gamma} \rho_H(\ell-j)^{\delta}$$
$$\trianglelefteq n^{-2} \sum_{i\neq j} \rho_H(j-i)^{r+1+\beta+\delta} \trianglelefteq n^{-2} \sum_{i\neq j} |j-i|^{(r+1+\beta+\delta)(2H-2)}$$
$$\trianglelefteq n^{-2} \sum_{i\neq j} |j-i|^{2H-2}$$
$$\trianglelefteq n^{-1} \sum_{j=1}^{n-1} j^{2H-2} \trianglelefteq n^{-1} + n^{2H-2}$$

(we used the fact that $r+1+\beta+\delta \geq 1$ since $r, \beta, \delta \geq 0$). For the fifth sum, we can proceed analogously and we again obtain the bound $n^{-1} + n^{2H-2}$.

For the sixth sum, we have

$$\begin{aligned}
&n^{-2}\sum_{\substack{i=j\\k=\ell\\k\neq i}}\rho_H(k-\ell)^{r+1}\rho_H(i-j)^{r+1}\rho_H(k-i)^{\alpha}\rho_H(k-j)^{\beta}\\
&\rho_H(\ell-i)^{\gamma}\rho_H(\ell-j)^{\delta}\\
&\trianglelefteq n^{-2}\sum_{k\neq i}\rho_H(k-i)^{2q-2-2r}\trianglelefteq n^{-2}\sum_{k\neq i}|k-i|^{(2q-2-2r)(2H-2)}\\
&\quad\times\trianglelefteq n^{-2}\sum_{k\neq i}|k-i|^{4H-4}\\
&\trianglelefteq n^{-1}\sum_{k=1}^{n-1}k^{4H-4}\trianglelefteq n^{-1}+n^{4H-4}
\end{aligned}$$

(here, we used $r\leq q-2$). For the seventh and the eighth sums, we can proceed analogously and we also obtain $n^{-1}+n^{4H-4}$ for bound.

For the ninth sum, we have

$$\begin{aligned}
&n^{-2}\sum_{\substack{i=j,k\neq i\\k\neq\ell,\ell\neq i}}\rho_H(k-\ell)^{r+1}\rho_H(i-j)^{r+1}\rho_H(k-i)^{\alpha}\rho_H(k-j)^{\beta}\\
&\rho_H(\ell-i)^{\gamma}\rho_H(\ell-j)^{\delta}\\
&\trianglelefteq n^{-2}\sum_{\substack{k\neq i\\k\neq\ell,\ell\neq i}}\rho_H(k-\ell)^{r+1}\rho_H(k-i)^{q-r-1}\rho_H(\ell-i)^{q-r-1}.
\end{aligned}$$

Now, let us decompose the sum $\sum_{k\neq i,k\neq\ell,\ell\neq i}$ into

$$\sum_{k>\ell>i}+\sum_{k>i>\ell}+\sum_{\ell>i>k}+\sum_{\ell>k>i}+\sum_{i>\ell>k}+\sum_{i>k>\ell}.$$

For the first term (for instance), we have

$$\begin{aligned}
&n^{-2}\sum_{k>\ell>i}\rho_H(k-\ell)^{r+1}\rho_H(k-i)^{q-r-1}\rho_H(\ell-i)^{q-r-1}\\
&\trianglelefteq n^{-2}\sum_{k>\ell>i}(k-\ell)^{(r+1)(2H-2)}(k-i)^{(q-r-1)(2H-2)}(\ell-i)^{(q-r-1)(2H-2)}\\
&\trianglelefteq n^{-2}\sum_{k>\ell>i}(k-\ell)^{q(2H-2)}(\ell-i)^{(q-r-1)(2H-2)}\quad\text{since } k-i>k-\ell\\
&=n^{-2}\sum_{k}\sum_{\ell<k}(k-\ell)^{q(2H-2)}\sum_{i<\ell}(\ell-i)^{(q-r-1)(2H-2)}\\
&\trianglelefteq n^{-2}\sum_{k}\sum_{\ell<k}(k-\ell)^{q(2H-2)}\sum_{i<\ell}(\ell-i)^{2H-2}\quad\text{since } q-r-1\geq 1
\end{aligned}$$

$$\trianglelefteq n^{-1}\sum_{\ell=1}^{n-1}\ell^{2qH-2q}\sum_{i=1}^{n-1}i^{2H-2}$$
$$\trianglelefteq n^{-1}(1+n^{2qH-2q+1})(1+n^{2H-1}) \trianglelefteq n^{-1}+n^{2H-2} \quad \text{since } 2qH-2q+1<0.$$

We obtain the same bound for the other terms. By proceeding in the same way than for the ninth term, we also obtain the bound $n^{-1}+n^{2H-2}$ for the tenth, eleventh, twelfth, thirteenth and fourteenth terms.

For the fifteenth (and last!) sum, we decompose $\sum_{(i,j,k,\ell \text{ are all different})}$ as follows

$$\sum_{k>\ell>i>j} + \sum_{k>\ell>j>i} + \cdots . \tag{4.55}$$

For the first term, we have:

$$n^{-2}\sum_{k>\ell>i>j}\rho_H(k-\ell)^{r+1}\rho_H(i-j)^{r+1}\rho_H(k-i)^{\alpha}$$
$$\rho_H(k-j)^{\beta}\rho_H(\ell-i)^{\gamma}\rho_H(\ell-j)^{\delta}$$
$$\trianglelefteq n^{-2}\sum_{k>\ell>i>j}(k-\ell)^{q(2H-2)}(i-j)^{(r+1)(2H-2)}(\ell-i)^{(q-r-1)(2H-2)}$$
$$= n^{-2}\sum_{k}\sum_{\ell<k}(k-\ell)^{q(2H-2)}\sum_{i<\ell}(\ell-i)^{(q-r-1)(2H-2)}\sum_{j<i}(i-j)^{(r+1)(2H-2)}$$
$$\trianglelefteq n^{-1}\sum_{\ell=1}^{n-1}\ell^{q(2H-2)}\sum_{i=1}^{n-1}i^{(q-r-1)(2H-2)}\sum_{j=1}^{n-1}j^{(r+1)(2H-2)}$$
$$\trianglelefteq n^{-1}(1+n^{2qH-2q+1})(1+n^{(q-r-1)(2H-2)+1})(1+n^{(r+1)(2H-2)+1})$$
$$\trianglelefteq n^{-1}(1+n^{2H-1}+n^{2qH-2q+2})$$
$$\text{since } 2qH-2q+1<0 \quad \text{and} \quad r+1, q-r-1\geq 1$$
$$\trianglelefteq n^{-1}+n^{2H-2}+n^{2qH-2q+1}.$$

