

Fisher Information and the Fourth Moment Theorem

Ivan Nourdin* and David Nualart†

December 19, 2013

Abstract

Using a representation of the score function by means of the divergence operator we exhibit a sufficient condition, in terms of the negative moments of the norm of the Malliavin derivative, under which convergence in Fisher information to the standard Gaussian of sequences belonging to a given Wiener chaos is actually equivalent to convergence of only the fourth moment. Thus, our result may be considered as a further building block associated to the recent but already rich literature dedicated to the Fourth Moment Theorem of Nualart and Peccati [31]. To illustrate the power of our approach we prove a local limit theorem together with some rates of convergence for the normal convergence of a standardized version of the quadratic variation of the fractional Brownian motion.

Keywords: Fisher information; total variation distance; relative entropy; Fourth Moment Theorem; Fractional Brownian motion; Malliavin calculus.

1 Introduction

Measuring the discrepancy between the law of a given real-valued random variable F and that of its *Gaussian* counterpart N is arguably an important and recurrent problem both in probability and statistics. For instance, one faces this situation when trying to prove a central limit type theorem, or when wanting to check the asymptotic normality of an estimator. And quite often, the choice of a suitable probability metric reveals to be a crucial step.

In the present paper, we are concerned with this question within the framework of the Malliavin calculus. More precisely, we will focus on the Wiener chaos of a given order and, as a way to measure the proximity between laws, we will work either with the L^r -distance between densities (especially for $r = 1$ and $r = \infty$), or with the relative entropy $D(F||N)$,

*I. Nourdin is partially supported by the ANR grant ‘Malliavin, Stein and Stochastic Equations with Irregular Coefficients’ [ANR- 10-BLAN-0121].

†D. Nualart is supported by the NSF grant DMS1208625

or with the relative Fisher information $J(F) - 1$. These three notions, that we will recall now, are strongly related to each other.

Let F be a centered real-valued random variable with unit variance and density p_F . We suppose throughout that all needed assumptions on p_F (such as its strictly positivity, differentiability, etc.) are always satisfied when required. Let also $N \sim N(0, 1)$ be standard Gaussian, with density $p_N(x) = e^{-x^2/2}/\sqrt{2\pi}$, $x \in \mathbb{R}$.

The L^r -distance between densities of F and N is given by

$$\begin{aligned} \|p_F - p_N\|_r &= \left(\int_{\mathbb{R}} |p_F(x) - p_N(x)|^r dx \right)^{\frac{1}{r}}, \quad r \in [1, \infty); \\ \|p_F - p_N\|_\infty &= \sup_{x \in \mathbb{R}} |p_F(x) - p_N(x)| \quad (\text{assuming, say, that } p_F \text{ is continuous}). \end{aligned} \quad (1.1)$$

Actually, in what follows we will only consider the particular cases $r = 1$ and $r = \infty$. This is because the bounds we will produce are going to be of the same order. So, a bound for the L^r -distance will simply follow from the crude estimate:

$$\|p_F - p_N\|_r \leq \|p_F - p_N\|_1^{1/r} \|p_F - p_N\|_\infty^{1-1/r}.$$

When $r = 1$ in (1.1), it is an easy exercise (sometimes referred to as the Scheffé's theorem) to show that $\|p_F - p_N\|_1 = 2d_{TV}(F, N)$, where $d_{TV}(F, N)$ is the *total variation distance* defined as

$$d_{TV}(F, N) = \sup_{A \in \mathcal{B}(\mathbb{R})} |P(F \in A) - P(N \in A)|. \quad (1.2)$$

It is clear from its very definition (1.2) that $d_{TV}(F, N)$ represents a strong measure on how close the laws of F and N are.

The *relative entropy* $D(F\|N)$ of F with respect to N is given by

$$D(F\|N) = \int_{\mathbb{R}} p_F(x) \log(p_F(x)/p_N(x)) dx. \quad (1.3)$$

Our interest in this quantity comes from its link with the total variation distance, as provided by the celebrated Csiszár-Kullback-Pinsker inequality, according to which:

$$2(d_{TV}(F, N))^2 \leq D(F\|N). \quad (1.4)$$

(In particular, note that $D(F\|N) \geq 0$.) See, e.g., [5] for a proof of (1.4) and original references.

Inequality (1.4) shows that bounds on the relative entropy translate directly into bounds on the total variation distance. Hence, it makes perfectly sense to quantify the discrepancy between the law of F and that of the standard Gaussian N in terms of its relative entropy. Actually, one can go even further by considering the *Fisher information* $J(F)$ of F . Let us recall its definition. Let $s_F(F)$ denote the *score* associated to F . This is the F -measurable random variable uniquely determined by the following integration by parts:

$$E[\phi'(F)] = -E[s_F(F)\phi(F)] \quad \text{for all test function } \phi : \mathbb{R} \rightarrow \mathbb{R}. \quad (1.5)$$

When it makes sense, it is easy to compute that $s_F = p'_F/p_F$. Set $J(F) = E[s_F(F)^2]$ if the random variable $s_F(F)$ is square-integrable and $J(F) = +\infty$ otherwise. In the former case, it is a straightforward exercise to check that

$$J(F) - 1 = E[(s_F(F) + F)^2].$$

In particular, $J(F) \geq 1 = J(N)$ with equality if and only if F is standard Gaussian. Our interest in the relative Fisher information $J(F) - 1$ comes from its link with the relative entropy through the following de Bruijn's formula (stated in an integral and rescaled version due to Barron [4]; see also [14, Theorem C.1]). Assume, without loss of generality, that F and N are independent; then

$$D(F\|N) = \int_0^1 \frac{J(\sqrt{t}F + \sqrt{1-t}N) - 1}{2t} dt. \quad (1.6)$$

Since from, e.g., [14, Lemma 1.21] one has $J(\sqrt{t}F + \sqrt{1-t}N) \leq tJ(F) + (1-t)J(N) = 1 + t(J(F) - 1)$, we deduce that

$$D(F\|N) \leq \frac{1}{2}(J(F) - 1). \quad (1.7)$$

By comparing (1.7) with (1.4), we observe that the gap between $J(F)$ and $1 = J(N)$ is an even stronger measure of how the law of F is close to the standard Gaussian N . This claim is even more supported by the Shimizu's inequality [34], which gives a L^∞ -bound between p_F and p_N provided p_F is continuous and satisfies $x^2 p_F(x) \rightarrow 0$ as $x \rightarrow \pm\infty$:

$$\|p_F - p_N\|_\infty \leq \sqrt{J(F) - 1}. \quad (1.8)$$

(In the original statement of Shimizu [34], there is actually an extra factor $(1 + \sqrt{6/\pi})$ in the right-hand side of (1.8); but this latter was removed by Ley and Swan in [18]).

