

# Multivariate normal approximation on the Wiener space: new bounds in the convex distance

Ivan Nourdin, Giovanni Peccati and Xiaochuan Yang

January 8, 2020

## Abstract

We establish explicit bounds on the convex distance between the distribution of a vector of smooth functionals of a Gaussian field, and that of a normal vector with a positive definite covariance matrix. Our bounds are commensurate to the ones obtained by Nourdin, Peccati and Réveillac (2010) for the (smoother) 1-Wasserstein distance, and do not involve any additional logarithmic factor. One of the main tools exploited in our work is a recursive estimate on the convex distance recently obtained by Schulte and Yukich (2019). We illustrate our abstract results in two different situations: (i) we prove a quantitative multivariate fourth moment theorem for vectors of multiple Wiener-Itô integrals, and (ii) we characterise the rate of convergence for the finite-dimensional distributions in the functional Breuer-Major theorem.

**Keywords:** Breuer-Major Theorem; Convex Distance; Fourth Moment Theorems; Gaussian Fields; Malliavin-Stein Method; Multidimensional Normal Approximations.

**AMS 2010 Classification:**

## 1 Introduction

Fix  $m \geq 1$ , and consider random vectors  $\mathbf{F}$  and  $\mathbf{G}$  with values in  $\mathbb{R}^m$ . The **convex distance** between the distributions of  $\mathbf{F}$  and  $\mathbf{G}$  is defined as

$$d_c(\mathbf{F}, \mathbf{G}) := \sup_{h \in \mathcal{I}_m} |\mathbb{E}h(\mathbf{F}) - \mathbb{E}h(\mathbf{G})|, \quad (1)$$

where the supremum runs over the class  $\mathcal{I}_m$  of indicator functions of the measurable convex subsets of  $\mathbb{R}^m$ . For  $m \geq 2$ , the distance  $d_c$  represents a natural counterpart to the well-known **Kolmogorov distance** on the class of probability distributions on the real line, and enjoys a number of desirable invariance properties that make it well-adapted to applications<sup>1</sup>.

The aim of the present note is to establish explicit bounds on the quantity  $d_c(\mathbf{F}, \mathbf{G})$ , in the special case where  $\mathbf{F}$  is a vector of smooth functionals of an infinite-dimensional Gaussian field, and  $\mathbf{G} = N_\Sigma$  is a  $m$ -dimensional centered Gaussian vector with covariance  $\Sigma > 0$ . Our main tool is the so-called **Malliavin-Stein method** for probabilistic approximations [17], that we will combine with some powerful recursive estimates on  $d_c$ , recently derived in [29] in the context of multidimensional second-order Poincaré inequalities on the Poisson space — see Lemma 2.1 below.

<sup>1</sup>For instance, one has that  $d_c(T\mathbf{F}, T\mathbf{G}) = d_c(\mathbf{F}, \mathbf{G})$ , whenever the mapping  $T : \mathbb{R}^m \rightarrow \mathbb{R}^m$  is an invertible affine mapping — see e.g. [4, 28] for more details.

Multidimensional normal approximations in the convex distance have been the object of an intense study since several decades, mostly in connection with multivariate central limit theorems (CLTs) for sums of independent random vectors — see e.g. [4, 10, 13, 14] for some classical references, as well as [28] for recent advances and for a discussion of further relevant literature. The specific challenge we are setting ourselves in the present work is to establish bounds on the quantity  $d_c(\mathbf{F}, N_\Sigma)$  that coincide (up to an absolute multiplicative constant) with the bounds deduced in [19] on the **1-Wasserstein distance**

$$d_W(\mathbf{F}, N_\Sigma) := \sup_{h \in \text{Lip}(1)} |\mathbb{E}h(\mathbf{F}) - \mathbb{E}h(N_\Sigma)|, \quad (2)$$

where  $\text{Lip}(1)$  denotes the class of  $C^1$  mappings  $h : \mathbb{R}^m \rightarrow \mathbb{R}$  with Lipschitz constant not exceeding 1. We will see that our estimates systematically improve the bounds that one can infer from the general inequality

$$d_c(\mathbf{F}, N_\Sigma) \leq K \sqrt{d_W(\mathbf{F}, N_\Sigma)}, \quad (3)$$

where  $K$  is an absolute constant uniquely depending on  $\Sigma$ . For the sake of completeness, a full proof of (3) is presented in Appendix A, where one can also find more details on the constant  $K$ .

*Remark 1.1.* In order for the quantity  $d_W(\mathbf{F}, N_\Sigma)$  to be well-defined, one needs that  $\mathbb{E}\|\mathbf{F}\|_{\mathbb{R}^m} < \infty$ . In Appendix A we will also implicitly use the well-known representation

$$d_W(\mathbf{F}, N_\Sigma) = \inf_{(\mathbf{U}, \mathbf{V})} \mathbb{E} \|\mathbf{U} - \mathbf{V}\|_{\mathbb{R}^m},$$

where the infimum runs over all couplings  $(\mathbf{U}, \mathbf{V})$  of  $\mathbf{F}$  and  $N_\Sigma$ . See [30, Ch. I-6] for a proof of this fact, and for further relevant properties of Wasserstein distances.

The main contributions of our paper are described in full detail in Section 1.4 and Section 1.5. Section 1.1 contains some elements of Malliavin calculus that are necessary in order to state our findings. Section 1.2 discusses some estimates on the smooth distance  $d_2$  (to be defined therein) that can be obtained by interpolation techniques, whereas Section 1.3 provides an overview of the main results of [19].

*Remark on notation.* From now on, every random element is assumed to be defined on a common probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , with  $\mathbb{E}$  denoting expectation with respect to  $\mathbb{P}$ . For  $p \geq 1$ , we write  $L^p(\Omega) := L^p(\Omega, \mathcal{F}, \mathbb{P})$ .

## 1.1 Elements of Malliavin calculus

The reader is referred e.g. to the monographs [17, 23, 24] for a detailed discussion of the concepts and results presented in this subsection.

Let  $\mathfrak{H}$  be a real separable Hilbert space, and write  $\langle \cdot, \cdot \rangle_{\mathfrak{H}}$  for the corresponding inner product. In what follows, we will denote by  $X = \{X(h) : h \in \mathfrak{H}\}$  an **isonormal Gaussian process** over  $\mathfrak{H}$ , that is,  $X$  is a centered Gaussian family indexed by the elements of  $\mathfrak{H}$  and such that  $\mathbb{E}[X(h)X(g)] = \langle h, g \rangle_{\mathfrak{H}}$  for every  $h, g \in \mathfrak{H}$ . For the rest of the paper, we will assume without loss of generality that  $\mathcal{F}$  coincides with the  $\sigma$ -field generated by  $X$ .

