

Berry-Esseen bounds in the Breuer-Major CLT and Gebelein's inequality

Ivan Nourdin* Giovanni Peccati† Xiaochuan Yang‡

Abstract

We derive explicit Berry-Esseen bounds in the total variation distance for the Breuer-Major central limit theorem, in the case of a subordinating function φ satisfying minimal regularity assumptions. Our approach is based on the combination of the Malliavin-Stein approach for normal approximations with Gebelein's inequality, bounding the covariance of functionals of Gaussian fields in terms of maximal correlation coefficients.

Keywords: Breuer-Major theorem; rate of convergence; Gebelein's inequality; Malliavin-Stein approach.

AMS MSC 2010: 60F05; 60G15; 60H07.

Submitted to ECP on February 10, 2019, final version accepted on May 10, 2019.

1 Introduction

1.1 Motivation and main results

Let $X = (X_k)_{k \in \mathbb{Z}}$ be a centered stationary Gaussian sequence with covariance function $\mathbb{E}[X_k X_j] = \rho(k - j)$ satisfying $\rho(0) = 1$. Let $\varphi \in L^2(\mathbb{R}, \gamma)$, where γ is the standard Gaussian measure on the real line, and assume without loss of generality that $\mathbb{E}[\varphi(X_1)] = \int_{\mathbb{R}} \varphi d\gamma = 0$. By exploiting the orthogonality and completeness of Hermite polynomials in $L^2(\mathbb{R}, \gamma)$ (see, e.g., [12, p. 13]), we can write

$$\varphi = \sum_{\ell \geq d} a_\ell H_\ell, \tag{1.1}$$

where H_ℓ is the Hermite polynomial of order ℓ , the coefficient a_d is different from zero, $d \geq 1$ is the *Hermite rank* of φ , and the series converges in $L^2(\mathbb{R}, \gamma)$. Consider the sequence of normalized sums

$$F_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n \varphi(X_k), \quad n \geq 1. \tag{1.2}$$

*Université du Luxembourg, Unité de Recherche en Mathématiques, Maison du Nombre, 6 avenue de la Fonte, L-4364 Esch-sur-Alzette, Grand Duché du Luxembourg E-mail: ivan.nourdin@uni.lu

†Université du Luxembourg, Unité de Recherche en Mathématiques, Maison du Nombre, 6 avenue de la Fonte, L-4364 Esch-sur-Alzette, Grand Duché du Luxembourg E-mail: giovanni.peccati@uni.lu

‡Université du Luxembourg, Unité de Recherche en Mathématiques, Maison du Nombre, 6 avenue de la Fonte, L-4364 Esch-sur-Alzette, Grand Duché du Luxembourg E-mail: xiaochuan.yang@uni.lu

The celebrated *Breuer-Major theorem* [2], stated below, provides sufficient conditions on the covariance function ρ , in order for F_n to exhibit Gaussian fluctuations, as $n \rightarrow \infty$ (see also Taggu [22] for a related work). Throughout the paper, the symbol $N(a, b)$ denotes a Gaussian random variable with mean $a \in \mathbb{R}$ and variance $b \geq 0$, and \xrightarrow{d} the convergence in distribution.

Theorem 1.1 (Breuer-Major Theorem). *Let the previous assumptions on X and φ prevail, and suppose moreover that $\sum_{k \in \mathbb{Z}} |\rho(k)|^d < \infty$. Then $F_n \xrightarrow{d} N(0, \sigma^2)$, where*

$$\sigma^2 = \sum_{\ell \geq d} a_\ell^2 \ell! \sum_{k \in \mathbb{Z}} \rho(k)^\ell < \infty. \tag{1.3}$$

Here, and for the rest of the paper, \xrightarrow{d} denotes convergence in distribution of random variables.

The Breuer-Major theorem has far-reaching applications in many different areas, such as mathematical statistics, signal processing or geometry of random nodal sets, see e.g. [5, 16, 21, 23] and references therein. It has been generalized and refined in various aspects [3, 4, 11, 13, 14].

Now let $\sigma_n^2 := \text{Var}(F_n)$ and $V_n := F_n / \sqrt{\text{Var}(F_n)}$. The aim of the present paper is to develop a novel method for obtaining explicit upper bounds on the sequence

$$d_{\text{TV}}(V_n, N(0, 1)) := \sup_{A \in \mathcal{B}(\mathbb{R})} |\mathbb{P}(V_n \in A) - \mathbb{P}(N(0, 1) \in A)|, \quad n \geq 1,$$

where $\mathcal{B}(\mathbb{R})$ is the Borel σ -algebra on \mathbb{R} , under *minimal regularity assumptions* on the function φ . Our strategy for doing so is to combine the *Malliavin-Stein method for probabilistic approximations* (as described in Section 2.2 below) and the powerful *Gebelein's inequality* for correlation of Gaussian functionals (see [6, 24], as well as Section 5, for a self-contained proof), as applied to non-linear transformations of correlated Gaussian sequences. To the best of our knowledge, our use of Gebelein's inequality is new: it is reasonable to expect that the content of the present work might constitute the blueprint for further applications of such general a bound to probabilistic approximations in a Gaussian setting.

We recall that, for every n , the quantity $d_{\text{TV}}(V_n, N(0, 1))$ corresponds to the *total variation distance* between the distributions of V_n and $N(0, 1)$ — see e.g. [13, Appendix C], and the references therein, for a discussion of the properties of d_{TV} . Any statement yielding the existence of an explicit numerical sequence $\{\alpha_n\}$ such that $\alpha_n \rightarrow 0$ and $d(V_n, N(0, 1)) \leq \alpha_n$, for some distance d , is called a *quantitative Breuer-Major Theorem*.

One of the first quantitative Breuer-Major theorems is contained in the work by Nourdin, Peccati and Podolskij [14] — see, in particular, [14, Cor. 2.4], where the focus is on the Kolmogorov and 1-Wasserstein distances and on the case where φ is a Hermite polynomial of order q . The rates obtained in [14] are, in general, not optimal. We stress that, according to [14, Corollary 2.4], the convergence in distribution in Theorem 1.1 always takes place in the sense of the Kolmogorov and 1-Wasserstein distances.