Let us now come to the description of the main results contained in the present paper. From now on, we will systematically assume that F belongs to a Wiener chaos \mathcal{H}_q of order $q \geq 2$, that is, has the form of a q th multiple Wiener-Itô integral (see Section 2 below for precise definitions). Our first result is the following, with $\|DF\|$ the norm of the Malliavin derivative of F (again, see Section 2 for details).

Theorem 1.1 *Let $q \geq 2$ be an integer and let $F \in \mathcal{H}_q$ have variance one. Assume in addition that $\varepsilon > 0$ and $\eta \geq 1$ satisfy*

$$E[\|DF\|^{-4-\varepsilon}] \leq \eta. \quad (1.9)$$

Then, there exists a constant $c > 0$, depending on q , ε and η but not on F , such that

$$J(F) - 1 \leq c(E[F^4] - 3). \quad (1.10)$$

In the next result, we take advantage of the conclusion (1.10) of Theorem 1.1 to complete the current state of the art related to the *Fourth Moment Theorem* of Nualart and Peccati [31]. See also the discussion located just after the statement of Corollary 1.2.

Corollary 1.2 *Fix an integer $q \geq 2$, and let $(F_n) \subset \mathcal{H}_q$ be a sequence of random variables satisfying $E[F_n^2] = 1$ for all n . Then, the following four assertions are equivalent as $n \rightarrow \infty$:*

- (a) $E[F_n^4] \rightarrow 3$;
- (b) $F_n \xrightarrow{\text{law}} N \sim N(0, 1)$;
- (c) $d_{TV}(F_n, N) = \frac{1}{2} \|p_{F_n} - p_N\|_1 \rightarrow 0$;
- (d) $D(F_n \| N) \rightarrow 0$.

Moreover, there exists $c_1, c_2, c_3, c_4 > 0$ (independent of n) such that, for all n large enough,

$$c_1 \max\{|E[F_n^3]|, E[F_n^4] - 3\} \leq d_{TV}(F_n, N) \leq c_2 \max\{|E[F_n^3]|, E[F_n^4] - 3\} \quad (1.11)$$

$$\leq c_3 \sqrt{E[F_n^4] - 3}; \quad (1.12)$$

$$D(F_n \| N) \leq c_4 (E[F_n^4] - 3) |\log(E[F_n^4] - 3)|. \quad (1.13)$$

Suppose in addition that, for some $\varepsilon > 0$,

$$\limsup_{n \rightarrow \infty} E[\|DF_n\|^{-4-\varepsilon}] < \infty. \quad (1.14)$$

Then, the four previous assertions (a) – (d) are equivalent to the following two further assertions:

- (e) $\|p_{F_n} - p_N\|_\infty \rightarrow 0$;
- (f) $J(F_n) \rightarrow 1$.

More precisely, one has the existence of $c_5 > 0$ (independent of n) such that, for all n large enough,

$$\|p_{F_n} - p_N\|_\infty^2 \leq c_5 (E[F_n^4] - 3) \quad (1.15)$$

$$J(F_n) - 1 \leq c_5 (E[F_n^4] - 3). \quad (1.16)$$

Equivalence between (a) and (b) in Corollary 1.2 is known as the *Fourth Moment Theorem*. This striking result, discovered by Nualart and Peccati in [31], has been the starting point of a new and fruitful line of research, consisting in using the Malliavin calculus to prove limit theorems. It has led to a burst of new research in many different fields, such as information theory [27], stochastic geometry [16, 32], Markov operator [2, 17], random matrices of large size [23], free probability [15, 25], q -calculus [1, 11], computer science

[9, 10], cosmology [19, 20], statistics [3, 13], or spin glasses [28, 35], to name a few. One can also consult the constantly updated webpage

<http://www.iecn.u-nancy.fr/~nourdin/steinmalliavin.htm>

for literally hundreds of results related to the Fourth Moment Theorem and its ramifications.

Equivalence between (a) and (c) in Corollary 1.2, coming from the bound (1.12), is due to Nourdin and Peccati [22]. By combining Malliavin calculus with the Stein's method, they were indeed able to show that, for any random variable $F \in \mathcal{H}_q$ such that $E[F^2] = 1$,

$$d_{TV}(F, N) \leq 2E\left|1 - \frac{1}{q}\|DF\|^2\right| \leq \sqrt{\frac{4q-4}{3q}} \sqrt{E[F^4] - 3}, \quad (1.17)$$

see [24, Theorem 5.2.6]. (Note that $E[F^4] > 3$ in (1.17): see, e.g., [24, Lemma 5.2.4].) The refinement (1.11) of (1.12), leading to optimal rates, is taken from [26].

Equivalence between (a) and (d) in Corollary 1.2, as well as the bound (1.13), was shown by Nourdin, Peccati and Swan in [27]. Their strategy of proof relied on the discovery a novel representation formula for the relative entropy, namely,

$$D(F\|N) = \frac{1}{2} \int_0^1 \frac{t}{1-t} E[E[N(1 - \frac{1}{q}\|DF\|^2)|\sqrt{t}F + \sqrt{1-t}N]^2] dt$$

for any $F \in \mathcal{H}_q$ with unit variance and where $N \sim N(0, 1)$ is supposed to be independent of F .

When (1.14) is satisfied for $\varepsilon = 2$, the inequality (1.15) (leading to the equivalence between (a) and (e) in Corollary 1.2) was proved by Hu, Lu and Nualart in [12], after adapting Stein's method to handle the supremum distance. Note that combining our Theorem 1.1 with Shimizu inequality (1.8) allows to recover (1.15) (which corresponds to Theorem 4.1 in [12]).