Every  $F \in L^2(\Omega)$  admits a **Wiener-Itô chaos expansion** of the form

$$F = \mathbb{E}F + \sum_{q=1}^{\infty} I_q(f_q),$$

where  $f_q$  belongs to the symmetric  $q$ th tensor product  $\mathfrak{H}^{\odot q}$  (and is uniquely determined by  $F$ ), and  $I_q(f_q)$  is the  $q$ -th **multiple Wiener-Itô integral** of  $f_q$  with respect to  $X$ . One writes  $F \in \mathbb{D}^{1,2}(\Omega)$  if

$$\sum_{q \geq 1} q q! \|f_q\|_{\mathfrak{H}^{\otimes q}}^2 < \infty.$$

For  $F \in \mathbb{D}^{1,2}(\Omega)$ , we denote by  $DF$  the **Malliavin derivative** of  $F$ , and recall that  $DF$  is by definition a random element with values in  $\mathfrak{H}$ . The operator  $D$  satisfies a fundamental **chain rule**: if  $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}$  is  $C^1$  and has bounded derivatives and if  $F_1, \dots, F_m \in \mathbb{D}^{1,2}(\Omega)$ , then  $\varphi(F_1, \dots, F_m) \in \mathbb{D}^{1,2}(\Omega)$  and

$$D\varphi(F_1, \dots, F_m) = \sum_{i=1}^m \partial_i \varphi(F_1, \dots, F_m) DF_i. \quad (4)$$

For general  $p > 2$ , one writes  $F \in \mathbb{D}^{1,p}(\Omega)$  if  $F \in L^p(\Omega) \cap \mathbb{D}^{1,2}(\Omega)$  and  $\mathbb{E}[\|DF\|_{\mathfrak{H}}^p] < \infty$ .

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

$$\mathbb{E}[\delta(u)F] = \mathbb{E}[\langle u, DF \rangle_{\mathfrak{H}}] \quad (5)$$

for all  $F \in \mathbb{D}^{1,2}(\Omega)$ , whenever  $u : \Omega \rightarrow \mathfrak{H}$  is in the domain  $\text{Dom}(\delta)$  of  $\delta$ .

The **generator of the Ornstein-Uhlenbeck semigroup**, written  $L$ , is defined by the relation  $LF = -\sum_{q \geq 1} q I_q(f_q)$  for every  $F$  such that  $\sum_{q \geq 1} q^2 q! \|f_q\|_{\mathfrak{H}^{\otimes q}}^2 < \infty$ . The **pseudo-inverse** of  $L$ , denoted by  $L^{-1}$ , is the operator defined for any  $F \in L^2(\Omega)$  as  $L^{-1}F = -\sum_{q \geq 1} \frac{1}{q} I_q(f_q)$ . The crucial relation that links the objects introduced above is the identity

$$F = \mathbb{E}F - \delta(DL^{-1}F), \quad (6)$$

which is valid for any  $F \in L^2(\Omega)$  (in particular, one has that, for every  $F \in L^2(\Omega)$ ,  $DL^{-1}F \in \text{Dom}(\delta)$ ).

## 1.2 Bounds on the smooth distance $d_2$

Fix  $m \geq 1$  and assume that  $\mathbf{F} = (F_1, \dots, F_m)$  is a centered random vector in  $\mathbb{R}^m$  whose components belong to  $\mathbb{D}^{1,2}(\Omega)$ . Without loss of generality, we can assume that each  $F_i$  is of the form  $F_i = \delta(u_i)$  for some  $u_i \in \text{Dom}(\delta)$ ; indeed, by virtue of (6) one can always set  $u_i = -DL^{-1}F_i$  (although this is by no means the only possible choice). Let also  $N_\Sigma = (N_1, \dots, N_m)$  be a centered Gaussian vector with invertible  $m \times m$  covariance matrix  $\Sigma = \{\Sigma(i, j)\} = \{\Sigma(i, j) : i, j = 1, \dots, m\}$ . Finally, consider the so-called  **$d_2$  distance** (between the distributions of  $\mathbf{F}$  and  $N_\Sigma$ ) defined by

$$d_2(\mathbf{F}, N_\Sigma) = \sup_h |\mathbb{E}h(\mathbf{F}) - \mathbb{E}h(N_\Sigma)|,$$

where the supremum is taken over all  $C^2$  functions  $h : \mathbb{R}^m \rightarrow \mathbb{R}$  that are 1-Lipschitz and such that  $\sup_{\mathbf{x} \in \mathbb{R}^m} \|(\text{Hess } h)(\mathbf{x})\|_{\text{H.S.}} \leq 1$ ; here,  $\text{Hess } h$  stands for the Hessian matrix of  $h$ , whereas  $\|\cdot\|_{\text{H.S.}}$  denotes the Hilbert-Schmidt norm, that is,  $\|A\|_{\text{H.S.}}^2 = \text{Tr}(AA^T) = \sum_{1 \leq i, j \leq m} A(i, j)^2$  for any  $m \times m$  matrix  $A = \{A(i, j)\}$ .

For a given  $h : \mathbb{R}^m \rightarrow \mathbb{R} \in C^2$  with bounded partial derivatives, let us introduce its mollification at level  $\sqrt{t}$ , defined by

$$h_t(\mathbf{x}) := \mathbb{E}[h(\sqrt{t}N_\Sigma + \sqrt{1-t}\mathbf{x})], \quad \mathbf{x} \in \mathbb{R}^m. \quad (7)$$

Supposing in addition (and without loss of generality) that  $\mathbf{F}$  and  $N_\Sigma$  are independent, one can write

$$\begin{aligned} \mathbb{E}h(N_\Sigma) - \mathbb{E}h(\mathbf{F}) &= \int_0^1 \frac{d}{dt} \mathbb{E}[h_t(\mathbf{F})] dt \\ &= \frac{1}{2} \sum_{i=1}^m \int_0^1 \left\{ \frac{1}{\sqrt{t}} \mathbb{E}[\partial_i h(\sqrt{t}N_\Sigma + \sqrt{1-t}\mathbf{F})N_i] - \frac{1}{\sqrt{1-t}} \mathbb{E}[\partial_i h(\sqrt{t}N_\Sigma + \sqrt{1-t}\mathbf{F})F_i] \right\} dt \\ &= \frac{1}{2} \sum_{i,j=1}^m \int_0^1 \mathbb{E}[\partial_{ij}^2 h(\sqrt{t}N_\Sigma + \sqrt{1-t}\mathbf{F})(\mathbb{E}[N_i N_j] - \langle DF_i, u_j \rangle_{\mathfrak{H}})] dt. \end{aligned}$$

Combining the duality formula (5) with the chain rule (4) implies

$$\mathbb{E}h(N_\Sigma) - \mathbb{E}h(\mathbf{F}) = \frac{1}{2} \int_0^1 \mathbb{E}[\langle (\text{Hess } h)(\sqrt{t}N_\Sigma + \sqrt{1-t}\mathbf{F}), \Sigma - M_F \rangle_{\text{H.S.}}] dt,$$

where  $M_F$  is the random  $m \times m$  matrix given by

$$M_F(i, j) = \langle DF_i, u_j \rangle_{\mathfrak{H}}. \quad (8)$$

It is then immediate that

$$d_2(\mathbf{F}, N_\Sigma) \leq \frac{1}{2} \mathbb{E}\|M_F - \Sigma\|_{\text{H.S.}}. \quad (9)$$

Inequalities in the spirit of (9) were derived e.g. in [18] (in the context of limit theorems for homogeneous sums) and [27] (in the framework of multivariate normal approximations on the Poisson space) — see also [29] and the references therein.