The same bound also holds for the other terms in (4.55). By combining all these bounds, we obtain

$$\max_{r=1,\ldots,q-1} E|A_r(n)|^2 \trianglelefteq n^{-1}+n^{2H-2}+n^{2qH-2q+1},$$

that finally gives:

$$E\left(\frac{1}{q}\|DZ_n\|_{\mathfrak{H}}^2-1\right)^2 \trianglelefteq n^{-1}+n^{2H-2}+n^{2qH-2q+1}.$$

The proof of Theorem 4.1 is now completed by means of Proposition 3.2.

5 Some remarks about χ^2 approximations

The following statement illustrates a natural application of the results about χ^2 approximations (as discussed in Sect. 3.3) in order to obtain upper bounds in non-central limit theorems for multiple integrals. Observe that we focus on double integrals but, at the cost of some heavy notation, everything can be straightforwardly extended to the case of integrals of any order $q \geq 2$. Recall that the class of functions $\mathscr{H}_1$ is defined in formula (1.14).

Proposition 5.1 *Let $F_n = I_2(f_n)$, $n \geq 1$, where $f_n \in \mathfrak{H}^{\odot 2}$, be a sequence of double Wiener-Itô integrals. Suppose that $E(F_n^2) \longrightarrow 1$ and $\|f_n \otimes_1 f_n\|_{\mathfrak{H}^{\otimes 2}} \longrightarrow 0$ as $n \to \infty$. Then, by defining*

$$H_n = I_4\left(f_n \widetilde{\otimes} f_n\right), \quad n \geq 1,$$

one has that

$$E\left[\left(2 + 2H_n - \frac{1}{4}\|DH_n\|^2_{\mathfrak{H}}\right)^2\right] \longrightarrow 0, \quad \text{as } n \to \infty, \tag{5.56}$$

and

$$d_{\mathscr{H}_1}(F_n^2 - 1, N^2 - 1) \leq 8\sqrt{2}\|f_n \otimes_1 f_n\|_{\mathfrak{H}^{\otimes 2}} + \sqrt{2\pi\, E\left[\left(2 + 2H_n - \frac{1}{4}\|DH_n\|^2_{\mathfrak{H}}\right)^2\right]}, \tag{5.57}$$

where $N \sim \mathscr{N}(0, 1)$.

Proof First, we have that $F_n \overset{\text{Law}}{\longrightarrow} N$ by Theorem 3.3. Now, use the multiplication formula (2.29) to deduce that

$$F_n^2 - 1 = 8\, I_2(f_n \otimes_1 f_n) + H_n.$$

Since

$$E\left(I_2(f_n \otimes_1 f_n)^2\right) = 2\,\|f_n \otimes_1 f_n\|^2_{\mathfrak{H}^{\otimes 2}} \longrightarrow 0, \quad \text{as } n \to \infty,$$

we infer that $H_n \overset{\text{Law}}{\longrightarrow} N^2 - 1$, and therefore that (5.56) must take place, due to Theorem 3.14. By the definition of the class $\mathscr{H}_1$, one also deduces that

$$d_{\mathscr{H}_1}(F_n^2 - 1, N^2 - 1) \leq d_{\mathscr{H}_1}(H_n, N^2 - 1) + 8\, E\,|I_2(f_n \otimes_1 f_n)|\,.$$

The final result is obtained by combining (3.49)–(3.50) with the relation

$$E|I_2(f_n \otimes_1 f_n)| \leq \sqrt{E\left(I_2(f_n \otimes_1 f_n)^2\right)}.$$

□

We conclude this section with a simple example, showing how one can apply our techniques to deduce bounds in a non-central limit theorem, involving quadratic functionals of i.i.d. Gaussian random variables.

Example Let $(G_k)_{k\geq 0}$ be a sequence of centered i.i.d. standard Gaussian random variables. Also, let $(a_k)_{k\in\mathbb{Z}}$ be a sequence of real numbers such that

$$a(0) = 1, \quad a(r) = a(-r),\ r \in \mathbb{Z}, \quad \text{and} \quad \sum_{r\in\mathbb{Z}} |a(r) - 1| < \infty.$$

In particular, this implies that a is bounded (say, by $\|a\|_\infty$). Set

$$F_n = \frac{1}{n} \sum_{k,l=0}^{n-1} a(k-l)\left(G_k G_l - \delta_{kl}\right), \quad n \geq 1,$$

where δ_{kl} denotes the Kronecker symbol. We claim that $F_n \underset{n\to\infty}{\overset{\text{Law}}{\longrightarrow}} N^2 - 1$ with $N \sim \mathscr{N}(0,1)$, and our aim is to associate a bound with this convergence. Observe first that, without loss of generality, we can assume that $G_k = B_{k+1} - B_k$ where B is a standard Brownian motion [that B can be therefore regarded as an isonormal process over $\mathfrak{H} = \mathrm{L}^2(\mathbb{R}_+, dx)$]. We then have

$$DF_n = \frac{1}{n} \sum_{k,l=0}^{n-1} a(k-l)\left(G_k \mathbf{1}_{[l,l+1]} + G_l \mathbf{1}_{[k,k+1]}\right)$$