Finally, inequality (1.16) (leading to the equivalence between (a) and (f) in Corollary 1.2) is new and will be a direct consequence of Theorem 1.1. It is worth noting at this stage that validity of (1.9) is, unfortunately, far to be a small assumption. Let us discuss this point a little bit more. As it is well-known, the Bouleau-Hirsch criterion (see, e.g., [30, Theorem 2.1.3]) asserts that any (smooth and bounded enough) random variable F in the Wiener space admits a density as soon as $P(\|DF\| > 0) = 1$. This latter condition is always satisfied for $F \in \mathcal{H}_q$ with unit variance, see [33]. In fact, one can prove a far better statement, see [29, (3.19)]: there exists a constant $c_q > 0$ such that, for all $x > 0$ and all $F \in \mathcal{H}_q$ with unit variance,

$$P(\|DF\|^2 \geq x) \geq 1 - c_q x^{\frac{1}{2q-2}}. \quad (1.18)$$

As a consequence, using that

$$E[\|DF\|^{-r}] = \int_0^\infty P(\|DF\|^2 \leq u^{-\frac{2}{r}}) du \leq 1 + \int_1^\infty P(\|DF\|^2 \leq u^{-\frac{2}{r}}) du, \quad (1.19)$$

one deduces from (1.18) that, for all $F \in \mathcal{H}_q$ with unit variance,

$$E[\|DF\|^{-r}] \leq 1 + \frac{c_q}{\frac{1}{r(q-1)} - 1} \quad \text{provided } r < \frac{1}{q-1}. \quad (1.20)$$

Unfortunately, one cannot deduce (1.14) from (1.20). It means that verifying (1.14) has to be made on a case-by-case basis, and heavily depends on the particular sequence (F_n) we are dealing with. In [12], one can find an application for the least squares estimator of the parameter θ in the Ornstein-Uhlenbeck process $dX_t = \theta X_t dt + dW_t$, where W is a standard Brownian motion. In the present paper, we consider a more involved application to the quadratic variation of a fractional Brownian motion B^H of index H . We obtain optimal rates for the relative Fisher information when $H < \frac{5}{8}$, and (possibly suboptimal) rates when $H < \frac{3}{4}$. More precisely, let us introduce the so-called *fractional Gaussian noise* associated with B^H , which is the Gaussian sequence given by

$$\xi_k = B^H(k+1) - B^H(k), \quad k \in \mathbb{N} \cup \{0\}. \quad (1.21)$$

Set

$$F_n := \frac{1}{\sqrt{n} v_n} \sum_{k=0}^{n-1} (\xi_k^2 - 1), \quad (1.22)$$

with $v_n > 0$ chosen so that $E[F_n^2] = 1$. It is well-known (it is indeed a very particular case of the Breuer-Major Theorem [7], see also [21, Theorem 7.2]) that, as $n \rightarrow \infty$,

$$F_n \xrightarrow{\text{law}} N(0, 1) \iff H \in (0, 3/4]. \quad (1.23)$$

In Section 4, we will show that (1.14) is satisfied for F_n defined by (1.22). Then, as a consequence of (1.15) and (1.16) on one hand and of the estimates for $E[F_n^4] - 3$ computed in [6] on the other hand, we will be able to deduce the following *local* limit theorem for F_n .

Theorem 1.3 *Let F_n be as in (1.22). Then, there exists $c, C > 0$ independent of n such that, for all n large enough,*

$$\|p_{F_n} - p_N\|_\infty \leq \sqrt{J(F_n) - 1} \leq C \times \begin{cases} n^{-\frac{1}{2}} & \text{if } 0 < H < \frac{5}{8} \\ n^{-\frac{1}{2}} \log^{\frac{3}{2}} n & \text{if } H = \frac{5}{8} \\ n^{4H-3} & \text{if } \frac{5}{8} < H < \frac{3}{4} \end{cases}$$

and

$$\sqrt{J(F_n) - 1} \geq c n^{-\frac{1}{2}} \quad \text{if } H < \frac{5}{8}.$$

A brief outline of the paper is as follows. In Section 2, we introduce the language of the Malliavin calculus, which is the framework in which our study takes place. We also recall the Carbery-Wright inequality, which will play a key role in the proof of Theorem 1.3. Proofs of Theorem 1.1 and Corollary 1.2 are presented in Section 3. Finally, Section 4 contains the proof of Theorem 1.3.

2 Notation and preliminaries

2.1 The language of Gaussian analysis and Malliavin calculus

We start by briefly recalling some basic notation and results connected to Gaussian analysis and Malliavin calculus. The reader is referred to [24, 30] for details or missing proofs.

Let \mathfrak{H} be a real separable Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathfrak{H}}$. The norm of \mathfrak{H} will be denoted by $\| \cdot \| = \| \cdot \|_{\mathfrak{H}}$. Recall that we call *isonormal Gaussian process* over \mathfrak{H} any centered Gaussian family $X = \{X(h) : h \in \mathfrak{H}\}$, defined on a probability space (Ω, \mathcal{F}, P) and such that $E[X(h)X(g)] = \langle h, g \rangle_{\mathfrak{H}}$ for every $h, g \in \mathfrak{H}$. Assume from now on that \mathcal{F} is the σ -field generated by X .