### 1.3 Bounds on the 1-Wasserstein distance

For random vectors  $\mathbf{F}$  and  $N_\Sigma$  as in the previous section, we will now discuss a suitable method for assessing the quantity  $d_W(\mathbf{F}, N_\Sigma)$  defined in (2), that is, for uniformly bounding the absolute difference  $|\mathbb{E}h(\mathbf{F}) - \mathbb{E}h(N_\Sigma)|$  over all 1-Lipschitz functions  $h$  of class  $C^1$ .

Since we do not assume  $h$  to be twice differentiable, the method presented in Section 1.2 no longer works. A preferable approach is consequently the so-called ‘Malliavin-Stein method’, introduced in [16] in dimension 1, and later extended to the multivariate setting in [19]. Let us briefly recall how this works (see [17, Chapter 4 and Chapter 6] for a full discussion, and [1] for a constantly updated list of references).

Start by considering the following **Stein’s equation**, with  $h : \mathbb{R}^m \rightarrow \mathbb{R}$  given and  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  unknown:

$$\sum_{i,j=1}^m \Sigma(i, j) \partial_{ij}^2 f(\mathbf{x}) - \sum_{i=1}^m x_i \partial_i f(\mathbf{x}) = h(\mathbf{x}) - \mathbb{E}h(N_\Sigma), \quad \mathbf{x} \in \mathbb{R}^m. \quad (10)$$

When  $h \in C^1$  has bounded partial derivatives, it turns out that (10) admits a solution  $f = f_h$  of class  $C^2$  and whose second partial derivatives are bounded — see e.g. [17, Proposition 4.3.2] for a precise statement. Taking expectation with respect to the distribution of  $\mathbf{F}$  in (10) gives

$$\mathbb{E}h(\mathbf{F}) - \mathbb{E}h(N_\Sigma) = \sum_{i,j=1}^m \Sigma(i, j) \mathbb{E}[\partial_{ij}^2 f_h(\mathbf{F})] - \sum_{i=1}^m \mathbb{E}[F_i \partial_i f_h(\mathbf{F})].$$

We can apply again the duality formula (5) together with the chain rule (4), to deduce that

$$\mathbb{E}h(\mathbf{F}) - \mathbb{E}h(N_\Sigma) = \mathbb{E}[\langle (\text{Hess } f_h)(\mathbf{F}), M_F - \Sigma \rangle_{\text{H.S.}}],$$

where  $M_F$  is defined in (8). Taking the supremum over the set of all 1-Lipschitz functions  $h : \mathbb{R}^m \rightarrow \mathbb{R}$  of class  $C^1$ , we infer

$$d_W(\mathbf{F}, N_\Sigma) \leq c_1 \mathbb{E}\|M_F - \Sigma\|_{\text{H.S.}}, \quad (11)$$

with

$$c_1 = \sup_{h \in \text{Lip}(1)} \sup_{\mathbf{x} \in \mathbb{R}^m} \|(\text{Hess } f_h)(\mathbf{x})\|_{\text{H.S.}} \leq \sqrt{m} \|\Sigma^{-1}\|_{\text{op}} \|\Sigma\|_{\text{op}}^{1/2}, \quad (12)$$

and  $\|\cdot\|_{\text{op}}$  is the operator norm for  $m \times m$  matrices. The estimate (11) is the main result of [19] (see also [17, Theorem 6.1.1]), whereas a self-contained proof of (12) can be found in [17, Proposition 4.3.2].

#### 1.4 Main results: bounds on the convex distance

The principal aim of the present paper is to address the following natural question: *can one obtain a bound similar to (11) for distances based on **non-smooth** test functions  $h : \mathbb{R}^m \rightarrow \mathbb{R}$ , such as e.g. indicator functions of measurable convex subsets of  $\mathbb{R}^m$ ?*

If  $h$  is such an indicator function, then we recall e.g. from [29, Lemma 2.2] that, for all  $t \in (0, 1)$ ,

$$|\mathbb{E}h(\mathbf{F}) - \mathbb{E}h(N_\Sigma)| \leq \frac{4}{3} |\mathbb{E}h_t(\mathbf{F}) - \mathbb{E}h_t(N_\Sigma)| + \frac{20m}{\sqrt{2}} \frac{\sqrt{t}}{1-t},$$

where  $h_t$  stands for the mollification at level  $\sqrt{t}$  of  $h$ , as defined in (7). Let  $f_t = f_{h_t}$  be the solution of the Stein's equation (10) associated with  $h = h_t$ . In [6] (see also [29]), it is shown that

$$\max_{1 \leq i, j \leq m} \sup_{\mathbf{x} \in \mathbb{R}^m} |\partial_{ij}^2 f_t(\mathbf{x})| \leq c_2 |\log t|,$$

with  $c_2 = c_2(m, \Sigma)$  a constant depending only on  $m$  and  $\Sigma$ . Combining such an estimate with (7) yields the existence of a constant  $c_3 = c_3(m, \Sigma) > 0$  such that

$$|\mathbb{E}h(\mathbf{F}) - \mathbb{E}h(N_\Sigma)| \leq c_3 \left( \mathbb{E}\|M_F - \Sigma\|_{\text{H.S.}} |\log t| + \frac{\sqrt{t}}{1-t} \right). \quad (13)$$

From (13), it is straightforward to deduce the existence of  $c_4 = c_4(m, \Sigma) > 0$  such that

$$d_c(\mathbf{F}, N_\Sigma) \leq c_4 \mathbb{E}\|M_F - \Sigma\|_{\text{H.S.}} |\log \{\mathbb{E}\|M_F - \Sigma\|_{\text{H.S.}}\}|. \quad (14)$$

Comparing (14) with (9) and (11) shows that such a strategy yields a bound on  $d_c(\mathbf{F}, N_\Sigma)$  differing from those deduced above for the distances  $d_2$  and  $d_W$  by an additional logarithmic factor. See also [11, 20] for more inequalities analogous to (14) — that is, displaying a multiplicative logarithmic factor — related respectively to the (multivariate) Kolmogorov and total variation distances.

In this paper, we will show that one can actually remove the redundant logarithmic factor on the right-hand side of (14), thus yielding a bound on  $d_c(\mathbf{F}, N_\Sigma)$  that is commensurate to (9) and (11) (with moreover an explicit multiplicative constant). Our main result is the following:

**Theorem 1.2.** Let  $\mathbf{F} = (F_1, \dots, F_m) = (\delta(u_1), \dots, \delta(u_m))$  be a vector in  $\mathbb{R}^m$  of centered random variables such that  $u_i \in \text{Dom}(\delta)$ , for  $i = 1, \dots, m$ . Let also  $N_\Sigma = (N_1, \dots, N_m)$  be a centered Gaussian vector with invertible  $m \times m$  covariance matrix  $\Sigma = \{\Sigma(i, j)\}$ . Then

$$d_c(\mathbf{F}, N_\Sigma) \leq 402 \left( \left\| \Sigma^{-1} \right\|_{\text{op}}^{3/2} + 1 \right) m^{41/24} \mathbb{E} \|M_F - \Sigma\|_{\text{H.S.}},$$

with  $M_F$  defined in (8).