so that

$$\begin{aligned}
\|DF_n\|_{L^2}^2 &= \frac{1}{n^2} \sum_{i,j,k,l=0}^{n-1} a(k-l)a(i-j)\left\langle G_k \mathbf{1}_{[l,l+1]} + G_l \mathbf{1}_{[k,k+1]},\right. \\
&\quad \times \left. G_i \mathbf{1}_{[j,j+1]} + G_j \mathbf{1}_{[i,i+1]}\right\rangle_{L^2} \\
&= \frac{1}{n^2} \sum_{i,j,k,l=0}^{n-1} a(k-l)a(i-j) \\
&\quad \times \left(G_k G_i \delta_{lj} + G_k G_j \delta_{li} + G_l G_i \delta_{kj} + G_l G_j \delta_{ik}\right) \\
&= \frac{4}{n} \sum_{k,l=0}^{n-1} G_k G_l + \frac{1}{n^2} \sum_{i,j,k,l=0}^{n-1} \left(a(k-l)a(i-j) - 1\right) \\
&\quad \times \left(G_k G_i \delta_{lj} + G_k G_j \delta_{li} + G_l G_i \delta_{kj} + G_l G_j \delta_{ik}\right).
\end{aligned}$$

Hence,

$$\frac{1}{2}\|DF_n\|_{L^2}^2 - 2F_n - 2 = A_n + B_n$$

with

$$A_n = \frac{2}{n} \sum_{k,l=0}^{n-1} (1 - a(k-l))\, G_k G_l$$

$$B_n = \frac{1}{2n^2} \sum_{i,j,k,l=0}^{n-1} (a(k-l)a(i-j) - 1) \times \left(G_k G_i \delta_{lj} + G_k G_j \delta_{li} + G_l G_i \delta_{kj} + G_l G_j \delta_{ik} \right).$$

We have

$$E(A_n^2) = \frac{4}{n^2} \sum_{i,j,k,l=0}^{n-1} (1 - a(k-l))\, (1 - a(i-j))\, E\left(G_k G_l G_i G_j\right)$$
$$= \frac{8}{n^2} \sum_{i,k=0}^{n-1} (1 - a(k-i))^2 \le \frac{8}{n} (1 + \|a\|_\infty) \sum_{r \in \mathbb{Z}} |1 - a(r)| = O(1/n).$$

On the other hand, we have

$$B_n = B_n^1 + B_n^2 + B_n^3 + B_n^4$$

with

$$B_n^1 = \frac{1}{2n^2} \sum_{i,j,k,l=0}^{n-1} (a(k-l)a(i-j) - 1)\ G_k G_i \delta_{lj}$$

and similar computations hold for the other terms. Observe that

$$B_n^1 = \frac{1}{2n^2} \sum_{i,j,k=0}^{n-1} (a(k-j)a(i-j) - 1)\ G_k G_i$$
$$= \frac{1}{2n^2} \sum_{i,k=0}^{n-1} \alpha_{ki}\ G_k G_i, \quad \text{with } \alpha_{ki} = \sum_{j=0}^{n-1} (a(k-j)a(i-j) - 1)\,.$$

We have

$$|\alpha_{ki}| = \left| \sum_{j=0}^{n-1} (a(k-j) - 1) + \sum_{j=0}^{n-1} a(k-j)\, (a(i-j) - 1) \right|$$
$$\le (1 + \|a\|_\infty) \sum_{r \in \mathbb{Z}} |a(r) - 1|.$$

Consequently

$$E\left|B_n^1\right|^2 = \frac{1}{4n^4}\sum_{i,j,k,l=0}^{n-1} \alpha_{ki}\alpha_{jl}\, E\left(G_k G_i G_l G_j\right) = O(1/n^2).$$

Similarly, the same bound holds for $E\left|B_n^i\right|^2$, $i = 2, 3, 4$. Finally, by combining all the previous estimates, we obtain

$$\sqrt{E\left(\frac{1}{2}\|DF_n\|_{L^2}^2 - 2F_n - 2\right)^2} = O(1/\sqrt{n}),$$

and therefore, by using Theorem 3.11 and the fact that $N^2 - 1 \overset{\text{Law}}{=} F(1)$, we deduce that there exists a positive constant $C > 0$ (independent of n) such that

$$d_{\mathscr{H}_1}(F_n, N^2 - 1) \leq C/\sqrt{n},$$

where the class $\mathscr{H}_1$ is defined in (1.14). □

6 An attempt at unification

In this section, we show that the computations contained in Sects. 3.1 and 3.3, respectively, in the Gaussian case and the Gamma case, can be unified, by means of the general theory of approximations developed by Stein in [49, Lecture VI].

Let Z be a real-valued random variable having an absolutely continuous distribution with density $p(x)$, $x \in \mathbb{R}$. We make the following assumptions:

(A1) Z is integrable and centered, that is,

$$E|Z| < \infty \quad \text{and} \quad E(Z) = \int_{-\infty}^{+\infty} y p(y) dy = 0; \tag{6.58}$$

(A2) there exist (possibly infinite) numbers a, b such that $-\infty \leq a < 0 < b \leq +\infty$, and the support of the density p exactly coincides with the open interval (a, b), that is,

$$p(x) > 0 \quad \text{if, and only if,} \quad x \in (a, b). \tag{6.59}$$

Remark 6.1 At the cost of some heavier notation, one could easily generalize the results of this section, in order to accommodate the case of a density p whose support is a union of open (and possibly infinite) intervals.