For any integer $q \in \mathbb{N} \cup \{0\}$, we denote by \mathcal{H}_q the q th *Wiener chaos* of X . We recall that \mathcal{H}_0 is simply \mathbb{R} whereas, for any $q \geq 1$, \mathcal{H}_q is the closed linear subspace of $L^2(\Omega)$ generated by the family of random variables $\{H_q(X(h)), h \in \mathfrak{H}, \|h\|_{\mathfrak{H}} = 1\}$, with H_q the q th Hermite polynomial given by

$$H_q(x) = (-1)^q e^{\frac{x^2}{2}} \frac{d^q}{dx^q} \left(e^{-\frac{x^2}{2}} \right).$$

For any $q \geq 1$, we denote by $\mathfrak{H}^{\otimes q}$ (resp. $\mathfrak{H}^{\odot q}$) the q th tensor product (resp. the q th *symmetric* tensor product) of \mathfrak{H} . Then, the mapping $I_q(h^{\otimes q}) = H_q(X(h))$ can be extended to a linear isometry between $\mathfrak{H}^{\odot q}$ (equipped with the modified norm $\sqrt{q!} \| \cdot \|_{\mathfrak{H}^{\otimes q}}$) and \mathcal{H}_q . For $q = 0$ and $x \in \mathbb{R}$, we write $I_0(x) = x$. In the particular case where $\mathfrak{H} = L^2(A, \mathcal{A}, \mu)$, where (A, \mathcal{A}) is a measurable space and μ is a σ -finite and non-atomic measure, one has that $\mathfrak{H}^{\odot q} = L^2_s(A^q, \mathcal{A}^{\otimes q}, \mu^{\otimes q})$ 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 .

Recall that $L^2(\Omega) = \bigoplus_{q=0}^{\infty} \mathcal{H}_q$, meaning that every square-integrable random variable F measurable with respect to \mathcal{F} admits a unique decomposition of the type

$$F = E[F] + \sum_{q=1}^{\infty} I_q(f_q), \tag{2.24}$$

where the series converges in $L^2(\Omega)$, and $f_q \in \mathfrak{H}^{\odot q}$, for $q \geq 1$. Identity (2.24) is the so-called *Wiener-Itô chaotic decomposition* of F . According to a classical result of Shigekawa [33], when F is not zero and when the kernels f_q in (2.24) all equal zero except for a finite number, then the distribution of F necessarily admits a density with respect to the Lebesgue measure.

Let $\{e_i, i \geq 1\}$ be a complete orthonormal system in \mathfrak{H} . Given $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$, for every $r = 0, \dots, p \wedge q$, the *contraction* of f and g of order r is the element of $\mathfrak{H}^{\odot(p+q-2r)}$ defined by

$$f \otimes_r g = \sum_{i_1, \dots, i_r=1}^{\infty} \langle f, e_{i_1} \otimes \dots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}} \otimes \langle g, e_{i_1} \otimes \dots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}}.$$

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

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

Moreover, $f \otimes_0 g = f \otimes g$ equals the tensor product of f and g while, for $p = q$, $f \otimes_p g = \langle f, g \rangle_{\mathfrak{H}^{\otimes p}}$. The contraction $f \otimes_r g$ is not necessarily symmetric, and we denote by $f \widetilde{\otimes}_r g$ its symmetrization. We have the following product formula: if $f \in \mathfrak{H}^{\otimes p}$ and $g \in \mathfrak{H}^{\otimes q}$ then

$$I_p(f)I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \widetilde{\otimes}_r g). \quad (2.25)$$

We will now introduce some standard operators from Malliavin calculus. Let \mathcal{S} be the set of all cylindrical random variables of the form

$$F = g(X(h_1), \dots, X(h_n)),$$

where $n \geq 1$, $h_i \in \mathfrak{H}$, and g is infinitely differentiable such that all its partial derivatives have polynomial growth. The Malliavin derivative of F is the element of $L^2(\Omega; \mathfrak{H})$ defined by

$$DF = \sum_{i=1}^n \frac{\partial g}{\partial x_i}(X(h_1), \dots, X(h_n)) h_i.$$

By iteration, for every $m \geq 2$, we define the m th derivative $D^m F$ which is an element of $L^2(\Omega; \mathfrak{H}^{\otimes m})$. For $m \geq 1$ and $p \geq 1$, $\mathbb{D}^{m,p}$ denote the closure of \mathcal{S} with respect to the norm $\|\cdot\|_{m,p}$ defined by

$$\|F\|_{m,p}^p = E[|F|^p] + \sum_{j=1}^m E[\|D^j F\|_{\mathfrak{H}^{\otimes j}}^p].$$

One can then extend the definition of D^m to $\mathbb{D}^{m,p}$. When $m = 1$, one simply write D instead of D^1 . As a consequence of the hypercontractivity property of the Ornstein-Uhlenbeck semigroup (see, e.g., [24, Theorem 2.7.2]), all the $\|\cdot\|_{m,p}$ -norms are equivalent in any *finite* sum of Wiener chaoses. This is a crucial result that will be used all along the paper.

The Malliavin derivative D satisfies the following *chain rule*: if $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ is in \mathcal{C}_b^1 (that is, belongs to the set 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

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

Also, when $\mathfrak{H} = L^2(A, \mathcal{A}, \mu)$ (with μ non-atomic), one has, for any $f \in L_s^2(A^q, \mathcal{A}^{\otimes q}, \mu^{\otimes q})$,

$$D_x(I_q(f)) = qI_{q-1}(f(\cdot, x)), \quad x \in A.$$

The *divergence operator* δ , which will play a crucial role in our approach, is defined as the adjoint of D . Denoting by $\text{dom } \delta$ its domain, one has the so-called *integration by parts formula*: for every $D \in \mathbb{D}^{1,2}$ and every $u \in \text{dom } \delta$,

$$E[F\delta(u)] = E[\langle DF, u \rangle_{\mathfrak{H}}]. \quad (2.26)$$

We will moreover need the following two properties. For every $F \in \mathbb{D}^{1,2}$ and every $u \in \text{dom } \delta$ such that Fu and $F\delta(u) + \langle DF, u \rangle_{\mathfrak{H}}$ are square integrable, one has that $Fu \in \text{dom } \delta$ and

$$\delta(Fu) = F\delta(u) - \langle DF, u \rangle_{\mathfrak{H}}. \quad (2.27)$$

Also, one has a commutation relationship between the Malliavin derivative and the Skorohod integral:

$$D\delta(u) = u + \delta(Du), \quad (2.28)$$

for any $u \in \mathbb{D}^{2,2}(\mathfrak{H})$. In particular, for such an u ,

$$E[\delta(u)^2] = E[\|u\|_{\mathfrak{H}}^2] + E[\|Du\|_{\mathfrak{H}^{\otimes 2}}^2]. \quad (2.29)$$