As anticipated, to prove Theorem 1.2, we shall combine the somewhat classical smoothing estimate (13) with a remarkable bound by Schulte and Yukich [29].

## 1.5 Applications

We illustrate the use of Theorem 1.2 by developing two examples in full detail.

*Quantitative fourth moment theorems.* A **fourth moment theorem** (FMT) is a mathematical statement implying that a given sequence of centered and normalized random variables converges in distribution to a Gaussian limit, as soon as the corresponding sequence of fourth moments converges to 3 (that is, to the fourth moment of the standard Gaussian distribution). Distinguished examples of FMTs are e.g. de Jong's theorem for degenerate  $U$ -statistics (see [8, 9]) as well as the CLTs for multiple Wiener-Itô integrals proved in [25, 26]; the reader is referred to the webpage [1] for a list (composed of several hundreds of papers) of applications and extensions of such results, as well as to the lecture notes [31] for a modern discussion of their relevance in mathematical physics. Our first application of Theorem 1.2 is a quantitative multivariate fourth moment theorem for a vector of multiple Wiener-Itô integrals, considerably extending the qualitative multivariate results proved in [26]. Note that such a result was already obtained by Nourdin and Rosiński [22, Theorem 4.3] for the 1-Wasserstein distance  $d_W$ . Thanks to Theorem 1.2, it is not difficult to generalize their result to the  $d_c$  metric.

**Corollary 1.3.** Fix  $m \geq 1$  as well as  $q_1, \dots, q_m \geq 1$ . Let  $\mathbf{F} = (F_1, \dots, F_m)$  where  $F_i = I_{q_i}(f_{q_i})$  with  $f_{q_i} \in \mathfrak{H}^{\odot q_i}$ . Let  $N_\Sigma$  be a centered Gaussian vector with covariance matrix  $\Sigma = (\mathbb{E} F_i F_j)_{i,j \in [m]}$ . Then

$$d_c(\mathbf{F}, N_\Sigma) \leq 402 \left( \left\| \Sigma^{-1} \right\|_{\text{op}}^{3/2} + 1 \right) m^{41/24} \sqrt{\mathbb{E} \|\mathbf{F}\|^4 - \mathbb{E} \|N_\Sigma\|^4}.$$

In particular, for a vector  $\mathbf{F}$  of multiple Wiener-Itô integrals to be close in the convex distance to a centered Gaussian vector  $N_\Sigma$  with matching covariance matrix, it is enough that  $\mathbb{E} \|\mathbf{F}\|^4 \approx \mathbb{E} \|N_\Sigma\|^4$ .

*The multivariate Breuer-Major theorem.* The second example concerns the convergence towards a Brownian motion occurring in the **Breuer-Major theorem** proved in [5]. Let us briefly recall this fundamental result (see [17, Chapter 7] for an introduction to the subject, as well as [15, 7] for recent advances in a functional setting). Let  $\{G_k : k \in \mathbb{Z}\}$  be a centered Gaussian stationary sequence with  $\rho(j - k) = \mathbb{E}[G_j G_k]$  and  $\rho(0) = 1$ ; in particular,  $G_k \sim N(0, 1)$  for all  $k$ . Let  $\varphi \in L^2(\mathbb{R}, \gamma)$  where  $\gamma(dx) = (2\pi)^{-1/2} e^{-x^2/2} dx$  denotes the standard Gaussian measure on  $\mathbb{R}$ . Since the Hermite polynomials  $\{H_k : k \geq 0\}$  form an orthonormal basis of  $L^2(\mathbb{R}, \gamma)$ , one has

$$\varphi = \sum_{k \geq d} a_k H_k,$$

with  $d \in \mathbb{N}$  and  $a_d \neq 0$ . The index  $d$  is known as the Hermite rank of  $\varphi \in L^2(\mathbb{R}, \gamma)$ . Suppose in addition that  $\int_{\mathbb{R}} \varphi d\gamma = \mathbb{E}[\varphi(G_0)] = 0$ , that is, suppose  $d \geq 1$ . The Breuer-Major theorem [5] states the following: if  $\sum_{k \in \mathbb{Z}} |\rho(k)|^d < \infty$ , then

$$\left\{ \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} \varphi(G_k) : t \geq 0 \right\} \xrightarrow{f.d.d.} \{\sigma W(t) : t \geq 0\} \quad (15)$$

where  $W$  is a standard Brownian motion,  $\xrightarrow{f.d.d.}$  indicates convergence in the sense of finite-dimensional distributions, and

$$\sigma^2 := \sum_{k \geq d} a_k^2 k! \sum_{j \in \mathbb{Z}} \rho(j)^k \in [0, \infty),$$

(That  $\sigma^2$  is a well-defined positive real number is part of the conclusion.) We refer to our note [21] and references therein for results on the rate of convergence in the total variation distance for one-dimensional marginal distributions (that is, in dimension 1). We intend to apply Theorem 1.2 to address the rate of convergence for the following multivariate CLT implied by (15): for every  $0 = t_0 < t_1 < \dots < t_m = T < \infty$ ,

$$\left( \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt_1 \rfloor} \varphi(G_k), \dots, \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt_m \rfloor} \varphi(G_k) \right) \xrightarrow{d} N(0, \Sigma(t_1, \dots, t_m))$$

where  $\xrightarrow{d}$  indicates converges in distribution, and  $N(0, \Sigma(t_1, \dots, t_m))$  is a  $m$ -dimensional centred Gaussian vector with covariance  $\Sigma(t_1, \dots, t_m)$  having entries  $\sigma^2 t_i \wedge t_j$ ,  $i, j = 1, \dots, m$ . Notice that for any  $m \times m$  invertible matrix  $A$ ,

$$d_c(\mathbf{F}, \mathbf{G}) = d_c(A\mathbf{F}, A\mathbf{G}).$$

Therefore, choosing  $A$  appropriately, it suffices to consider the vector  $\mathbf{F}_n = (F_{n,1}, \dots, F_{n,m})$  with

$$F_{n,i} = \frac{1}{\sqrt{n}} \sum_{k=\lfloor nt_{i-1} \rfloor + 1}^{\lfloor nt_i \rfloor} \varphi(G_k), \quad i \in [m]$$

and obtain the rate of convergence for

$$\mathbf{F}_n \xrightarrow{d} N(0, \sigma^2 \text{Diag}(t_1 - t_0, \dots, t_m - t_{m-1})) =: N_{\Sigma}. \quad (16)$$

The following result provides a quantitative version of this CLT with respect to the distance  $d_c$ . Recall from [21] that the minimal regularity assumption over  $\varphi$  for obtaining rates of convergence via the Malliavin-Stein method is that  $\varphi \in \mathbb{D}^{1,4}(\mathbb{R}, \gamma)$ , meaning that  $\varphi$  is absolutely continuous and both  $\varphi$  and its derivative  $\varphi'$  belong to  $L^4(\mathbb{R}, \gamma)$ . We say that  $\varphi$  is 2-sparse if its expansion in Hermite polynomials does not have consecutive non-zero coefficients. In particular, even functions are 2-sparse.