Determining whether the Breuer-Major CLT holds in the topology of the distance d_{TV} is a much more delicate matter, since – unlike convergence in the Kolmogorov or 1-Wasserstein distances – convergence in total variation cannot take place in full generality, and requires extra regularity assumptions on φ . Our specific aim is therefore to tackle the following problem:

Problem P. Letting the notation and assumptions of Theorem 1.1 prevail, find conditions on φ and ρ in order to have that

$$d_{\text{TV}}(F_n / \sqrt{\text{Var}(F_n)}, N(0, 1)) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

To appreciate the subtlety of Problem P, one should recall the following two facts:

- (i) according to the main findings in [17], if φ is a polynomial, then the convergence in Theorem 1.1 always takes place in the sense of total variation;
- (ii) on the other hand, if one considers independent $X_k \sim N(0, 1)$, then it is immediate to build counterexamples, for instance by setting $\varphi(x) = \text{sign}(x)$ — in which case the assumptions of Theorem 1.1 are satisfied, but $d_{\text{TV}}(F_n/\sqrt{\text{Var}(F_n)}, N(0, 1)) = 1$ for all n .

As anticipated, the content of Points (i) and (ii) suggests that there exists a *minimal* amount of regularity for the function φ , below which convergence in total variation in the Breuer-Major Theorem ceases to take place. Exactly locating such a threshold is the ultimate goal of the line of research inaugurated by the present paper.

As already discussed, in what follows we will be concerned with upper bounds on the rate of convergence in the Breuer-Major theorem when the function φ possibly displays an *infinite* Hermite expansion (1.1), and belongs to the Sobolev space $\mathbb{D}^{1,4}$ — where we adopted the usual notation $\mathbb{D}^{p,q}$ in order to indicate the Sobolev space of those random variables on a Gaussian space that are p times differentiable in the sense of Malliavin, and whose Malliavin derivative is q -integrable (see Section 2 for a precise definition). We consider that the property of belonging to some space $\varphi \in \mathbb{D}^{1,q}$, $q \geq 1$, is somehow unavoidable, in the sense that it is the least requirement on φ that allows one to directly apply the Malliavin-Stein method outlined in Section 2.

The following statement is the main result of the paper:

Theorem 1.2. *Let $X = (X_k)_{k \in \mathbb{Z}}$ be a centered stationary Gaussian sequence with covariance function $\mathbb{E}[X_k X_j] = \rho(k - j)$ satisfying $\rho(0) = 1$, and let $\varphi \in \mathbb{D}^{1,4} \subset L^2(\mathbb{R}, \gamma)$ be such that $\mathbb{E}[\varphi(X_1)] = \int_{\mathbb{R}} \varphi d\gamma = 0$. Let F_n be given by (1.2) and set $\sigma_n^2 = \text{Var}(F_n)$ and $V_n = F_n/\sigma_n$. Then, for a finite constant $C(\varphi)$, whose explicit value is given in (2.6) below:*

- (i) For every n ,

$$d_{\text{TV}}(V_n, N(0, 1)) \leq \frac{4C(\varphi)}{\sigma_n^2} n^{-\frac{1}{2}} \left(\sum_{|k| < n} |\rho(k)| \right)^{\frac{3}{2}}. \tag{1.4}$$

- (ii) If φ is symmetric (or, more generally, 2-sparse, as defined in Section 3.1) then, for all $b \in [1, 2]$ and all n ,

$$d_{\text{TV}}(V_n, N(0, 1)) \leq \frac{4C(\varphi)}{\sigma_n^2} n^{-(\frac{1}{b} - \frac{1}{2})} \left(\sum_{|k| < n} |\rho(k)|^2 \right)^{\frac{1}{2}} \left(\sum_{|k| < n} |\rho(k)|^b \right)^{\frac{1}{b}}. \tag{1.5}$$

Remark 1.3. (1) Recall that, according e.g. to the terminology adopted in [12, Chapter 9], a numerical sequence $\alpha_n \downarrow 0$ is said to provide an *optimal rate* (for $d_{\text{TV}}(V_n, N(0, 1))$), whenever there exist non-zero finite constants $k < K$ such that

$$k\alpha_n \leq d_{\text{TV}}(V_n, N(0, 1)) \leq K\alpha_n,$$

for n large enough. The rate provided in Theorem 1.2-(i) for functions φ with Hermite rank 1 is optimal in this sense. Indeed, in the trivial case where $\rho(j) = 0$ for every $j \neq 0$

and using e.g. the *reverse Berry-Esseen inequality* from [1], it is easy to build a centered smooth function φ with Hermite rank 1 and such that

$$d_{\text{TV}}(V_n, N(0, 1)) \geq Cn^{-1/2},$$

for some absolute constant $C > 0$.

(2) For a function φ having Hermite rank equal to 2, the sufficient condition for asymptotic normality in Theorem 1.1 is that $\rho \in \ell^2(\mathbb{Z})$. Theorem 1.2-(ii) refines such a result by yielding that, in the case of a symmetric φ , convergence in total variation takes place whenever $\rho \in \ell^b(\mathbb{Z})$ ($\subset \ell^2$), for some $b \in [1, 2)$. We also observe that Theorem 1.2-(ii) yields an upper bound on $d_{\text{TV}}(V_n, N(0, 1))$, explicitly interpolating all the cases $\rho \in \ell^b(\mathbb{Z})$, for $1 \leq b < 2$.

(3) By inspection of our forthcoming proof, it will be clear that our techniques do not allow us to deal with the case of a general function $\varphi \in \mathbb{D}^{1,4}$ having Hermite rank equal to 2. This implies that the requirement that φ is 2-sparse cannot easily be removed.

We will now compare our findings with further results in the literature.