2.2 Carbery-Wright inequality

In the proof of Theorem 1.3, we will make use of the following inequality due to Carbery and Wright [8, Theorem 8]: there is an absolute constant $c > 0$ such that, for any $d, n \geq 1$, any polynomial $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ of degree at most d and any Gaussian random vector (X_1, \dots, X_n) ,

$$E[|Q(X_1, \dots, X_n)|^{\frac{1}{d}} P(|Q(X_1, \dots, X_n)| \leq x)] \leq c d x^{\frac{1}{d}}, \quad x > 0. \quad (2.30)$$

3 Proofs of Theorem 1.1 and Corollary 1.2

In what follows, c denote positive constants which may depend of q , ε and η but not of F , and whose values may change from one appearance to the next. Also, $\langle \cdot, \cdot \rangle$ ($\|\cdot\|$, respectively) always stands for inner product (the norm, respectively) in an appropriate tensor product $\mathfrak{H}^{\otimes s}$.

3.1 Proof of Theorem 1.1

Observe first that, without loss of generality, we may and will assume that X is an isonormal process over some Hilbert space of the type $\mathfrak{H} = L^2(A, \mathcal{A}, \mu)$ (with μ non atomic).

Due to (1.9) and the fact that F has moments of all order by hypercontractivity, it is straightforward to check that $DF\|DF\|^{-2} \in \text{dom}\delta$ with $E[\delta(DF\|DF\|^{-2})^2] < \infty$. Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be a test function. We have, on one hand,

$$E[\delta(DF\|DF\|^{-2})\phi(F)] = E[\langle D\phi(F), DF\|DF\|^{-2} \rangle] = E[\phi'(F)].$$

After setting $\Sigma = 1 - \frac{1}{q}\|DF\|^2$ and because $\delta DF = qF$ for any $F \in \mathcal{H}_q$, we deduce from (1.5) that

$$s_F(F) + F = -E \left[\delta \left(DF \left(\|DF\|^{-2} - \frac{1}{q} \right) \right) \middle| F \right] = -E[\delta(DF\|DF\|^{-2}\Sigma)|F].$$

Using the formula

$$\delta(GDF) = G\delta(DF) - \langle DF, DG \rangle = qFG - \langle DF, DG \rangle,$$

one can write for $G = \|DF\|^{-2}\Sigma$,

$$-\delta(DF\|DF\|^{-2}\Sigma) = -qF\|DF\|^{-2}\Sigma + \langle DF, DG \rangle.$$

Notice that

$$DG = -2\|DF\|^{-4}(D^2F \otimes_1 DF)\Sigma - \frac{2}{q}\|DF\|^{-2}(D^2F \otimes_1 DF).$$

Therefore,

$$\langle DF, DG \rangle = -2\|DF\|^{-4}\langle D^2F, DF \otimes DF \rangle \Sigma - \frac{2}{q}\|DF\|^{-2}\langle DF, D^2F \otimes_1 DF \rangle.$$

This leads to the estimate

$$|\langle DF, DG \rangle| \leq 2\|DF\|^{-2}\|D^2F\|\|\Sigma\| + \frac{2}{q}\|DF\|^{-1}\|D^2F \otimes_1 DF\|.$$

As a consequence,

$$\delta(DF\|DF\|^{-2}\Sigma)^2 \leq 2q^2F^2\|DF\|^{-4}\Sigma^2 + 16\|DF\|^{-4}\|D^2F\|^2\Sigma^2 + \frac{16}{q^2}\|DF\|^{-2}\|D^2F \otimes_1 DF\|^2.$$

Thus, using among other properties the hypercontractivity for F , $\|D^2F\|^2$ and Σ ,

$$\begin{aligned} J(F) - 1 &= E[(s_F(F) + F)^2] \leq E[\delta(DF\|DF\|^{-2}\Sigma)^2] \\ &\leq c E[\|DF\|^{-4-\varepsilon}]^{\frac{4}{4+\varepsilon}} (E[\Sigma^2] + E[\Sigma^2]E[\|D^2F\|^4]^{\frac{1}{2}} + E[\|D^2F \otimes_1 DF\|^4]^{\frac{1}{2}}). \end{aligned} \tag{3.31}$$

Now, use the product formula to get that, for any $x \in A$,

$$\begin{aligned}
(D^2F \otimes_1 DF)(x) &= q^2(q-1) \int_A I_{q-2}(f(x, y, \cdot)) I_{q-1}(f(y, \cdot)) d\mu(y) \\
&= q^2(q-1) \sum_{r=0}^{q-2} r! \binom{q-1}{r} \binom{q-2}{r} I_{2q-3-2r} \left(\int_A f(x, y, \cdot) \otimes_r f(y, \cdot) d\mu(y) \right) \\
&= q^2(q-1) \sum_{r=1}^{q-1} (r-1)! \binom{q-1}{r-1} \binom{q-2}{r-1} I_{2q-1-2r}((f \otimes_r f)(x, \cdot)).
\end{aligned}$$

As a result, using again the product formula and with $c_{q,r,s,a}$ some constant whose exact value is useless here,

$$\begin{aligned}
&\|D^2F \otimes_1 DF\|^2 \\
&= \sum_{r,s=1}^{q-1} \sum_{a=0}^{2q-1-2(r \vee s)} c_{q,r,s,a} I_{4q-2-2r-2s-2a} \left(\int_A \left((f \otimes_r f)(x, \cdot) \otimes_a (f \otimes_s f)(x, \cdot) \right) d\mu(x) \right),
\end{aligned}$$