With a Z verifying assumptions **(A1)–(A2)**, we associate the real-valued mapping $\tau(\cdot)$, defined as

$$x \mapsto \tau(x) = \frac{\int_x^\infty yp(y)dy}{p(x)} \mathbf{1}_{x\in(a,b)} = -\frac{\int_{-\infty}^x yp(y)dy}{p(x)} \mathbf{1}_{x\in(a,b)}, \quad x \in \mathbb{R}. \tag{6.60}$$

Note that τ is well-defined on $\mathbb{R}$, due to assumptions (6.58)–(6.59). Also, relation (6.58) implies that $\tau(x) \geq 0$ for every x and $\tau(x) > 0$ if, and only if, $x \in (a,b)$. The following result, which is proved in [49], states that, under some additional assumptions, the mapping τ completely characterizes the density p, and therefore the law of Z.

Lemma 6.2 (Lemma 3, p. 61 in [49]) *Let the reals a, b be such that $-\infty \leq a < 0 < b \leq +\infty$, and consider a continuous function $\tau(\cdot) \geq 0$ on $\mathbb{R}$ such that*

$$\tau(x) > 0 \quad \textit{if, and only if,} \quad x \in (a,b). \tag{6.61}$$

Then, if

$$\int_0^b (y/\tau(y))dy = +\infty \quad \textit{and} \quad \int_a^0 (y/\tau(y))dy = -\infty, \tag{6.62}$$

there exists a unique (up to sets of zero Lebesgue measure) probability density $p_\tau(\cdot)$ on $\mathbb{R}$ such that the support of p_τ exactly coincides with the interval (a,b) and

$$\int_{-\infty}^{+\infty} yp_\tau(y)dy = 0 \quad \textit{and} \quad \tau(x) = \frac{\int_x^{+\infty} yp_\tau(y)dy}{p_\tau(x)} \mathbf{1}_{x\in(a,b)}, \quad x \in \mathbb{R}. \tag{6.63}$$

The explicit form of p is given by

$$p_\tau(x) = \frac{1}{C} \times \frac{\mathrm{e}^{-\int_0^x \frac{ydy}{\tau(y)}}}{\tau(x)} \mathbf{1}_{x\in(a,b)}, \quad x \in \mathbb{R}, \tag{6.64}$$

where $C = \int_a^b \frac{\mathrm{e}^{-\int_0^x \frac{ydy}{\tau(y)}}}{\tau(x)} dx$, and we used the notational convention $\int_0^x = -\int_x^0$ whenever $x < 0$.

We will see later on that property (6.62) is verified by the functions τ associated with densities in the Pearson's family of continuous distributions. Now let X have a density p verifying assumptions (**A1**)–(**A2**) above, and let τ be the mapping given by (6.60) [for the time being, we do not suppose that (6.62) is verified]. We define the *Stein operator* T_τ, associated with p and τ, as the differential operator

$$T_\tau f(x) = \tau(x)f'(x) - xf(x), \quad x \in \mathbb{R}, \tag{6.65}$$

acting on differentiable functions f. Now fix a function h which is piecewise continuous on $\mathbb{R}$ and such that $E(h(Z))$ is well-defined. The *Stein equation*, associated with p, τ and h, is the first order differential equation

$$h(x) - E(h(Z)) = T_\tau f(x), \quad x \in \mathbb{R}, \tag{6.66}$$

where $T_\tau f$ is defined in (6.65). If τ verifies (6.62), then (due to Lemma 6.2) $E(Z) = E_\tau(h)$, where $E_\tau(h) = \int h(y)p_\tau(y)dy$, and p_τ is the density given by (6.64). It follows that, in this case, one can rewrite (6.66) as

$$h(x) - E_\tau(h) = T_\tau f(x), \quad x \in \mathbb{R}, \tag{6.67}$$

in order to emphasize the role of τ. The next result, whose (rather straightforward) proof is once again given by Stein [49], states that, under (6.62), the Eq. (6.66) admits a unique continuous and bounded solution.

Lemma 6.3 (Lemma 4, p. 62 in [49]) *Let τ satisfy (6.61) and (6.62), and let p_τ be the density associated with τ via (6.64). Then, since τ has support in (a, b), every solution f of (6.67) must necessarily be such that*

$$f(x) = \frac{h(x) - E_\tau(h)}{x}, \quad x \in \mathbb{R}\backslash(a, b). \tag{6.68}$$

Moreover, whenever h is bounded and piecewise continuous, the Eq. (6.67) admits a unique solution f which is bounded and continuous on (a, b). This unique solution is defined by (6.68) on $\mathbb{R}\backslash(a, b)$, and by

$$f(x) = \int_a^x (h(y) - E_\tau(h)) \frac{e^{\int_y^x \frac{z dz}{\tau(z)}}}{\tau(y)} dy, \quad \textit{for every} \quad x \in (a, b). \tag{6.69}$$

Given a bounded and piecewise continuous function h on $\mathbb{R}$, we define the function $U_\tau h$ as

$$U_\tau h(x) = \begin{cases} \frac{h(x) - E_\tau(h)}{x}, & \text{if } x \in \mathbb{R}\backslash(a, b) \\ \int_a^x (h(y) - E_\tau(h)) \frac{e^{\int_y^w \frac{z dz}{\tau(z)}}}{\tau(y)} dy, & \text{if } x \in (a, b), \end{cases} \tag{6.70}$$

so that one can rephrase Lemma 6.3 by saying that $U_\tau h$ is the unique solution of (6.67) which is bounded and continuous on (a, b) (note that $U_\tau h$ can be discontinuous only at a or b, whenever they are finite). We also record the following consequence of the calculations contained in [49, formulae (34)–(35), pp. 64–65]: *if h is bounded and piecewise continuous, then*

$$\sup_{x \in (a,b)} [|x U_\tau h(x)| + |\tau(x) U_\tau' h(x)|] \le 6 \sup_{x \in (a,b)} |h(x)|, \tag{6.71}$$

where $U'_\tau h = (U_\tau h)'$. Note that, due to (6.68), one deduces immediately from (6.71) that

$$\sup_{x\in\mathbb{R}}[|xU_\tau h(x)| + |\tau(x)U'_\tau h(x)|] \leq K \sup_{x\in\mathbb{R}} |h(x)|, \tag{6.72}$$

where $K = 2\max\{3; 1/a; 1/b\}$ (with $1/\pm\infty = 0$). The next statement provides a typical "Stein-type characterization" of the law of Z. It is a general version of Lemmas 1.2(i) and 1.3(i).