1.2 Discussion

In the case where φ has a possibly infinite Hermite expansion (1.1), and under some extra smoothness assumptions, Nourdin, Peccati and Reinert [15], Nualart and Zhou [20] and Vidotto [25] obtained total variation error bounds that are better than those derived in [14]. The rates of convergence deduced in [14] and [15, 20, 25] (that are sometimes optimal, and sometimes not) are all obtained via some variation of the Malliavin-Stein approach described in Section 2.2.

In [20] (the closest reference to the present note), the following general quantitative result is proved (see [20, Th. 4.2 and Th. 4.3(v)]): as $n \rightarrow \infty$, one has that

$$d_{\text{TV}}(V_n, N(0, 1)) = O(n^{-1/2}),$$

provided that

- (a) either φ has Hermite rank 1 ($d = 1$ in (1.1)), $\varphi \in \mathbb{D}^{2,4}$ and $\rho \in \ell^1(\mathbb{Z})$, or
- (b) φ has Hermite rank 2 ($d = 2$ in (1.1)), $\varphi \in \mathbb{D}^{6,8}$ and $\rho \in \ell^{\frac{3}{2}}(\mathbb{Z}) \subset \ell^2(\mathbb{Z})$.

The regularity assumptions on φ required at Points (a) and (b) above are clearly more restrictive than ours. On the other hand, disregarding the regularity of φ , the upper bound of the order $n^{-1/2}$ obtained in [20] is optimal for the set of assumptions at Point (a) and (b) above. The optimality for Point (a) follows from the same argument used in Remark 1.3-(1). Similarly, the order $n^{-1/2}$ under the set of assumptions at Point (b) cannot be improved in general, since it coincides with the *third/fourth cumulant barrier* for the total variation distance, between the laws of a sequence of random variables in a fixed chaos and the standard normal distribution. Such a result was established in full generality in [13, Theorem 11.2], and is presented in the next proposition in the simple case of polynomials of order 2. Here and after, $a(n) \asymp b(n)$ means that the ratio $a(n)/b(n)$ is bounded from above and below by positive finite constants.

Proposition 1.4. [13, Proposition 4.2] *Let F_n be given by (1.2) with $\varphi = H_2$. Set $V_n = F_n/\sqrt{\text{Var}(F_n)}$. Then,*

$$d_{\text{TV}}(V_n, N(0, 1)) \asymp \frac{1}{\sqrt{n}} \left(\sum_{|k| < n} |\rho(k)|^{\frac{3}{2}} \right)^2$$

as $n \rightarrow \infty$. In particular, $d_{\text{TV}}(V_n, N(0, 1)) \asymp \frac{1}{\sqrt{n}}$ if $\rho \in \ell^{\frac{3}{2}}(\mathbb{Z})$.

One interesting subordinating function φ entering the scope of our paper is $\varphi(x) = |x| - \sqrt{2/\pi}$. The Breuer-Major CLT associated with such a mapping has been recently applied in a geometric setting in [3], where φ arose in the approximation of the length of a smooth regularization of the sample paths of a Gaussian process with stationary increments. Note that $\varphi \in \mathbb{D}^{1,q}$ for any $q \geq 1$, but $\varphi \notin \mathbb{D}^{2,2}$. Also, φ has an infinite expansion (1.1) with Hermite rank $d = 2$. Such a case is not covered by the findings of [20] or [14, 15, 24] (due to the lack of sufficient regularity for the function φ), and enters indeed the framework of our main result, stated in Theorem 1.2. The case of such a mapping is also covered by the recent reference [9], where convergence in total variation is deduced for a class much smaller than $\mathbb{D}^{1,4}$, containing however $\varphi(x) = |x| - \sqrt{2/\pi}$.

The higher regularity requirement for φ which is necessary in [20] stems from the method, used therein, of applying integration by parts several times. On the other hand, our approach requires that we only perform one integration by parts in the Malliavin-Stein approach, since our final estimate makes use of the intrinsic correlation bound given by Gebelein's inequality. The use of Gebelein's inequality, which is the main technological breakthrough of the present paper, requires much less regularity on φ .

Although the focus of our paper is on finding minimal regularity assumptions on φ for having convergence in total variation in the Breuer-Major Theorem, a natural question one might ask is whether the rates of convergence implied by our bounds are optimal. In view of Proposition 1.4, applying the upper bound in Theorem 1.2-(ii) to the case $\varphi = H_2$ (and $\rho \in \ell^b(\mathbb{Z})$, for some $1 \leq b < 2$), one obtains a rate which is not optimal. The already mentioned reference [9] shows that our results are, in general, not optimal also for the case $\varphi(x) = |x| - \sqrt{2/\pi}$. Further discussions around this problem are gathered at the end of the paper — see Section 4.

The present paper is organised as follows. We start by reviewing some basic elements of stochastic analysis on the Wiener space and of the Malliavin-Stein approach. Then we introduce the new ingredient, Gebelein's inequality for correlated isonormal Gaussian processes, in Section 2. We apply a Gebelein-Malliavin-Stein bound to prove our main theorem in Section 3. A discussion on optimality is provided in Section 4, thus concluding the paper.

Every random object considered below is defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with \mathbb{E} denoting mathematical expectation with respect to \mathbb{P} .

2 Preliminaries

2.1 Stochastic analysis on the Wiener space

The content of this subsection can be found in [12] or [11]. An *isonormal Gaussian process* $\{W(h) : h \in \mathfrak{H}\}$ is a family of centered Gaussian random variables indexed by a real separable Hilbert space \mathfrak{H} such that the covariance satisfies

$$\mathbb{E}[W(g)W(h)] = \langle g, h \rangle_{\mathfrak{H}}.$$

Let F be a square-integrable functional of an isonormal Gaussian process W . Then, F has a unique Wiener-Itô chaos expansion

$$F = \mathbb{E}[F] + \sum_{k \geq 1} I_k(f_k) \text{ in } L^2(\Omega), \tag{2.1}$$

where $f_k \in \mathfrak{H}^{\otimes k}$ is a symmetric kernel, and $I_k(f_k)$ is the k -th multiple Wiener-Itô integral, $k \geq 1$. By convention we write $I_0(f_0) = f_0 = \mathbb{E}[F]$. By orthogonality between multiple integrals of different orders, we have $\mathbb{E}[F^2] = \sum_{k \geq 0} k! \|f_k\|_{\mathfrak{H}^{\otimes k}}^2$. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be of class C^∞ , and such that all its partial derivatives have at most polynomial growth.