implying in turn

$$\begin{aligned}
&E\|D^2F \otimes_1 DF\|^4 \\
&\leq c \sum_{r,s=1}^{q-1} \sum_{a=0}^{2q-1-2(r \vee s)} \left\| \int_A \left((f \otimes_r f)(x, \cdot) \otimes_a (f \otimes_s f)(x, \cdot) \right) d\mu(x) \right\|^2 \\
&\leq c \sum_{r,s=1}^{q-1} \sum_{a=0}^{2q-1-2(r \vee s)} \left(\int_A \left\| (f \otimes_r f)(x, \cdot) \otimes_a (f \otimes_s f)(x, \cdot) \right\| d\mu(x) \right)^2 \\
&\leq c \sum_{r,s=1}^{q-1} \left(\int_A \|(f \otimes_r f)(x, \cdot)\| \|(f \otimes_s f)(x, \cdot)\| d\mu(x) \right)^2 \\
&\leq c \left(\sum_{r=1}^{q-1} \int_A \|(f \otimes_r f)(x, \cdot)\|^2 d\mu(x) \right)^2 = c \left(\sum_{r=1}^{q-1} \|f \otimes_r f\|^2 \right)^2 \leq c(E[F^4] - 3)^2,
\end{aligned}$$

the last inequality following from [24, identities (5.2.5)-(5.2.6)]. On the other hand, we can also write

$$\begin{aligned}
\|D^2F\|^2 &= q^2(q-1)^2 \int_{A^2} I_{q-2}(f(x, y, \cdot))^2 d\mu(x) d\mu(y) \\
&= q^2(q-1)^2 \sum_{r=0}^{q-2} r! \binom{q-2}{r}^2 I_{2q-4-2r} \left(\int_{A^2} f(x, y, \cdot) \otimes_r f(x, y, \cdot) d\mu(x) d\mu(y) \right) \\
&= q^2(q-1)^2 \sum_{r=2}^q (r-2)! \binom{q-2}{r-2}^2 I_{2q-2r}(f \otimes_r f),
\end{aligned}$$

so that, using moreover that $\|f\tilde{\otimes}_r f\|^2 \leq \|f\|^4 = 1/q!^2$,

$$E[\|D^2F\|^4] = q^4(q-1)^4 \sum_{r=2}^q (r-2)!^2 \binom{q-2}{r-2}^4 (2q-2r)! \|f\tilde{\otimes}_r f\|^2 \leq c.$$

Finally, recall from [24, Lemma 5.2.4] that $E[\Sigma^2] \leq \frac{q-1}{3q}(E[F^4] - 3)$. Hence, by plugging all the previous estimates in (3.31), one finally obtains the desired inequality (1.10). \square

3.2 Proof of Corollary 1.2

As we said in the Introduction, the equivalences between (a), (b), (c), (d) and (e) (provided (1.9) holds true for the latter one), as well as the estimates (1.11) and (1.13), are straightforward consequences of the main results contained in [12], [22], [27] and [31].

Note in passing that Shimizu's inequality (1.8) indeed takes place for the random variable we are considering. This is because, if $F \in \mathcal{H}_q$ satisfies $E[\|DF\|^{-4-\varepsilon}] < \infty$ then, by [30, Proposition 2.1.1], F has a continuous density given by

$$p_F(x) = E[\mathbf{1}_{\{F > x\}} \delta(DF\|DF\|^{-2})].$$

Hence, using moreover that $E[\delta(DF\|DF\|^{-2})] = 0$, one deduces that

$$p_F^2(x) \leq P(|F| > x) E[\delta(DF\|DF\|^{-2})^2] \leq |x|^{-r} E[|F|^r] E[\delta(DF\|DF\|^{-2})^2],$$

implying in turn that $p(x) = o(|x|^{-n})$ as $|x| \rightarrow \infty$.

The equivalence between (f) and (a), provided condition (1.14) is fulfilled, follows immediately from Theorem 1.1 (for one implication) and the fact that bounds on the Fisher information translates directly into bounds on the total variation distance through (1.10) and (1.4) (for the other implication).

Inequality (1.16) also follows immediately from (1.10). \square

4 Proof of Theorem 1.3

4.1 Preparation to the proof

For F_n as in the statement, recall from [6, 26] that there exist $c, C > 0$ independent of n such that, for all n large enough,

$$E[F_n^4] - 3 \leq C \times \begin{cases} n^{-1} & \text{if } 0 < H < \frac{5}{8} \\ n^{-1} \log^3 n & \text{if } H = \frac{5}{8} \\ n^{8H-6} & \text{if } \frac{5}{8} < H < \frac{3}{4} \end{cases}$$

and

$$d_{TV}(F_n, N) \geq c n^{-1} \quad \text{if } H > \frac{5}{8}.$$

Assume for an instant that (1.14) has been checked. Then, as far as the upper (resp. lower) bound is concerned, the desired conclusion directly follows from (1.15) and (1.16) (from (1.4) and (1.7), respectively).

So, in order to complete the proof of Theorem 1.3, it remains to check that (1.9) holds true.

4.2 Checking (1.9)

For simplicity, throughout all the proof, we write B instead of B^H to indicate the fractional Brownian motion of index $H \in (0, 1)$ we are dealing with. We know that B has an integral representation of the form

$$B_t = \int_0^t K_1(t, s) dW_s, \quad t \geq 0, \quad (4.32)$$

where $\{W_t, t \geq 0\}$ is a standard Brownian motion. By convention we will assume that $K_1(t, s) = 0$ if $t \leq s$. Also, we set $\Delta K_1(t, s) = K_1(t, s) - K_1(t-1, s)$.

Recall the definition (1.21) of ξ_k and the definition (1.22) of F_n . We claim that, for any $p \geq 1$, there exist n_0 such that

$$\sup_{n \geq n_0} E(\|DF_n\|^{-p}) < \infty. \quad (4.33)$$

Note that (4.33) implies (1.14).

The proof of our claim (4.33) is based on the following approach. First we will derive a lower bound for $\|DF_n\|^2$ (in distribution) denoted by B_n and defined in (4.34), which can be expressed as the sum of the square norms of n Gaussian random variables. Then, we fix an integer $N \geq 1$ and we decompose B_n into the sum of N blocks B_n^i , $i = 0, \dots, N-1$ of size $[n/N]$. The basic inequality (4.35) reduces the problem to estimate negative moments of order $\frac{p}{N}$ of each block B_n^i , provided these blocks are independent. These negative moments can be estimated by the Carbery-Wright inequality if N is large enough. Actually, the blocks are not independent, but we can control the conditional expectation of each block given the previous ones, using the properties of the fractional Brownian motion. Then, it suffices to show that these conditional expectations do not vanish as n tends to infinity (condition (4.40)), which is done in the two final steps of the proof.