Proposition 6.4 *Let Z be a random variable having a density p verifying assumptions* (**A1**)–(**A2**). *Let τ be related to p by* (6.60).

(i) *For every differentiable f such that $E|\tau(Z)f'(Z)| < \infty$, one has that $E|Zf(Z)| < \infty$ and*

$$E[T_\tau(Z)] = E[\tau(Z)f'(Z) - Zf(Z)] = 0. \tag{6.73}$$

(ii) *Suppose in addition that τ verifies* (6.62). *Let Y be a real-valued random variable with an absolutely continuous distribution. Suppose that, for every differentiable f such that the mapping $x \mapsto |\tau(x)f'(x)| + |xf(x)|$ $(x \in \mathbb{R})$ is bounded, one has that*

$$E[T_\tau(Y)] = E[\tau(Y)f'(Y) - Yf(Y)] = 0. \tag{6.74}$$

Then, $Y \stackrel{\text{Law}}{=} Z$.

Proof Part (i) is proved in [49, Lemma 1, p. 69]. Part (ii) is a consequence of the fact that, if (6.74) is in order, then [due to (6.69)–(6.71)], for every bounded and piecewise continuous function h on $\mathbb{R}$, $0 = E[\tau(Y)U'_\tau h(Y) - YU_\tau h(Y)] = E[h(Y)] - E_\tau(h) = E[h(Y)] - E[h(Z)]$. □

The following corollary can be proved along the lines of Theorems 3.1 and 3.11.

Corollary 6.5 *Let Z be a random variable having a density p verifying assumptions* (**A1**)–(**A2**). *Let τ be related to p by* (6.60). *Let $F \in \mathbb{D}^{1,2}$ be a smooth functional of some isonormal Gaussian process. Assume moreover that $E(F) = 0$ and the law of F is absolutely continuous with respect to the Lebesgue measure. Then, for every bounded and piecewise continuous function h, we have*

$$E(h(F)) - E(h(Z)) = E[\tau(F)(U_\tau h)'(F) - FU_\tau h(F)] \tag{6.75}$$

$$= E[(U_\tau h)'(F)(\tau(F) - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})]. \tag{6.76}$$

Also,

$$|E(h(F)) - E(h(Z))| \leq E[(U_\tau h)'(F)^2]^{1/2} E[(\tau(F) - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})^2]^{1/2}. \tag{6.77}$$

It is not difficult to see that the conclusions of Theorems 3.1 and 3.11 are indeed corollaries of formula (6.77), corresponding, respectively, to $\tau(x) = 1$ and $\tau(x) = 2(x+\nu)_+$. Plainly, a study of general expressions such as the RHS of (6.77) would require a fine analysis of the properties of the solutions to the Stein equation (6.66) (similar to the ones performed in the Gamma case by Luk and Pickett, respectively, in [26,41]). This topic is clearly outside the scope of the present paper. However, we conjecture that such a study could be successfully performed in the case where the density p belongs to the Pearson's family of curves. Indeed, in this case the function τ can be neatly characterized in terms of polynomials of degree 2.

To see this, let Z satisfy **(A1)–(A2)**, and let τ satisfy (6.62). We say that Z is a (centered) member of the *Pearson's family* of continuous distributions, whenever the density $p = p_\tau$ [see (6.64)] satisfies the differential equation

$$\frac{p'(x)}{p(x)} = \frac{a_0 + a_1 x}{b_0 + b_1 x + b_2 x^2}, \quad x \in (a, b), \tag{6.78}$$

for some real numbers a_0, a_1, b_0, b_1, b_2. We refer the reader e.g. to [15, Sect. 5.1] for an introduction to the Pearson's family. Here, we shall only observe that there are basically five families of distributions satisfying (6.78): the centered normal distributions, centered Gamma and beta distributions, and distributions that are obtained by centering densities of the type $p(x) = Cx^{-\alpha}\exp(-\beta/x)$ or $p(x) = C(1+x)^{-\alpha}\exp(\beta \arctan(x))$ (C being a suitable normalizing constant). The next result, proved in [49, Theorem 1, p. 65], states that a density belongs to the class of the Pearson's curves if, and only if, its associated mapping τ is a polynomial of degree ≤ 2. The reader is also referred to [46, Sects. 2 and 4] for several related results and explicit computations involving orthogonal polynomials.

Theorem 6.6 (Stein) *Let Z satisfy* ***(A1)–(A2)****, and let τ satisfy* (6.62)*. Then, the density $p = p_\tau$ is such that $\tau(x) = \alpha x^2 + \beta x + \gamma$, $x \in (a, b)$ (with α, β, γ constants) if, and only if, p satisfies* (6.78) *for every $x \in (a, b)$ and for $a_0 = \beta$, $a_1 = 2\alpha + 1$, $b_0 = \gamma$, $b_1 = \beta$ and $b_2 = \alpha$.*

Of course, in order for (6.62) to be satisfied, one must have that the $\tau(a) = 0$ (whenever a is finite) and $\tau(b) = 0$ (whenever b is finite). As already discussed, the centered Gaussian distribution is a member of the Pearson's family, corresponding to the case $a = -\infty$, $b = +\infty$ and $\tau(x) = 1$. Analogously, a centered Gamma random variable $F(\nu)$ as in (1.4) has a density of the Pearson type, with characteristics $a = -\nu$, $b = +\infty$ and $\tau(x) = 2(x+\nu)_+$.