Consider a smooth functional of the form $F = f(W(h_1), \dots, W(h_n))$ with $h_1, \dots, h_n \in \mathfrak{H}$. We define the Malliavin derivative of F as

$$DF = \sum_{i=1}^n \partial_i f(W(h_1), \dots, W(h_n)) h_i.$$

The set of smooth functionals F introduced above is dense in $L^q(\Omega)$, $q \geq 1$, and the operator D is closable. Therefore, D can be extended to $\mathbb{D}^{1,q}$, the set of F such that there exists a sequence of smooth functionals $(F_n)_{n \geq 1}$ satisfying $\mathbb{E}[|F_n - F|^q] \rightarrow 0$ and $\mathbb{E}[\|DF_n - \eta\|_{\mathfrak{H}}^q] \rightarrow 0$, for some $\eta \in L^q(\Omega, \mathfrak{H})$, that we rewrite as $\eta := DF$. One defines similarly \mathbb{D}^p and $\mathbb{D}^{p,q}$. When $q = 2$, these spaces are Hilbert spaces and we have the following characterization in terms of the chaos expansion (2.1):

$$\mathbb{D}^{p,2} = \{F \in L^2(\Omega) : \sum_{k \geq p} k^p k! \|f_k\|_{\mathfrak{H}^{\otimes k}}^2 < \infty\}.$$

The adjoint of D , customarily called the divergence operator or the Skorohod integral, is denoted by δ and satisfies the duality formula,

$$\mathbb{E}[\delta(u)F] = \mathbb{E}[\langle u, DF \rangle_{\mathfrak{H}}] \tag{2.2}$$

for all $F \in \mathbb{D}^{1,2}$, whenever $u : \Omega \rightarrow \mathfrak{H}$ is in the domain of δ . The Ornstein-Uhlenbeck semigroup $(P_t)_{t \geq 0}$ is defined by Mehler's formula for all $F \in L^1(\Omega)$ by

$$P_t F = \mathbb{E}'[F(e^{-t}W + \sqrt{1 - e^{-2t}}W')],$$

where W' is an independent copy of W and \mathbb{E}' denotes the expectation with respect to W' . For $F \in L^2(\Omega)$ given by the chaos expansion (2.1), the Ornstein-Uhlenbeck semigroup takes the form

$$P_t F = \sum_{k \geq 0} e^{-kt} I_k(f_k).$$

The generator of $(P_t)_{t \geq 0}$ is denoted by L and acts on the chaos expansion in a simple way,

$$-LF = \sum_{k \geq 1} k I_k(f_k),$$

with $\text{dom } L = \{F : \sum_{k \geq 1} k^2 k! \|f_k\|_{\mathfrak{H}^{\otimes k}}^2 < \infty\}$. The pseudo-inverse of L is defined by

$$-L^{-1}F = \sum_{k \geq 1} \frac{1}{k} I_k(f_k)$$

for all $F \in L^2(\Omega)$. We have $LL^{-1}F = F - \mathbb{E}[F]$ for all $F \in L^2(\Omega)$. The key identity that links the objects defined above is $L = -\delta D$; in particular, we have $-DL^{-1}F \in \text{dom}(\delta)$ for all $F \in L^2(\Omega)$.

We end this subsection with a fundamental product formula for multiple integrals.

Proposition 2.1 (Product formula). *Let p, q be non-negative integers. Let $f \in \mathfrak{H}^{\otimes p}$ and $g \in \mathfrak{H}^{\otimes q}$ be symmetric kernels. We have*

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

where $f \tilde{\otimes}_r g$ is the symmetrized r -th contraction of f and g , see [12, p. 208] for a definition.

2.2 Malliavin-Stein approach

We make use of an identity (labeled below as (2.3)) first noted by Jaramillo and Nualart in [8].

First of all, we observe that any stationary centered Gaussian sequence $X = \{X_k : k \in \mathbb{Z}\}$ is embedded in an isonormal Gaussian process $W = \{W(h) : h \in \mathfrak{H}\}$. This means that there always exists a Hilbert space \mathfrak{H} and an isonormal Gaussian process W (defined on the same probability space) such that, for some $\{e_k : k \geq 1\} \subset \mathfrak{H}$, $W(e_k) = X_k$ for all k , and consequently $\mathbb{E}[W(e_k)W(e_l)] = \langle e_k, e_l \rangle_{\mathfrak{H}} = \rho(k - l)$, for all k, l (see, e.g., [10, Section 1] for a justification of this fact).

For $\varphi = \sum_{\ell \geq 0} a_\ell H_\ell \in L^2(\mathbb{R}, \gamma)$, we define the shift mapping $\varphi_1 := \sum_{\ell \geq 1} a_\ell H_{\ell-1}$ and set

$$u_n := \frac{1}{\sigma_n \sqrt{n}} \sum_{m=1}^n \varphi_1(X_m) e_m$$

Then,

$$\delta u_n = V_n. \tag{2.3}$$

To prove this, just observe that $u_n = -DL^{-1}V_n$, and then apply the relations $L = -\delta D$ and $LL^{-1}F = F$, valid for any centered random variable $F \in L^2(\Omega)$. By Stein's lemma (see [12, Th. 3.3.1]) for d_{TV} and then by integration by parts via (2.2), we have that

$$\begin{aligned} d_{TV}(V_n, N(0, 1)) &\leq \sup_{g \in \mathcal{G}} |\mathbb{E}[V_n g(V_n)] - \mathbb{E}g'(V_n)| \\ &= \sup_{g \in \mathcal{G}} |\mathbb{E}[\delta(u_n)g(V_n)] - \mathbb{E}g'(V_n)| \\ &= \sup_{g \in \mathcal{G}} |\mathbb{E}g'(V_n)(1 - \langle DV_n, u_n \rangle_{\mathfrak{H}})| \\ &\leq 2\sqrt{\text{Var}(\langle DV_n, u_n \rangle_{\mathfrak{H}})}. \end{aligned} \tag{2.4}$$

where we used the fact that $\mathbb{E}\langle DV_n, u_n \rangle_{\mathfrak{H}} = \mathbb{E}V_n^2 = 1$, and the class \mathcal{G} is composed of those $g : \mathbb{R} \rightarrow \mathbb{R}$ such that $\|g\|_\infty < \frac{\sqrt{2\pi}}{2}$ and $\|g'\|_\infty \leq 2$.