**Corollary 1.4.** *Let  $\mathbf{F}_n$  and  $N_{\Sigma}$  be given in (16). Suppose that  $\varphi \in \mathbb{D}^{1,4}(\mathbb{R}, \gamma)$  with Hermite rank  $d \geq 1$ . Then,*

i) *There exists a constant  $C$  depending only on  $\varphi, m, \Sigma$  such that for each  $n \in \mathbb{N}$ ,*

$$d_c(\mathbf{F}_n, N_{\Sigma}) \leq C \sum_{i,j=1}^m |\Sigma(i,j) - \mathbb{E}[F_i F_j]| + Cn^{-\frac{1}{2}} \left( \sum_{|k| < n} |\rho(k)| \right)^{\frac{3}{2}}.$$

ii) If  $d = 2$ ,  $\varphi$  is 2-sparse and  $b \in [1, 2]$ , then there exists a constant  $C$  depending only on  $\varphi, m, \Sigma$  such that for each  $n \in \mathbb{N}$ ,

$$d_c(\mathbf{F}_n, N_\Sigma) \leq C \sum_{i,j=1}^m |\Sigma(i,j) - \mathbb{E}[F_i F_j]| + C n^{-(\frac{1}{b} - \frac{1}{2})} \left( \sum_{|k| < n} |\rho(k)|^2 \right)^{\frac{1}{2}} \left( \sum_{|k| < n} |\rho(k)|^b \right)^{\frac{1}{b}}.$$

iii) If  $d = 2$ ,  $\varphi$  is 2-sparse, and  $\sum_{k \in \mathbb{Z}} |\rho(k)|^2 < \infty$ , then as  $n \rightarrow \infty$ ,

$$d_c(\mathbf{F}_n, N_\Sigma) \rightarrow 0.$$

The rest of the note is organized as follows. The proof of Theorem 1.2 is given in Section 2.1, Corollary 1.3 in Section 2.2, Corollary 1.4 in Section 2.3. We use  $C$  to denote a generic constant whose value may change from line to line.

## 1.6 Acknowledgments

I. Nourdin is supported by the FNR grant APOGee (R-AGR-3585-10) at Luxembourg University; G. Peccati is supported by the FNR grant FoRGES (R-AGR-3376-10) at Luxembourg University; X. Yang is supported by the FNR Grant MISSILE (R-AGR-3410-12-Z) at Luxembourg and Singapore Universities.

## 2 Proofs

### 2.1 Proof of Theorem 1.2

We divide the proof into several steps.

Step 1 (smoothing). For any bounded and measurable  $h$  and  $t \in (0, 1)$ , recall its mollification at level  $\sqrt{t}$  from (7). Then it is plain that  $h_t$  is  $C^\infty$  with bounded derivatives of all orders and the solution to (10) with  $h = h_t$  is given by

$$f_t(\mathbf{x}) := -\frac{1}{2} \int_t^1 \frac{1}{1-s} (\mathbb{E}[h_t(\sqrt{s}N_\Sigma + \sqrt{1-s}\mathbf{x})] - \mathbb{E}[h_t(N_\Sigma)]) ds,$$

see [29, p.12]. Finally, recall from e.g. [29, Lemma 2.2] that, for any  $t \in (0, 1)$ ,

$$d_c(\mathbf{F}, N_\Sigma) \leq \frac{4}{3} \sup_{h \in \mathcal{I}_m} |\mathbb{E}h_t(\mathbf{F}) - \mathbb{E}h_t(N_\Sigma)| + \frac{20m}{\sqrt{2}} \frac{\sqrt{t}}{1-t}.$$

Step 2 (integration by parts). An integration by parts by (5) and (4) (see [17, Chapter 4] for more details), together with Cauchy-Schwarz's inequality, implies,

$$\begin{aligned} |\mathbb{E}h_t(\mathbf{F}) - \mathbb{E}h_t(N_\Sigma)| &= \left| \mathbb{E} \sum_{i,j=1}^d \Sigma(i,j) \partial_{ij}^2 f_t(\mathbf{F}) - \mathbb{E} \sum_{k=1}^d F_k \partial_k f_t(\mathbf{F}) \right| \\ &= \left| \mathbb{E} \sum_{i,j=1}^d (\Sigma(i,j) - \langle DF_i, u_j \rangle) \partial_{ij}^2 f(\mathbf{F}) \right| \\ &\leq \sqrt{\sum_{i,j=1}^d \mathbb{E}[(\Sigma(i,j) - \langle DF_i, u_j \rangle)^2]} \sqrt{\sum_{i,j=1}^d \mathbb{E}[\partial_{ij}^2 f(\mathbf{F})^2]}. \end{aligned} \tag{17}$$

The following remarkable estimate is due to M. Schulte and J. Yukich.

**Lemma 2.1** (Proposition 2.3 in [29]). *Let  $\mathbf{Y}$  be an  $\mathbb{R}^m$ -valued random vector and  $\Sigma$  be an invertible  $m \times m$  covariance matrix. Then,*

$$\sup_{h \in \mathcal{I}_m} \mathbb{E} \sum_{i,j=1}^m |\partial_{ij}^2 f_t(\mathbf{Y})|^2 \leq \left\| \Sigma^{-1} \right\|_{\text{op}}^2 \left( m^2 (\log t)^2 d_c(\mathbf{Y}, N_\Sigma) + 530m^{17/6} \right).$$

where the left-hand side depends on  $h$  through the function  $f_t$  solving Stein's equation with test function  $h_t$  given by (7).

*Remark 2.2.* Lemma 2.1 improves upon the uniform bound (see [6] or [29])

$$|\partial_{ij}^2 f_t(\mathbf{x})| \leq C(m, \Sigma) \|h\|_\infty |\log t|,$$

when some a priori estimate on  $d_c(\mathbf{Y}, N_\Sigma)$  is available.

As consequence,

$$\begin{aligned} & |\mathbb{E} h_t(\mathbf{F}) - \mathbb{E} h_t(N_\Sigma)| \\ & \leq \left\| \Sigma^{-1} \right\|_{\text{op}} \left( m |\log t| d_c(\mathbf{F}, N_\Sigma)^{1/2} + 24m^{17/12} \right) \sqrt{\sum_{i,j=1}^d \mathbb{E}[(\Sigma(i,j) - \langle DF_i, u_j \rangle)^2]}. \end{aligned}$$

Letting

$$\begin{aligned} \kappa &= d_c(\mathbf{F}, N_\Sigma), \\ \gamma &= \sqrt{\sum_{i,j=1}^d \mathbb{E}[(\Sigma(i,j) - \langle DF_i, u_j \rangle)^2]}, \end{aligned}$$

we have thus established

$$\kappa \leq \frac{4}{3} \left\| \Sigma^{-1} \right\|_{\text{op}} (m |\log t| \sqrt{\kappa} + 24m^{17/12}) \gamma + \frac{20m}{\sqrt{2}} \frac{\sqrt{t}}{1-t}. \quad (18)$$