7 Two proofs

7.1 Proof of Lemma 1.3

Proof of Point (i) One could use directly Proposition 6.4 in the case $\tau(x) = 2(x+\nu)_+$. Alternatively, observe first that, for every $\nu > 0$, the random variable $F^*(\nu) := F(\nu) + \nu$ has a non-centered Gamma law with parameter $\nu/2$. The fact that

$$E[2F^*(\nu)f'(F^*(\nu)-\nu)] = E[2(F^*(\nu)-\nu)f'(F^*(\nu))],$$

for every f as in the statement, is therefore an immediate consequence of [46, Proposition 1 and Sect. 4(2)]. Now suppose that W verifies (1.13). By choosing f with support in $(-\infty,-\nu)$, one deduces immediately that $P(W \le -\nu) = 0$. To conclude, we apply once again the results contained in [46], to infer that the relations

$$P(W \le -\nu) = 0 \quad \text{and} \quad E[2(W+\nu)f'(W) - Wf(W)] = 0$$

imply that, necessarily, $W+\nu \overset{\text{Law}}{=} F^*(\nu)$.

Proof of Point (ii) Fix $\nu > 0$, consider a function h as in the statement and define $h_\nu(y) = h(y-\nu)$, $y > 0$. Plainly, h_ν is twice differentiable, and $|h_\nu(y)| \le c\exp\{-\nu a\}\exp\{ay\}$, $y > 0$ (recall that $a > 1/2$). In view of these properties, according to Luk [26, Theorem 1], the second-order Stein equation

$$h_\nu(y) - E(h_\nu(F^*(\nu))) = 2yg''(y) - (y-\nu)g'(y), \quad y > 0, \tag{7.79}$$

(where, as before, we set $F^*(\nu) = F(\nu)+\nu$) admits a solution g such that $\|g'\|_\infty \le 2\|h'\|_\infty$ and $\|g''\|_\infty \le \|h''\|_\infty$. Since $f(x) = g'(x+\nu)$, $x > -\nu$, is a solution of (1.12), the conclusion is immediately obtained.

Proof of Point (iii) According to a result of Pickett [41], as reported in [43, Lemma 3.1], when $\nu \ge 1$ is an integer, the ancillary Stein equation (7.79) admits a solution g such that $\|g'\|_\infty \le \sqrt{2\pi/\nu}\|h\|_\infty$ and $\|g''\|_\infty \le \sqrt{2\pi/\nu}\|h'\|_\infty$. The conclusion is obtained as in the proof of Point (ii).

7.2 Proof of Theorem 1.5

We begin with a technical lemma.

Lemma 7.1 *Let $F = I_2(f)$ be a random variable living in the second Wiener chaos of an isonormal Gaussian process X (over a real Hilbert space $\mathfrak{H}$). Then*

$$E\left(\|DF\|_{\mathfrak{H}}^4\right) = \frac{2}{3}E(F^4) + 2E(F^2)^2. \tag{7.80}$$

Proof Without loss of generality, we can assume that $\mathfrak{H} = \mathrm{L}^2(A,\mathscr{A},\mu)$, where $(A,\mathscr{A})$ is a measurable space, and μ is a σ-finite and non-atomic measure. On one hand, thanks to the multiplication formula (2.29), we can write

$$F^2 = I_4(f\otimes f) + 4\,I_2(f\otimes_1 f) + E\left(F^2\right).$$

In particular, this yields

$$L(F^2) = -4\, I_4(f \otimes f) - 8\, I_2(f \otimes_1 f).$$

On the other hand, (2.32) implies that $D_a F = 2\, I_1\,(f(\cdot, a))$. Consequently, again by (2.29):

$$\begin{aligned}
\|DF\|^2_{\mathfrak{H}} &= 4 \int\limits_A I_1\,(f(\cdot, a))^2\, \mu(da) \\
&= 4 \int\limits_A I_2\,(f(\cdot, a) \otimes f(\cdot, a))\, \mu(da) + E\left(\|DF\|^2_{\mathfrak{H}}\right) \\
&= 4\, I_2(f \otimes_1 f) + 2E(F^2), \\
&\quad \text{by (2.34) and since } \textstyle\int_A f(\cdot, a) \otimes f(\cdot, a)\mu(da) = f \otimes_1 f. \qquad (7.81)
\end{aligned}$$

Taking into account the orthogonality between multiple stochastic integrals of different orders, we deduce

$$\begin{aligned}
E\left[\|DF\|^2_{\mathfrak{H}}\, L(F^2)\right] &= -32\, E\left[(I_2(f \otimes_1 f))^2\right] \\
&= -2\, E\left[\|DF\|^2_{\mathfrak{H}} \left(F^2 - E(F^2)\right)\right]. \qquad (7.82)
\end{aligned}$$

Finally, we have

$$\begin{aligned}
E\left[\|DF\|^4_{\mathfrak{H}}\right] &= E\left[\|DF\|^2_{\mathfrak{H}} \langle DF, DF\rangle_{\mathfrak{H}}\right] \\
&= E\left[\|DF\|^2_{\mathfrak{H}} \left(\delta DF \times F - \frac{1}{2}\delta D(F^2)\right)\right] \quad \text{by identity (2.31)}, \\
&= 2\, E\left[\|DF\|^2_{\mathfrak{H}} F^2\right] + \frac{1}{2}\, E\left[\|DF\|^2_{\mathfrak{H}}\, L(F^2)\right] \quad \text{using } \delta D = -L, \\
&= E\left[\|DF\|^2_{\mathfrak{H}} F^2\right] + E(F^2) E\left[\|DF\|^2_{\mathfrak{H}}\right] \quad \text{using (7.82)}, \\
&= \frac{2}{3}\, E\left(F^4\right) + 2\, E\left(F^2\right)^2 \quad \text{by (2.34)}.
\end{aligned}$$

□

Now, let us go back to the proof of the first point in Theorem 1.5. In view of Theorem 3.1, it is sufficient to prove that

$$E\left(\left|1 - \frac{1}{2}\|DZ_n\|^2_{\mathfrak{H}}\right|^2\right) \le \frac{1}{6}\left|E(Z_n^4) - 3\right| + \frac{3 + E(Z_n^2)}{2}\left|E(Z_n^2) - 1\right|. \qquad (7.83)$$