Now we estimate from above the variance in the above bound. Note that, by the chain rule and the relation $DX_k = e_k$,

$$\langle DV_n, u_n \rangle_{\mathfrak{H}} = \frac{1}{\sigma_n^2 n} \sum_{k, \ell=1}^n \varphi'(X_k) \varphi_1(X_\ell) \rho(k - \ell).$$

Hence,

$$\begin{aligned} &\text{Var}(\langle DV_n, u_n \rangle_{\mathfrak{H}}) \tag{2.5} \\ &= \frac{1}{\sigma_n^4 n^2} \sum_{k, \ell, k', \ell'=1}^n \text{Cov}(\varphi'(X_k) \varphi_1(X_\ell), \varphi'(X_{k'}) \varphi_1(X_{\ell'})) \rho(k - \ell) \rho(k' - \ell'). \end{aligned}$$

The following relation is a consequence of Meyer's inequality and of the equivalence of Sobolev norms [18, p.72], justifying our integrability assumption on φ . Its proof is given in [20, Lem. 2.2].

Lemma 2.2. *Let $q > 1$. The shift $\varphi \mapsto \varphi_1$ is a bounded operator from $L^q(\mathbb{R}, \gamma)$ to $L^q(\mathbb{R}, \gamma)$.*

Note that

$$\begin{aligned} \sqrt{\text{Var}(\varphi'(X_k) \varphi_1(X_\ell))} &\leq \sqrt{\mathbb{E} \varphi'(X_k)^2 \varphi_1(X_\ell)^2} \\ &\leq \mathbb{E}[\varphi'(X_0)^4]^{1/4} \mathbb{E}[\varphi_1(X_0)^4]^{1/4} =: C(\varphi) < \infty, \end{aligned} \tag{2.6}$$

so that the covariance in (2.5) is finite.

2.3 Gebelein’s inequality

Up to some slight adaptation, Theorem 2.3 can be deduced from Veraar’s paper [24]. For the sake of completeness, in the Appendix contained in Section 5 we will however present an independent proof of such a result (inspired by the approach of [24]), using tools and concepts that are directly connected to the framework of isonormal Gaussian processes.

Recall that an L^2 functional of an isonormal Gaussian process is said to have Hermite rank d if its projection to the first $d - 1$ chaoses is zero, and its projection to the d -th chaos is non trivial.

Theorem 2.3 (Gebelein’s inequality for isonormal processes). *Let $W = \{W(h) : h \in \mathfrak{H}\}$ be an isonormal Gaussian process over some real separable Hilbert space \mathfrak{H} , and let $\mathfrak{H}_1, \mathfrak{H}_2$ be two Hilbert subspaces of \mathfrak{H} . Define W_1 and W_2 , respectively, to be the restriction of W to \mathfrak{H}_1 and \mathfrak{H}_2 . Now consider two measurable mappings $F_i : \mathbb{R}^{\mathfrak{H}_i} \rightarrow \mathbb{R}$, $i = 1, 2$, and assume that each $F_i(W_i)$ is centred and square-integrable. If F_1 has Hermite rank equal to $p \geq 1$, one has that*

$$|\mathbb{E}[F_1(W_1)F_2(W_2)]| \leq \theta^p \text{Var}(F_1(W_1))^{1/2} \text{Var}(F_2(W_2))^{1/2}, \tag{2.7}$$

where $\theta := \sup_{h \in \mathfrak{H}_1, g \in \mathfrak{H}_2: \|g\|, \|h\|=1} |\langle h, g \rangle| \in [0, 1]$.

3 Proof of the main result

3.1 k -sparsity

As we will see in the next subsection, combining Gebelein’s inequality with the Malliavin-Stein approach will lead to effective upper bounds for the total variation distance in the Breuer-Major CLT. To this end, we need information on the Hermite rank of functionals of the type $F := \varphi'(W(h))\varphi_1(W(g))$ for $h, g \in \mathfrak{H}$ with unit norm, and $\varphi \in \mathbb{D}^{1,4}$. We introduce the notion of k -sparsity.

Definition 3.1. *Let $\varphi \in L^2(\mathbb{R}, \gamma)$ be given by the series expansion $\varphi = \sum_{q \geq d} a_q H_q$. We say the φ is k -sparse if $\min\{j - i : j > i \geq d, a_i \neq 0, a_j \neq 0\} \geq k$.*

Remark 3.2. Symmetric functions are 2-sparse. Indeed, since $H_q(-x) = (-1)^q H(x)$ for all $q \geq 1$, the expansion of a symmetric function satisfies $a_{2k-1} = 0$ for $k \in \mathbb{N}$.