The random variables ξ_k form a centered stationary Gaussian sequence with covariance

$$\rho(k) = E[\xi_r \xi_{r+k}] = \frac{1}{2} (|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}).$$

We can thus write, with D the Malliavin derivative with respect to B ,

$$\|DF_n\|^2 = \frac{4}{nv_n} \sum_{j,k=1}^n \xi_j \xi_k \rho(j-k).$$

Suppose that $\{\tilde{\xi}_j, 1 \leq j \leq n\}$ is an independent copy of the sequence $\{\xi_j, 1 \leq j \leq n\}$. Then,

$$\|DF_n\|^2 = \frac{4}{nv_n} \tilde{E} \left(\left| \sum_{j=1}^n \xi_j \tilde{\xi}_j \right|^2 \right),$$

with obvious notations. The sequence $\tilde{\xi}_j$ can be chosen of the form $\tilde{\xi}_j = \tilde{B}_j - \tilde{B}_{j-1}$, where \tilde{B} is a fractional Brownian motion of Hurst index H which is independent of B . We know that \tilde{B} has a representation of the form (different in nature from (4.32))

$$\tilde{B}_t = \kappa_H \left(\int_0^t (t-s)^{H-\frac{1}{2}} d\tilde{W}_s + Z_t \right),$$

where \tilde{W} is a standard Brownian motion (independent of W), Z is a process independent of \tilde{W} (and of W) and κ_H is a constant only depending on H . We set $K_2(t) = t^{H-\frac{1}{2}}$ if $t > 0$ and $K_2(t) = 0$ if $t \leq 0$, and $\Delta K_2(t) = K_2(t) - K_2(t-1)$. With this notation we can write

$$\tilde{\xi}_j = \kappa_H \left(\int_0^j \Delta K_2(j-s) d\tilde{W}_s + Z_j - Z_{j-1} \right).$$

As a consequence, and since $v_n = \frac{2}{n} \sum_{k,l=1}^n \rho(k-l)^2 \leq 2 \sum_{j \in \mathbb{Z}} \rho(j)^2 < \infty$ for $H \in (0, \frac{3}{4})$,

$$\begin{aligned} \|DF_n\|^2 &= \frac{c_H}{n} \tilde{E} \left(\left| \sum_{j=1}^n \xi_j \int_0^j \Delta K_2(j-s) d\tilde{W}_s \right|^2 \right) \\ &= \frac{c_H}{n} \tilde{E} \left(\left| \sum_{j=1}^n \xi_j \sum_{h=1}^j \int_{h-1}^h \Delta K_2(j-s) d\tilde{W}_s \right|^2 \right) \\ &= \frac{c_H}{n} \tilde{E} \left(\left| \sum_{h=1}^n \int_{h-1}^h \left(\sum_{j=h}^n \xi_j \Delta K_2(j-s) \right) d\tilde{W}_s \right|^2 \right) \\ &= \frac{c_H}{n} \sum_{h=1}^n \int_{h-1}^h \left(\sum_{j=h}^n \xi_j \Delta K_2(j-s) \right)^2 ds. \end{aligned}$$

Making a change of indices, we obtain

$$\|DF_n\|^2 \geq A_n := \frac{c_H}{n} \sum_{h=1}^n \int_{n-h}^{n-h+1} \left(\sum_{j=1}^h \xi_{n-j+1} \Delta K_2(n-j+1-s) \right)^2 ds.$$

The sequences $\{\xi_j, 1 \leq j \leq n\}$ and $\{\xi_{n-j+1}, 1 \leq j \leq n\}$ have the same law, so A_n has the same law as

$$B_n := \frac{c_H}{n} \sum_{h=1}^n \int_{n-h}^{n-h+1} \left(\sum_{j=1}^h \xi_j \Delta K_2(n-j+1-s) \right)^2 ds. \quad (4.34)$$

With the change of variable $s \mapsto n + 1 - s$, we get

$$B_n = \frac{c_H}{n} \sum_{h=1}^n \int_h^{h+1} \left(\sum_{j=1}^h \xi_j \Delta K_2(s-j) \right)^2 ds.$$

Fix an integer $N \geq 1$ and let $M = \lfloor n/N \rfloor$ be the integer part of n/N . Then, $n \geq NM$. As a consequence,

$$B_n \geq \frac{c_H}{n} \sum_{i=0}^{N-1} \sum_{h=iM+1}^{(i+1)M} \int_h^{h+1} \left(\sum_{j=1}^h \xi_j \Delta K_2(s-j) \right)^2 ds.$$

Set

$$B_n^i = \frac{c_H}{n} \sum_{h=iM+1}^{(i+1)M} \int_h^{h+1} \left(\sum_{j=1}^h \xi_j \Delta K_2(s-j) \right)^2 ds, \quad i = 0, \dots, N-1.$$

We are going to use the estimate

$$(B_n)^{-p} \leq \prod_{i=0}^{N-1} (B_n^i)^{-\frac{p}{N}}. \quad (4.35)$$

Consider again the representation of the sequence ξ as stochastic integrals with respect to a Brownian motion W , and denote by $\{\mathcal{F}_t^W\}$ the filtration generated by the Brownian motion W . Then,

$$E[(B_n)^{-p}] \leq E \left[\prod_{i=0}^{N-1} (B_n^i)^{-\frac{p}{N}} \right] = E \left[E[(B_n^{N-1})^{-\frac{p}{N}} | \mathcal{F}_{(N-1)M}^W] \prod_{i=0}^{N-2} (B_n^i)^{-\frac{p}{N}} \right]. \quad (4.36)$$

Let us estimate the conditional expectation appearing in the right-hand side. In the same spirit that (1.19), it is immediate that

$$E[(B_n^{N-1})^{-\frac{p}{N}} | \mathcal{F}_{(N-1)M}^W] \leq 1 + \frac{p}{N} \int_0^1 P(B_n^{N-1} \leq x | \mathcal{F}_{(N-1)M}^W) x^{-\frac{p}{N}-1} dx.$$