We have

$$E\left(\left|1-\frac{1}{2}\|DZ_n\|_{\mathfrak{H}}^2\right|^2\right)=1-E(\|DZ_n\|_{\mathfrak{H}}^2)+\frac{1}{4}E(\|DZ_n\|_{\mathfrak{H}}^4)$$
$$=1-2E(Z_n^2)+\frac{1}{6}E(Z_n^4)$$
$$+\frac{1}{2}E(Z_n^2)^2 \quad \text{by (2.34) and (7.80)}$$
$$=\frac{1}{6}(E(Z_n^4)-3)+(E(Z_n^2)-1)\left(\frac{1}{2}E(Z_n^2)-\frac{3}{2}\right).$$

The estimate (7.83) follows immediately.

Similarly, for the second point of Theorem 1.5, it is sufficient to prove (see Proposition 3.13) that

$$E\left(\left|2Z_n-2\nu-\frac{1}{2}\|DZ_n\|_{\mathfrak{H}}^2\right|^2\right)\le\frac{1}{6}\left|E(Z_n^4)-12E(Z_n^3)-12\nu^2+48\nu\right|$$
$$+\frac{\left|8-6\nu+E(Z_n^2)\right|}{2}\left|E(Z_n^2)-2\nu\right|. \tag{7.84}$$

By using the relations

$$E\left(\left|2Z_n-2\nu-\frac{1}{2}\|DZ_n\|_{\mathfrak{H}}^2\right|^2\right)$$
$$=4E(Z_n^2)+4\nu^2+\frac{1}{4}E(\|DZ_n\|_{\mathfrak{H}}^4)-2E(Z_n\|DZ_n\|_{\mathfrak{H}}^2)-2\nu E(\|DZ_n\|_{\mathfrak{H}}^2)$$
$$=4(1-\nu)E(Z_n^2)+4\nu^2+\frac{1}{6}E(Z_n^4)+\frac{1}{2}E(Z_n^2)^2-2E(Z_n^3) \quad \text{by (2.34) and (7.80)}$$
$$=(E(Z_n^2)-2\nu)\left(4-3\nu+\frac{1}{2}E(Z_n^2)\right)+\frac{1}{6}\left(E(Z_n^4)-12E(Z_n^3)-12\nu^2+48\nu\right),$$

the estimate (7.84) follows immediately. □

Acknowledgments We thank an anonymous referee for interesting suggestions and remarks. We are grateful to G. Reinert for bringing to our attention references [26,41]. We also thank J. Dedecker for useful discussions. *This paper is dedicated to the memory of Livio Zerbini.*

References

1. Baldi, P., Kerkyacharian, G., Marinucci, D., Picard, D.: Subsampling spherical needlets on the sphere (2007, preprint)
2. Barndorff-Nielsen, O., Corcuera, J., Podolskij, M.: Power variation for Gaussian processes with stationary increments (2007, preprint)
3. Barbour, A.D.: Stein's method for diffusion approximations. Probab. Theory Relat. Fields **84**(3), 297–322 (1990)

4. Breton, J.-C., Nourdin, I.: Error bounds on the non-normal approximation of Hermite power variations of fractional Brownian motion (2008, preprint)
5. Breuer, P., Major, P.: Central limit theorems for nonlinear functionals of Gaussian fields. J. Multivar. Anal. **13**(3), 425–441 (1983)
6. Cacoullos, T., Papathanasiou, V., Utev, S.A.: Variational inequalities with examples and an applications to the central limit theorem. Ann. Prob. **22**(3), 1607–1618 (1994)
7. Chambers, D., Slud, E.: Central limit theorems for nonlinear functionals of stationary Gaussian processes. Probab. Theory Relat. Fields **80**, 323–349 (1989)
8. Chatterjee, S.: A new method of normal approximation. Ann. Prob. (to appear) (2007)
9. Chatterjee, S.: Fluctuation of eigenvalues and second order Poincaré inequalities. Probab. Theory Relat. Fields (to appear) (2007)
10. Chen, L., Shao, Q.-M.: Normal approximation under local dependence. Ann. Probab. **32**(3A), 1985–2028 (2004)
11. Chen, L., Shao, Q.-M.: Stein's method for normal approximation. In: An introduction to Stein's method. Lecture Notes Series International Mathematical Sciences National University Singapore, vol. 4, pp. 1–59. Singapore University Press, Singapore (2005)
12. Corcuera, J.M., Nualart, D., Woerner, J.H.C.: Power variation of some integral fractional processes. Bernoulli **12**, 713–735 (2006)
13. Decreusefond, L., Savy, N.: Rubinstein distance on configuration space (2007, preprint)
14. Deheuvels, P., Peccati, G., Yor, M.: On quadratic functionals of the Brownian sheet and related processes. Stoch. Proc. Appl. **116**, 493–538 (2006)
15. Diaconis, P., Zabell, S.: Closed form summation for classical distributions: variations on a theme of de Moivre. Stat. Sci. **6**(3), 284–302 (1991)
16. Dudley, R.M.: Real Analysis and Probability, 2nd edn. Cambridge University Press, Cambridge (2003)
17. Feller, W.: An Introduction to Probability Theory and its Applications, vol. II, 2nd edn. Wiley, London (1970)
18. Giraitis, L., Surgailis, D.: CLT and other limit theorems for functionals of Gaussian processes. Z. Wahrsch. Verw. Geb. **70**, 191–212 (1985)
19. Götze, F.: On the rate of convergence in the multivariate CLT. Ann. Probab. **19**(2), 724–739 (1991)
20. Goldstein, L., Reinert, G.: Stein's method and the zero bias transformation with application to simple random sampling. Ann. Appl. Probab. **7**(4), 935–952 (1997)
21. Goldstein, L., Rinott, Y.: On multivariate normal approximations by Stein's method and size bias couplings. J. Appl. Probab. **33**, 1–17 (1996)
22. Houdré, C., Pérez-Abreu, V.: Covariance identities and inequalities for functionals on Wiener and Poisson spaces. Ann. Probab. **23**, 400–419 (1995)
23. Hsu, E.P.: Characterization of Brownian motion on manifolds through integration by parts. In: Stein's method and applications. Lecture Notes Series International Mathematical Sciences National University Singapore, vol. 5, pp. 195–208. Singapore University Press, Singapore (2005)
24. Hu, Y., Nualart, D.: Renormalized self-intersection local time for fractional Brownian motion. Ann. Probab. **33**(3), 948–983 (2005)
25. Janson, S.: Gaussian Hilbert Spaces. Cambridge University Press, London (1997)
26. Luk, H.M.: Stein's method for the Gamma Distribution and Related Statistical Applications. Ph.D. Dissertation, University of Southern California, Los Angeles, USA (1994)
27. Major, P.: Multiple Wiener–Itô Integrals. Lecture Notes in Math., vol. 849. Springer, Berlin (1981)
28. Malliavin, P.: Stochastic Analysis. Springer, Berlin (1997)
29. Marinucci, D., Peccati, G.: High-frequency asymptotics for subordinated stationary fields on an Abelian compact group. Stoch. Proc. Appl. **118**(4), 585–613 (2007a)
30. Marinucci, D., Peccati, G.: Group representations and high-resolution central limit theorems for subordinated spherical random fields (2007b, preprint)
31. Meyer, P.-A.: Note sur les processus de Ornstein-Uhlenbeck. In: Séminaire de Probabilités XVI. Lecture Notes in Math., vol. 920, pp. 95–132. Springer, Berlin (1982)
32. Neuenkirch, A., Nourdin, I.: Exact rate of convergence of some approximation schemes associated to SDEs driven by fractional Brownian motion. J. Theoret. Probab. **20**(4), 871–899 (2007)
33. Nourdin, I., Peccati, G.: Non-central convergence of multiple integrals (2007, preprint)
34. Nourdin, I., Peccati, G.: Weighted variations of iterated Brownian motion. Electron. J. Probab. (to appear) (2007)
35. Nualart, D.: The Malliavin Calculus and Related Topics, 2nd edn. Springer, Berlin (2006)