Lemma 3.3. *Assume that $\varphi \in \mathbb{D}^{1,4}$ is 2-sparse and set $F := \varphi'(W(h))\varphi_1(W(g))$, for $h, g \in \mathfrak{H}$ with unit norm. Then $F - \mathbb{E}[F]$ has Hermite rank at least 2.*

Proof. By [12, Th. 2.7.7], we have $H_p(W(e)) = I_p(e^{\otimes p})$ for $e \in \mathfrak{H}$ with $\|e\|_{\mathfrak{H}} = 1$. Thus,

$$\varphi'(W(h))\varphi_1(W(g)) = \sum_{q \geq d} \sum_{p \geq d} q a_q a_p I_{q-1}(h^{\otimes q-1}) I_{p-1}(g^{\otimes p-1}),$$

where the series convergence is in $L^2(\Omega)$. By 2-sparsity, only those products of multiple integrals with indices (p, q) satisfying $p = q$ or $|p - q| \geq 2$ remain. Assume $|p - q| \geq 2$. By Proposition 2.1, the multiple integral of lowest order in the chaos expansion for the product is $I_{|p-q|}(\cdot)$, hence the projection of $I_{q-1}(h^{\otimes q-1}) I_{p-1}(g^{\otimes p-1})$ to the first chaos is zero. If $p = q$, Proposition 2.1 shows that the chaos expansion for the product contains only multiple integrals of even order, ending the proof. \square

3.2 Gebelein-Malliavin-Stein upper bound

Putting things together, we have the following Gebelein-Malliavin-Stein upper bound for the total variation distance.

Proposition 3.4. Let $\varphi(X_1) \in \mathbb{D}^{1,4}$ have Hermite rank $d \geq 1$, and define $V_n = F_n/\sigma_n$ according to (1.2) and $\sigma_n^2 := \text{Var}(F_n)$. We have

$$d_{\text{TV}}(V_n, N(0, 1)) \leq \frac{4C(\varphi)}{\sigma_n^2} \sqrt{\frac{1}{n^2} \sum_{i,j,k,\ell=0}^{n-1} |\rho(j-k)\rho(i-j)\rho(k-\ell)|}.$$

If, in addition, φ is 2-sparse, then

$$d_{\text{TV}}(V_n, N(0, 1)) \leq \frac{4C(\varphi)}{\sigma_n^2} \sqrt{\frac{1}{n^2} \sum_{i,j,k,\ell=0}^{n-1} |\rho(j-k)^2\rho(i-j)\rho(k-\ell)|}.$$

Proof. We evaluate the right-hand side of (2.5), by applying Theorem 2.3 in the specific situation where \mathfrak{H} is the linear span of $\{e_i, e_j, e_k, e_\ell\}$, \mathfrak{H}_1 the linear span of $\{e_i, e_j\}$, \mathfrak{H}_2 the linear span of $\{e_k, e_\ell\}$. It is straightforward that

$$\begin{aligned} |\theta| &= \max(|\rho(i-k)|, |\rho(i-\ell)|, |\rho(j-\ell)|, |\rho(j-k)|) \\ &\leq |\rho(i-k)| + |\rho(i-\ell)| + |\rho(j-\ell)| + |\rho(j-k)|. \end{aligned}$$

The conclusion follows from symmetry, and by using the estimate (2.6). □

3.3 End of the proof

We are now ready to finish the proof of Theorem 1.2. We set $\rho_n(k) = |\rho(k)|1_{|k|<n}$.

Proof of (i). We have

$$\begin{aligned} \sum_{i,j,k,\ell=0}^{n-1} |\rho(j-k)\rho(i-j)\rho(k-\ell)| &\leq \sum_{i,\ell=0}^{n-1} (\rho_n * \rho_n * \rho_n)(i-\ell) \\ &\leq n \|\rho_n * \rho_n * \rho_n\|_{\ell^1(\mathbb{Z})} \\ &\leq n \|\rho_n\|_{\ell^1(\mathbb{Z})}^3, \end{aligned}$$

the last inequality being obtained by applying twice Young’s inequality for convolutions. The result follows from Proposition 3.4. □

Proof of (ii). First we rewrite the sum of products as a sum of the product of convolutions by introducing the function $1_n(k) := 1_{|k|<n}$. We have

$$\begin{aligned} &\sum_{i,j,k,\ell=0}^{n-1} |\rho(j-k)^2\rho(i-j)\rho(k-\ell)| \\ &= \sum_{i,j,k,\ell=0}^{n-1} |\rho(j-k)^2\rho(i-j)\rho(k-\ell)1_n(\ell-i)| \\ &= \sum_{j,\ell=0}^{n-1} (\rho_n * 1_n)(\ell-j)(\rho_n * \rho_n^2)(\ell-j) \leq n \langle \rho_n * 1_n, \rho_n * \rho_n^2 \rangle_{\ell^2(\mathbb{Z})}. \end{aligned}$$

Let $b \in [1, 2]$. Applying successively Hölder’s inequality and Young’s inequality, we are led to

$$\begin{aligned} &\sum_{i,j,k,\ell=0}^{n-1} |\rho(j-k)^2\rho(i-j)\rho(k-\ell)| \\ &\leq n \|\rho_n * 1_n\|_{\ell^{\frac{b}{b-1}}(\mathbb{Z})} \|\rho_n * \rho_n^2\|_{\ell^b(\mathbb{Z})} \\ &\leq n \|\rho_n\|_{\ell^b(\mathbb{Z})} \|1_n\|_{\ell^{\frac{b}{2b-2}}(\mathbb{Z})} \|\rho_n\|_{\ell^b(\mathbb{Z})} \|\rho_n^2\|_{\ell^1(\mathbb{Z})} = n^{\frac{3b-2}{b}} \|\rho_n^2\|_{\ell^1(\mathbb{Z})} \|\rho_n\|_{\ell^b(\mathbb{Z})}^2. \end{aligned}$$

The result follows from Proposition 3.4. □

4 A remark on optimality

Our Gebelein-Malliavin-Stein upper bound (Proposition 3.4) could not provide the rate $n^{-1/2}$ in the case where ρ is square integrable but not summable. Indeed, restricting ourselves to the subset of indices $\{i = j = k\}$, we obtain that

$$\begin{aligned} & \frac{1}{n} \sum_{i,j,k,\ell=0}^{n-1} |\rho(j-k)^2 \rho(i-j) \rho(k-\ell)| \geq \frac{1}{n} \sum_{k,\ell=1}^n |\rho(k-\ell)| \\ & = 1 + 2 \sum_{\ell=1}^{n-1} \left(1 - \frac{\ell}{n}\right) |\rho(\ell)| \geq 1 + \sum_{\ell=1}^{n/2} |\rho(\ell)| \end{aligned}$$

goes to infinity as $n \rightarrow \infty$.