Step 3 (exploiting the recursive inequality). Suppose that  $\gamma \leq 1/e$ , otherwise the bound we intend to prove holds already (although not being informative). Let  $t = \gamma^2$ . Using the fact that  $\kappa \leq 1$  for the  $\kappa$  on the right-hand side of the (18), one has

$$\kappa \leq \frac{4}{3} \left\| \Sigma^{-1} \right\|_{\text{op}} (2m |\log \gamma| + 24m^{17/12}) \gamma + 20\sqrt{2}m\gamma.$$

Therefore,

$$|\log \gamma| \sqrt{\kappa} \leq \left\| \Sigma^{-1} \right\|_{\text{op}}^{1/2} \sqrt{\frac{8m}{3}} \gamma^{1/2} |\log \gamma|^{3/2} + \left( \left\| \Sigma^{-1} \right\|_{\text{op}} 32m^{17/12} + 20\sqrt{2}m \right)^{1/2} \gamma^{1/2} |\log \gamma|.$$

Since  $\sup_{x \in (0, 1/e]} x^{1/2} |\log x|^{3/2} \leq 4$ , one has

$$\begin{aligned} |\log \gamma| \sqrt{\kappa} &\leq \frac{8\sqrt{6}}{3} \left\| \Sigma^{-1} \right\|_{\text{op}}^{1/2} m^{1/2} + 16\sqrt{2} \left\| \Sigma^{-1} \right\|_{\text{op}}^{1/2} m^{17/24} + 8\sqrt{10}m^{1/2} \\ &\leq 58 \left( \left\| \Sigma^{-1} \right\|_{\text{op}}^{1/2} + 1 \right) m^{17/24}. \end{aligned}$$

Hence, putting the estimate back into (18) with  $t = \gamma^2$  gives

$$\begin{aligned}\kappa &\leq \frac{4}{3} \left\| \Sigma^{-1} \right\|_{\text{op}} \left( 2m \left( 58 \left( \left\| \Sigma^{-1} \right\|_{\text{op}}^{1/2} + 1 \right) m^{17/24} \right) + 24m^{17/12} \right) \gamma + 20\sqrt{2}m\gamma \\ &\leq \left( \frac{4}{3} \times 140 \times 2 \left( \left\| \Sigma^{-1} \right\|_{\text{op}}^{3/2} + 1 \right) m^{41/24} + 20\sqrt{2}m \right) \gamma \\ &\leq 402 \left( \left\| \Sigma^{-1} \right\|_{\text{op}}^{3/2} + 1 \right) m^{41/24} \gamma.\end{aligned}$$

The proof is complete.

## 2.2 Proof of Corollary 1.3

We will obtain the desired conclusion as a direct application of Theorem 1.2 with  $u_i = -DL^{-1}F_i$ , see (6). Indeed, recall that by Step 2 of [22, Proof of Theorem 4.3], for any  $i, j \in [m]$ ,

$$\mathbb{E}[(\mathbb{E}[F_i F_j] - \langle DF_i, -DL^{-1}F_j \rangle)^2] \leq \text{Cov}(F_i^2, F_j^2) - 2\mathbb{E}[F_i F_j]^2.$$

On the other hand, Step 3 of [22, Proof of Theorem 4.3] shows that

$$\sum_{i,j=1}^m \text{Cov}(F_i^2, F_j^2) - 2\mathbb{E}[F_i F_j]^2 = \mathbb{E} \|\mathbf{F}\|^4 - \mathbb{E} \|N_\Sigma\|^4.$$

Plugging these estimates into Theorem 1.2 gives the result.

## 2.3 Proof of Corollary 1.4

We follow closely the arguments of [21] and assume without loss of generality that  $T = 1$ . First, one can embed the Gaussian sequence in the statement in an isonormal Gaussian process  $\{X(h) : h \in \mathfrak{H}\}$ , in such a way that

$$\{G_k : k \in \mathbb{Z}\} \stackrel{d}{=} \{X(e_k) : k \in \mathbb{Z}\},$$

for some appropriate family  $\{e_k\} \subset \mathfrak{H}$  verifying  $\langle e_j, e_k \rangle_{\mathfrak{H}} = \rho(k - j)$  for all  $j, k$ . For  $\varphi = \sum_{\ell \geq d} a_\ell H_\ell \in L^2(\mathbb{R}, \gamma)$ , we define the shift mapping  $\varphi_1 := \sum_{\ell \geq 1} a_\ell H_{\ell-1}$  and set

$$u_{n,i} := \frac{1}{\sqrt{n}} \sum_{m=\lfloor nt_{i-1} \rfloor + 1}^{\lfloor nt_i \rfloor} \varphi_1(G_m) e_m, \quad i \in [m].$$

Then, standard computations using (6) leads to

$$\delta(u_{n,i}) = F_{n,i}. \tag{19}$$

Applying Theorem 1.2 and the triangle inequality implies that

$$d_c(\mathbf{F}, N_\Sigma) \leq C \sum_{i,j=1}^m |\Sigma(i, j) - \mathbb{E}[F_i F_j]| + C \sqrt{\sum_{i,j=1}^m \text{Var}(\langle DF_{n,i}, u_{n,j} \rangle)} =: I_1 + I_2.$$

Note that, by the chain rule and the relation  $D(G_k) = e_k$ ,

$$\langle DF_{n,i}, u_{n,j} \rangle_{\mathfrak{H}} = \frac{1}{n} \sum_{k \sim t_i} \sum_{\ell \sim t_j} \varphi'(G_k) \varphi_1(G_\ell) \rho(k - \ell),$$

where  $k \sim t_i$  means that the sum is taken over  $k \in \{\lfloor nt_{i-1} \rfloor + 1, \dots, \lfloor nt_i \rfloor\}$ , and similarly for the symbol  $\ell \sim t_j$ . Hence,

$$\begin{aligned} & \mathbb{V}\text{ar}(\langle DF_{n,i}, u_{n,j} \rangle_{\mathfrak{H}}) \\ &= \frac{1}{n^2} \sum_{k \sim t_i} \sum_{\ell \sim t_j} \sum_{k' \sim t_i} \sum_{\ell' \sim t_j} \mathbb{C}\text{ov}(\varphi'(X_k) \varphi_1(X_\ell), \varphi'(X_{k'}) \varphi_1(X_{\ell'})) \rho(k - \ell) \rho(k' - \ell') \\ &\leq \frac{1}{n^2} \sum_{k,k',\ell,\ell'=1}^n \mathbb{C}\text{ov}(\varphi'(X_k) \varphi_1(X_\ell), \varphi'(X_{k'}) \varphi_1(X_{\ell'})) \rho(k - \ell) \rho(k' - \ell'). \end{aligned} \quad (20)$$

The variance is bounded because of the assumption that  $\varphi \in \mathbb{D}^{1,4}$ . Once (20) is in place, one can apply Gebelein's inequality as in [21]. In particular, one infers that (see [21, Proposition 3.4])

$$d_c(\mathbf{F}_n, N_\Sigma) \leq C \sqrt{\frac{1}{n^2} \sum_{i,j,k,\ell=0}^{n-1} |\rho(j - k) \rho(i - j) \rho(k - \ell)|}.$$