36. Nualart, D., Ortiz-Latorre, S.: Central limit theorems for multiple stochastic integrals and Malliavin calculus. Stoch. Proc. Appl. **118**(4), 614–628 (2008)
37. Nualart, D., Peccati, G.: Central limit theorems for sequences of multiple stochastic integrals. Ann. Probab. **33**(1), 177–193 (2005)
38. Peccati, G.: Gaussian approximations of multiple integrals. Electron. Comm. Prob **12**, 350–364 (2007). (electronic)
39. Peccati, G., Tudor, C.A.: Gaussian limits for vector-valued multiple stochastic integrals. In: Séminaire de Probabilités XXXVIII. Lecture Notes in Mathematical, vol. 1857, pp. 247–262. Springer, Berlin (2005)
40. Picard, J.: A tree approach to p-variation and to integration. Ann. Probab. (to appear) (2007)
41. Pickett, A.H.: Rates of Convergence in χ^2 Approximations via Stein's Method. Ph.D. Dissertation, Oxford University. Oxford, UK (2004)
42. Privault, N., Reveillac, A.: Stein estimation for the drift of Gaussian processes using the Malliavin calculus. Ann. Stat. (to appear) (2007)
43. Reinert, G.: Three general approaches to Stein's method. In: An introduction to Stein's method. Lecture Notes Series International Mathematical Sciences National University Singapore, vol. 4, pp. 183–221. Singapore University Press, Singapore (2005)
44. Rinott, Y., Rotar, V.: Normal approximation by Stein's method. Decis. Econ. Financ. **23**, 15–29 (2000)
45. Rota, G.C., Wallstrom, T.C.: Stochastic integrals: a combinatorial approach. Ann. Probab. **25**(3), 1257–1283 (1997)
46. Schoutens, W.: Orthogonal polynomials in Stein's method. J. Math. Anal. Appl. **253**(2), 515–531 (2001)
47. Stein, Ch.: Inadmissibility of the usual estimator for the variance of a normal distribution with unknonwn mean. Ann. Inst. Statist. Math. **16**, 155–160 (1964)
48. Stein, Ch.: A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, vol. II: Probability theory, 583–602. University California Press, Berkeley (1972)
49. Stein, Ch.: Approximate computation of expectations. Institute of Mathematical Statistics Lecture Notes—Monograph Series, vol. 7. Institute of Mathematical Statistics, Hayward (1986)
50. Surgailis, D.: CLTs for polynomials of linear sequences: diagram formulae with applications. In: Long Range Dependence, pp. 111–128. Birkhäuser, Basel (2003a)
51. Surgailis, D.: Non-CLT's: U-statistics, multinomial formula and approximations of multiple Wiener–Itô integrals. In: Long Range Dependence, pp. 129–142. Birkhäuser, Basel (2003b)
52. Taqqu, M.S.: Weak convergence to fractional Brownian motion and to Rosenblatt process. Z. Wahrsch. Verw. Gebiete **31**, 287–302 (1975)
53. Taqqu, M.S.: Fractional Brownian lotion and long-range dependence. In: Long Range Dependence, pp. 5–38. Birkhäuser, Basel (2003)
54. Tikhomirov, A.N.: On the convergence of the remainder in the central limit theorem for weakly dependent random variables. Teor. Veroyatnost. i Primenen. Theory Probab. Appl. **25**(4), 800–818 (1980)
55. Tudor, C.A., Viens, F.: Variations and estimators for the selfsimilarity order through Malliavin calculus (2007, preprint)