5 Appendix: Proof of Theorem 2.3

We start by proving a similar result in a simpler setting, to which we can reduce the general case.

Proposition 5.1. *Let (X, Y) be a pair of jointly isonormal Gaussian processes over \mathfrak{H} , such that X, Y are rigidly correlated, in the following sense: there exists $\theta \in [-1, 1]$ such that, for every $h, g \in \mathfrak{H}$, one has $\mathbb{E}[X(h)Y(g)] = \theta \langle h, g \rangle$. Consider measurable mappings $F : \mathbb{R}^{\mathfrak{H}} \rightarrow \mathbb{R}$ and $G : \mathbb{R}^{\mathfrak{H}} \rightarrow \mathbb{R}$ such that $F(X)$ and $G(Y)$ are square-integrable and centred, and assume that F has Hermite rank $p \geq 1$. Then,*

$$|\mathbb{E}[F(X)G(Y)]| \leq |\theta|^p \text{Var}(F(X))^{1/2} \text{Var}(G(Y))^{1/2} \tag{5.1}$$

Proof. Let $\{e_i : i \geq 1\}$ be an orthonormal basis of \mathfrak{H} . We write α, β, \dots to indicate multi-indices; for a multi-index α , the symbol H_α indicates the corresponding multivariate polynomial. We also write

$$H_\alpha(X) = H_\alpha(X(e_i) : i = 1, 2, \dots) = \prod_{i=1}^{\infty} H_{\alpha_i}(X(e_i)),$$

where H_k stands for the k th Hermite polynomial in one variable; $H_\alpha(Y)$ is defined analogously. From the properties of Hermite polynomials and from the rigid correlation assumption, we infer that, for any choice of multi-indices α, β , one has that $\mathbb{E}[H_\alpha(X)H_\beta(Y)] = \theta^{|\alpha|} \alpha! \mathbf{1}_{\alpha=\beta}$. Now, by the chaotic representation property of isonormal processes, one has that

$$F(X) = \sum_{\alpha:|\alpha|\geq p} b_\alpha H_\alpha(X), \quad G(Y) = \sum_{\alpha:|\alpha|\geq 1} c_\alpha H_\alpha(Y),$$

with convergence in $L^2(\Omega)$. By virtue of the previous discussion,

$$|\mathbb{E}[F(X)G(Y)]| \leq \sum_{\alpha:|\alpha|\geq p} |b_\alpha c_\alpha| |\theta|^{|\alpha|} \alpha! \leq |\theta|^p \sum_{\alpha:|\alpha|\geq 1} |b_\alpha c_\alpha| \alpha!,$$

and the conclusion follows from an application of the Cauchy-Schwarz inequality. \square

We now turn to the proof of Theorem 2.3.

Proof of Theorem 2.3. Without loss of generality, we assume that $\theta \in (0, 1)$. For $i = 1, 2$, we denote by $\pi_{\mathfrak{H}_i}$ the orthogonal projection operator onto \mathfrak{H}_i . We will make use of the following estimate: for every $g \in \mathfrak{H}_2$ with unit norm,

$$\|\pi_{\mathfrak{H}_1}(g)\| \leq \theta, \tag{5.2}$$

which follows from the relation $\|\pi_{\mathfrak{H}_1}(g)\|^2 = |\langle g, \pi_{\mathfrak{H}_1}(g) \rangle| \leq \theta \|\pi_{\mathfrak{H}_1}(g)\|$. Now write $\mathfrak{H}_1 \oplus \mathfrak{H}_2$ to indicate the direct sum of \mathfrak{H}_1 and \mathfrak{H}_2 . The key of the proof is the definition of mappings $\tau_1 : \mathfrak{H}_2 \rightarrow \mathfrak{H}_1$, $\tau_2 : \mathfrak{H}_2 \rightarrow \mathfrak{H}_2$, and $\tau : \mathfrak{H}_2 \rightarrow \mathfrak{H}_1 \oplus \mathfrak{H}_2$ given by $g \mapsto \tau_1(g) \oplus \tau_2(g)$, with the following two properties:

- (i) for $h \in \mathfrak{H}_1$ and $g \in \mathfrak{H}_2$, $\langle h, \tau_1(g) \rangle = \theta^{-1} \langle h, g \rangle$;
- (ii) τ verifies the isometric property: $\langle \tau(g), \tau(k) \rangle_{\mathfrak{H}_1 \oplus \mathfrak{H}_2} = \langle g, k \rangle_{\mathfrak{H}_2}$, for every $g, k \in \mathfrak{H}_2$.

In order to define such a mapping τ , we first observe that, by virtue of (5.2), the positive self-adjoint and bounded operator U , from \mathfrak{H}_2 into itself, given by $g \mapsto U(g) = \pi_{\mathfrak{H}_2}(\pi_{\mathfrak{H}_1}(g))$, is such that

$$\|U\|_{op} = \sup_{g, k \in \mathfrak{H}_2: \|g\|, \|k\|=1} |\langle U(g), k \rangle| \leq \theta \sup_{g \in \mathfrak{H}_2: \|g\|=1} \|\pi_{\mathfrak{H}_1}(g)\|,$$

and therefore $\|U\|_{op} \leq \theta^2$, by virtue of (5.2). This implies that the operator $V(g) := \sqrt{\text{Id} - U/\theta^2}$ is well-defined. In particular one checks that a mapping τ satisfying the two properties (i) and (ii) listed above is given by $\tau_1(g) = \theta^{-1}\pi_{\mathfrak{H}_1}(g)$ and $\tau_2(g) = V(g)$, for every $g \in \mathfrak{H}_2$. We now consider two auxiliary independent isonormal Gaussian processes Y, Z over $\mathfrak{H}_1 \oplus \mathfrak{H}_2$, and we set $R := \theta Y + \sqrt{1 - \theta^2}Z$, in such a way that Y, R are rigidly correlated with parameter θ , in the sense made clear in the statement of Proposition 5.1. It is also easily verified that, by a direct covariance computation and with obvious notation,

$$(Y(\mathfrak{H}_1 \oplus \{0\}), R(\tau(\mathfrak{H}_2))) \stackrel{\text{law}}{=} (X_1, X_2).$$