If, in addition,  $\varphi$  is 2-sparse, then

$$d_c(\mathbf{F}_n, N_\Sigma) \leq C \sqrt{\frac{1}{n^2} \sum_{i,j,k,\ell=0}^{n-1} |\rho(j - k)^2 \rho(i - j) \rho(k - \ell)|}.$$

Items i)-ii) now follow from these inequalities, as shown in [21]; we include a proof for completeness. Applying twice Young's inequality for convolutions, one has

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

yielding Item i). 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

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

For  $b \in [1, 2]$ , we have

$$\begin{aligned}
\langle \rho_n * 1_n, \rho_n * \rho_n^2 \rangle_{\ell^2(\mathbb{Z})} &\leq \|\rho_n * 1_n\|_{\ell^{\frac{b}{b-1}}(\mathbb{Z})} \|\rho_n * \rho_n^2\|_{\ell^b(\mathbb{Z})} \\
&\leq \|\rho_n\|_{\ell^b} \|1_n\|_{\ell^{\frac{b}{2b-2}}(\mathbb{Z})} \|\rho_n\|_{\ell^b(\mathbb{Z})} \|\rho_n^2\|_{\ell^1(\mathbb{Z})} \\
&= n^{\frac{2b-2}{b}} \|\rho_n^2\|_{\ell^1(\mathbb{Z})} \|\rho_n\|_{\ell^b(\mathbb{Z})}^2,
\end{aligned} \tag{21}$$

yielding Item ii). Now we move to the proof of Item iii). Notice that taking  $b = 2$  for the right-hand side of (21), together with an application of Young's inequality, yields that

$$\langle \rho_n * 1_n, \rho_n * \rho_n^2 \rangle_{\ell^2(\mathbb{Z})} \leq \|\rho_n * 1_n\|_{\ell^2(\mathbb{Z})} \|\rho_n\|_{\ell^2(\mathbb{Z})}^3.$$

Thus,

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

To proceed, we handle the convolution involving  $1_n$  a bit differently. Set

$$\begin{aligned}
\tilde{\rho}_n(k) &= \rho(k) 1_{N \leq |k| < n}, \\
\widehat{\rho}_n(k) &= \rho(k) 1_{|k| \leq N}
\end{aligned}$$

so that  $\rho_n = \tilde{\rho}_n + \widehat{\rho}_n$ . One has

$$\begin{aligned}
\frac{1}{n} \|\rho_n * 1_n\|_{\ell^2(\mathbb{Z})} &\leq \frac{1}{n} \|\tilde{\rho}_n\|_{\ell^2(\mathbb{Z})} \|1_n\|_{\ell^1(\mathbb{Z})} + \frac{1}{n} \|\widehat{\rho}_n\|_{\ell^1(\mathbb{Z})} \|1_n\|_{\ell^2(\mathbb{Z})} \\
&\leq \left( \sum_{N \leq |k| < n} \rho(k)^2 \right)^{1/2} + (2N+1)n^{-1/2},
\end{aligned}$$

from which Item iii) follows. The proof is complete.

## A Proof and discussion of relation (3)

Inequality (3) is a direct consequence of the following statement, whose proof exploits a strategy already adopted in [2, Proof of Theorem 3.1].

**Proposition A.1.** *Fix  $m \geq 1$ , and let  $N_\Sigma$  denote a  $m$ -dimensional centered Gaussian vector with invertible covariance matrix  $\Sigma$ . Then, for any  $m$ -dimensional random vector  $\mathbf{F}$  one has that*

$$d_c(\mathbf{F}, N_\Sigma) \leq 2\sqrt{2} \Gamma(\Sigma)^{1/2} d_W(\mathbf{F}, N_\Sigma)^{1/2}, \tag{22}$$

where  $\Gamma(\Sigma)$  is the isoperimetric constant defined by

$$\Gamma(\Sigma) := \sup_{Q, \epsilon > 0} \frac{\mathbb{P}(N_\Sigma \in Q^\epsilon) - \mathbb{P}(N_\Sigma \in Q)}{\epsilon},$$

where  $Q$  ranges over all Borel measurable convex subsets of  $\mathbb{R}^m$ , and  $Q^\epsilon$  indicates the set of all elements of  $\mathbb{R}^m$  whose Euclidean distance from  $Q$  does not exceed  $\epsilon$ .

*Remark A.2.* In [14] it is proved that, for some absolute constants  $0 < c < C < \infty$ ,

$$c\sqrt{\|\Sigma\|_{\text{H.S.}}} \leq \Gamma(\Sigma) \leq C\sqrt{\|\Sigma\|_{\text{H.S.}}},$$

where  $\|\cdot\|_{\text{H.S.}}$  stands as above for the Hilbert-Schmidt norm. When  $\Sigma = I_m$  (identity matrix), one has also the well-known estimate  $\Gamma(I_m) \leq 4m^{1/4}$  (see [3]), as well as Nazarov's upper and lower bounds

$$e^{-5/4} \leq \liminf_m \frac{\Gamma(I_m)}{m^{1/4}} \leq \limsup_m \frac{\Gamma(I_m)}{m^{1/4}} \leq (2\pi)^{-1/4} < 0.64,$$

see [14, p. 170]. In [28, Theorem 1.2], it is proved that Nazarov's upper bound can be reduced from 0.64 to 0.59; see also [4] for related computations in the framework of the multivariate CLT.

*Proof of Proposition A.1.* We can assume that  $\mathbf{F}$  and  $N_\Sigma$  are defined on a common probability space, and that  $\mathbb{E}\|\mathbf{F} - N_\Sigma\|_{\mathbb{R}^m} = d_W(\mathbf{F}, N_\Sigma)$ . Fix a convex set  $Q$ , as well as  $\epsilon > 0$ . We have that

$$\begin{aligned} \mathbb{P}[\mathbf{F} \in Q] - \mathbb{P}[N_\Sigma \in Q] &\leq \mathbb{P}[\mathbf{F} \in Q, \|\mathbf{F} - N_\Sigma\|_{\mathbb{R}^m} \leq \epsilon] - \mathbb{P}[N_\Sigma \in Q] + \epsilon^{-1}\mathbb{E}[\|\mathbf{F} - N_\Sigma\|_{\mathbb{R}^m}] \\ &\leq \mathbb{P}[N_\Sigma \in Q^\epsilon] - \mathbb{P}[N_\Sigma \in Q] + \epsilon^{-1}d_W(\mathbf{F}, N_\Sigma) \\ &\leq \Gamma(\Sigma)\epsilon + \epsilon^{-1}d_W(\mathbf{F}, N_\Sigma). \end{aligned}$$

On the other hand, defining  $Q^{-\epsilon}$  as the set of those  $y \in Q$  such that the closed ball with radius  $\epsilon$  centered at  $y$  is contained in  $Q$ ,

$$\begin{aligned} \mathbb{P}[N_\Sigma \in Q] - \mathbb{P}[\mathbf{F} \in Q] &\leq \mathbb{P}[N_\Sigma \in Q, \|\mathbf{F} - N_\Sigma\|_{\mathbb{R}^m} \leq \epsilon] - \mathbb{P}[\mathbf{F} \in Q^{-\epsilon}] + \epsilon^{-1}\mathbb{E}[\|\mathbf{F} - N_\Sigma\|] \\ &\leq \Gamma(\Sigma)2\epsilon + \epsilon^{-1}d_W(\mathbf{F}, N_\Sigma), \end{aligned}$$

where we have used the inequality

$$\mathbb{P}[N_\Sigma \in Q, \|\mathbf{F} - N_\Sigma\|_{\mathbb{R}^m} \leq \epsilon] - \mathbb{P}[\mathbf{F} \in Q^{-\epsilon}] \leq \mathbb{P}[N_\Sigma \in Q] - \mathbb{P}[N_\Sigma \in Q^{-2\epsilon}].$$