To conclude the proof, we apply Proposition 5.1 as follows:

$$\begin{aligned} |\mathbb{E}[F_1(X_1)F_2(X_2)]| &= |\mathbb{E}[F_1(Y(\mathfrak{H}_1 \oplus \{0\}))F_2(R(\tau(\mathfrak{H}_2)))]| \\ &\leq \theta^p \text{Var}(F_1(X_1))^{1/2} \text{Var}(F_2(X_2))^{1/2}, \end{aligned}$$

where we have used the fact that $F_1(Y(\mathfrak{H}_1 \oplus \{0\}))$ has also Hermite rank p , as well as the relations $\text{Var}(F_1(X_1)) = \text{Var}(F_1(Y(\mathfrak{H}_1 \oplus \{0\})))$, and $\text{Var}(F_2(X_2)) = \text{Var}(F_2(R(\tau(\mathfrak{H}_2))))$. \square

References

- [1] A.D. Barbour and P. Hall (1984): Reversing the Berry-Esseen inequality *Proceedings of the American Mathematical Society*, 90(1), pp. 107–110. MR-0722426
- [2] P. Breuer and P. Major (1983): Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Anal.* **13**, no. 3, pp. 425–441. MR-0716933
- [3] S. Campese, I. Nourdin and D. Nualart (2018): Continuous Breuer-Major theorem: tightness and non-stationarity. To appear in *Annals of Probability*.
- [4] D. Chambers and E. Slud (1989): Central limit theorems for nonlinear functionals of stationary Gaussian processes. *Probab. Theory Related Fields* **80**, no. 3, pp. 323–346. MR-0976529
- [5] P. Doukhan (2018): *Stochastic Models for Time Series*. Springer, 308 pages. MR-3852404
- [6] H. Gebelein (1941): Das statistische Problem der Korrelation als Variations- und Eigenwertproblem und sein Zusammenhang mit der Ausgleichsrechnung. *Z. Angew. Math. Mech.* **21**, pp. 364–379. MR-0007220
- [7] Y. Hu and D. Nualart (2010): Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Statist. Probab. Lett.* **80**, no. 11-12, pp. 1030–1038. MR-2638974
- [8] A. Jaramillo and D. Nualart (2017): Asymptotic properties of the derivative of self-intersection local time of fractional Brownian motion. *Stochastic Process. Appl.* **127**, no. 2, pp. 669–700. MR-3583768
- [9] S. Kunzun and D. Nualart (2019): Rate of convergence in the Breuer-Major theorem via chaos expansions. Preprint.

- [10] I. Nourdin (2012): Lectures on Gaussian approximations with Malliavin calculus. *Sém. Probab.* **XLV**, pp. 3–89. MR-3185909
- [11] I. Nourdin and D. Nualart (2018): The functional Breuer-Major theorem. To appear in *Probab Theory Related Fields*.
- [12] I. Nourdin and G. Peccati (2012): *Normal Approximations with Malliavin calculus. From Stein's Method to Universality*. Cambridge Tracts in Mathematics, 192. Cambridge University Press, Cambridge. xiv+239 pp. MR-2962301
- [13] I. Nourdin and G. Peccati (2015): The optimal fourth moment theorem. *Proc. Amer. Math. Soc.* **143**, no. 7, pp. 3123–3133. MR-3336636
- [14] I. Nourdin, G. Peccati and M. Podolskij (2011): Quantitative Breuer-Major theorems. *Stochastic Process. Appl.* **121**, no. 4, pp. 793–812. MR-2770907
- [15] I. Nourdin, G. Peccati and G. Reinert (2009): Second order Poincaré inequalities and CLTs on Wiener space. *J. Funct. Anal.* **257**, pp. 593–609. MR-2527030
- [16] I. Nourdin, G. Peccati and M. Rossi (2017): Nodal statistics of planar random waves. Preprint. MR-3959555
- [17] I. Nourdin and G. Poly (2013): Convergence in total variation on Wiener chaos. *Stoch. Proc. App.*, 123(2), 651–674. MR-3003367
- [18] D. Nualart (2006): *Malliavin Calculus and Related Topics*. Springer-Verlag, Berlin, second edition. MR-1127886
- [19] D. Nualart and E. Nualart (2018): *Introduction to Malliavin Calculus*, IMS Textbooks, Cambridge University Press. MR-3838464
- [20] D. Nualart and H. Zhou (2018): Total variation estimates in the Breuer-Major theorem. Preprint arXiv:1807.09707.
- [21] V. Pipiras and M.S. Taqqu (2017): *Long-Range Dependence and Self-Similarity*. Cambridge University Press. MR-3729426
- [22] M. Taqqu (1979): Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch. Verw. Gebiete* **50**, no. 1, pp. 53–83. MR-0550123
- [23] C.A. Tudor (2013): *Analysis of Variations for Self-similar Processes*. Springer, 268 pages. MR-3112799
- [24] M. Veraar (2009): Correlation inequalities and applications to vector-valued Gaussian random variables and fractional Brownian motion. *Potential Anal.* **30**, no. 4, pp. 341–370. MR-2491457
- [25] A. Vidotto (2017): An improved second order Poincaré inequality for functionals of Gaussian fields. To appear in: *Journal of Theoretical Probability*.

Acknowledgments. We are grateful to David Nualart for stimulating discussions on this topic. Ivan Nourdin is supported by the FNR Grant R-AGR-3585-10 (APOGee) at the University of Luxembourg. Giovanni Peccati is supported by the FNR Grant R-AGR-3376-10 (FoRGES) at the University of Luxembourg. Xiaochuan Yang is supported by the FNR Grant R-AGR-3410-12-Z (MiSSILe) at Luxembourg and Singapore Universities.