The conclusion follows from a standard optimisation in  $\epsilon$ . □

*Remark A.3.* Fix  $m \geq 1$ , and let  $\mathcal{R}_m$  be the collection of all hyper-rectangles of the type  $R = (-\infty, t_1] \times \cdots \times (-\infty, t_m]$ . In [2, Theorem 3.1] it is proved that, if  $N$  is a  $m$ -dimensional centered Gaussian vector with identity covariance matrix and  $\mathbf{F}$  is any  $m$ -dimensional random vector, then

$$\sup_{R \in \mathcal{R}_m} |\mathbb{P}[\mathbf{F} \in R] - \mathbb{P}[N \in R]| \leq 3(\log m)^{1/4}d_W(\mathbf{F}, N)^{1/2}. \quad (23)$$

The left-hand side of the previous inequality is usually referred to as the **Kolmogorov distance** between the distributions of  $\mathbf{F}$  and  $N$ . The presence of the factor  $(\log m)^{1/4}$  is consistent with the fact that, for the standard Gaussian measure on  $\mathbb{R}^m$ , the isoperimetric constant associated with all hyper-rectangles of  $\mathbb{R}^m$  is bounded from above by  $\sqrt{\log m}$ , see [3, 14]. An estimate analogous to (23) is established by different methods in [12, Corollary 3.1].

## References

- [1] Malliavin-Stein approach: a webpage maintained by Ivan Nourdin. Adress: <https://sites.google.com/site/malliavinstein/home>
- [2] E. Azmoodeh, G. Peccati and G. Poly (2016): The law of iterated logarithm for subordinated Gaussian sequences: uniform Wasserstein bounds. *ALEA* **13**, 659–686.
- [3] K. Ball (1993): The reverse isoperimetric problem for the Gaussian measure. *Discrete Comput. Geom.*, **10**, no. 4, 4111–420.
- [4] V. Bentkus (2003): On the dependence of the Berry–Esseen bound on dimension. *Journal of Statistical Planning and Inference* **113**, 385–402.
- [5] P. Breuer and P. Major (1983): Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Anal.* **13**, no. 3, pp. 425–441.
- [6] L.H.Y. Chen, L. Goldstein and Q.-M. Shao (2011): *Normal approximation by Stein’s method*. Probability and its Applications (New York). Springer, Heidelberg.
- [7] S. Campese, I. Nourdin and D. Nualart (2020+): Continuous Breuer-Major theorem: tightness and non-stationarity. *Ann. Probab.*, to appear.
- [8] P. de Jong (1990): A central limit theorem for generalized multilinear forms. *J. Multivariate Anal.* **34**(2), 275–289.
- [9] Ch. Döbler and G. Peccati (2017): Quantitative de Jong theorems in any dimension. *Electron. J. Probab.* **22**(2), 1–35.
- [10] F. Götze (1991): On the rate of convergence in the multivariate CLT. *Ann. Probab.*, **19**(2), 724–739.
- [11] Y.T. Kim and H.S. Park (2015): Kolmogorov distance for multivariate normal approximation. *Korean J. Math.* **23**, no. 1, pp. 1–10.
- [12] Y. Koike (2019): High-dimensional central limit theorems for homogeneous sums. Preprint.
- [13] S.V. Nagaev (1976). An estimate of the remainder term in the multidimensional central limit theorem. In: *Proceedings of the Third Japan–USSR Symposium on Probability Theory (Tashkent, 1975)*. Lecture Notes in Math. **550**, 419–438. Berlin: Springer.
- [14] F. Nazarov (2003): On the Maximal Perimeter of a Convex Set in  $\mathbb{R}^n$  with Respect to a Gaussian Measure. In: Milman V.D., Schechtman G. (eds) *Geometric Aspects of Functional Analysis*. Lecture Notes in Mathematics, vol. **1807**, Springer, Berlin, Heidelberg, 169–187.
- [15] I. Nourdin and D. Nualart (2020+): The functional Breuer-Major theorem. *Probab. Th. Related Fields*, to appear.
- [16] I. Nourdin and G. Peccati (2009): Stein’s method on Wiener chaos. *Probab. Th. Related Fields* **145**, pp. 75–118.

- [17] 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.
- [18] I. Nourdin, G. Peccati and G. Reinert (2010): Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. *Ann. Probab.* **38**(5), 1947–1985.
- [19] I. Nourdin, G. Peccati and A. Réveillac (2010): Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Inst. Henri Poincaré Probab. Stat.* **46**(1), pp. 45–58.
- [20] I. Nourdin, G. Peccati and Y. Swan (2014): Entropy and the fourth moment phenomenon. *J. Funct. Anal.* **266**, no. 5, pp. 3170–3207.
- [21] I. Nourdin, G. Peccati and X. Yang (2019): Berry-Esseen bounds in the Breuer-Major CLT and Gebelein's inequality. *Electron. Commun. Probab.* **24**, Paper No. 34, 12 pp.
- [22] I. Nourdin and J. Rosiński (2014): Asymptotic independence of multiple Wiener-Itô integrals and the resulting limit laws. *Ann. Probab.* **42**, no. 2, pp. 497–526.
- [23] D. Nualart (2006). *The Malliavin calculus and related topics*. Springer-Verlag, Berlin, second edition.
- [24] D. Nualart and E. Nualart (2018). *Introduction to Malliavin calculus*. Cambridge University Press.
- [25] D. Nualart and G. Peccati (2005): Central limit theorems for sequences of multiple stochastic integrals, *Ann. Probab.* **33**(1), 177–193.
- [26] G. Peccati and C.A. Tudor (2004): Gaussian limits for vector-valued multiple stochastic integrals. *Séminaire de Probabilités XXXVIII*, 247–262.
- [27] G. Peccati and C. Zheng (2010): Multi-dimensional Gaussian fluctuations on the Poisson space. *Electron. J. Probab.* **15**, 1487–1527.
- [28] M. Raic (2019): A multivariate Berry–Esseen theorem with explicit constants. *Bernoulli* **25**(4A), 2824–2853.
- [29] M. Schulte and J.E. Yukich (2019): Multivariate second order Poincaré inequalities for Poisson functionals. *Electron. J. Probab.* **24**, no. 130, pp. 1–42.
- [30] C. Villani (2009). *Optimal transport. Old and new*. Grundlehren der mathematischen Wissenschaften **338**. Springer: Berlin.
- [31] N. Zygouras (2019). Discrete stochastic analysis. Lecture notes available on the webpage: <https://warwick.ac.uk/fac/sci/statistics/staff/academic-research/zygouras/>