By Carbery-Wright's inequality (2.30) with $d = 2$,

$$P(B_n^{N-1} \leq x | \mathcal{F}_{(N-1)M}^W) \leq c\sqrt{x} [E(B_n^{N-1} | \mathcal{F}_{(N-1)M}^W)]^{-\frac{1}{d}}. \quad (4.37)$$

The conditional expectation $E(B_n^{N-1} | \mathcal{F}_{(N-1)M}^W)$ is given by

$$E(B_n^{N-1} | \mathcal{F}_{(N-1)M}^W) = \frac{c_H}{n} \sum_{h=(N-1)M+1}^{NM} \int_h^{h+1} E \left[\left(\sum_{j=1}^h \xi_j \Delta K_2(s-j) \right)^2 | \mathcal{F}_{(N-1)M}^W \right] ds. \quad (4.38)$$

Taking into account that the sequence ξ_j is Gaussian, the conditional expectation appearing in the above equation can be bounded below by the conditional variance which is *not* random. More precisely,

$$\begin{aligned} E \left[\left(\sum_{j=1}^h \xi_j \Delta K_2(s-j) \right)^2 \middle| \mathcal{F}_{(N-1)M}^W \right] &\geq \text{Var} \left[\sum_{j=1}^h \xi_j \Delta K_2(s-j) \middle| \mathcal{F}_{(N-1)M}^W \right] \\ &= E \left[\left(\sum_{j=(N-1)M+1}^h \left(\int_{(N-1)M}^j \Delta K_1(j,u) dW_u \right) \Delta K_2(s-j) \right)^2 \right]. \end{aligned} \quad (4.39)$$

By plugging (4.37), (4.38) and (4.39) into (4.36) and then by proceeding by induction with the other terms, we see that (4.33) will follow as soon as, for any $i = 0, \dots, N-1$,

$$\liminf_{M \rightarrow \infty} \frac{1}{M} \sum_{h=iM+1}^{(i+1)M} \int_h^{h+1} E \left[\left(\sum_{j=iM+1}^h \left(\int_{iM}^j \Delta K_1(j,u) dW_u \right) \Delta K_2(s-j) \right)^2 \right] ds > 0. \quad (4.40)$$

Proof of (4.40). First we compute the expectation in (4.40):

$$\begin{aligned} &E \left[\left(\sum_{j=iM+1}^h \left(\int_{iM}^j \Delta K_1(j,u) dW_u \right) \Delta K_2(s-j) \right)^2 \right] \\ &= \sum_{j,k=iM+1}^h \Delta K_2(s-j) \Delta K_2(s-k) \int_{iM}^{j \wedge k} \Delta K_1(j,u) \Delta K_1(k,u) du. \end{aligned}$$

Set

$$\beta_{j,k}^M = \int_{iM}^{j \wedge k} \Delta K_1(j,u) \Delta K_1(k,u) du.$$

We can write, exchanging the order of the summation

$$\begin{aligned} &\sum_{h=iM+1}^{(i+1)M} \sum_{j,k=iM+1}^h \left(\int_h^{h+1} \Delta K_2(s-j) \Delta K_2(s-k) ds \right) \beta_{j,k}^M \\ &= \sum_{j,k=iM+1}^{(i+1)M} \sum_{h=j \vee k}^{(i+1)M} \left(\int_h^{h+1} \Delta K_2(s-j) \Delta K_2(s-k) ds \right) \beta_{j,k}^M = \sum_{j,k=iM+1}^{(i+1)M} \alpha_{j,k}^M \beta_{j,k}^M, \end{aligned}$$

where

$$\alpha_{j,k}^M = \int_{j \vee k}^{(i+1)M+1} \Delta K_2(s-j) \Delta K_2(s-k) ds.$$

Then we are interested in the liminf, as M tends to infinity, of

$$\frac{1}{M} \sum_{j,k=iM+1}^{(i+1)M} \alpha_{j,k}^M \beta_{j,k}^M.$$

We make now the change of indices $j \mapsto j - iM$ and $k \mapsto k - iM$ and we obtain the expression

$$\Psi_M := \frac{1}{M} \sum_{j,k=1}^M \tilde{\alpha}_{j,k}^M \tilde{\beta}_{j,k}^M,$$

where

$$\begin{aligned} \tilde{\beta}_{j,k}^M &= \int_{iM}^{(j+iM) \wedge (k+iM)} \Delta K_1(j+iM, u) \Delta K_1(k+iM, u) du \\ &= \int_0^{j \wedge k} \Delta K_1(j+iM, u+iM) \Delta K_1(k+iM, u+iM) du, \end{aligned}$$

and

$$\tilde{\alpha}_{j,k}^M = \int_{j \vee k}^{M+1} \Delta K_2(s-j) \Delta K_2(s-k) ds.$$

Step 1: Case $H > \frac{1}{2}$. In this case ΔK_1 and ΔK_2 are nonnegative. On the other hand, by [30, (5.10)], we have

$$\frac{\partial K_1}{\partial t}(t, s) = c_H (t/s)^{H-\frac{1}{2}} (t-s)^{H-\frac{3}{2}},$$

where $c_H = \sqrt{\frac{H(2H-1)}{\beta(2-2H, H-1/2)}}$. Therefore, assuming $j \geq k \geq 2$,

$$\begin{aligned} \tilde{\beta}_{j,k}^M &\geq \int_1^{k-1} \Delta K_1(j+iM, u+iM) \Delta K_1(k+iM, u+iM) du \\ &\geq c_H^2 \int_1^{k-1} \left(\int_{j+iM-1}^{j+iM} \left(\frac{x}{u+iM} \right)^{H-\frac{1}{2}} (x-u-iM)^{H-\frac{3}{2}} dx \right) \\ &\quad \times \left(\int_{k+iM-1}^{k+iM} \left(\frac{y}{u+iM} \right)^{H-\frac{1}{2}} (y-u-iM)^{H-\frac{3}{2}} dy \right) du \\ &\geq c_H^2 \int_1^{k-1} \left(\frac{(j+iM-1)(k+iM-1)}{(u+iM)^2} \right)^{H-\frac{1}{2}} (j-u)^{H-\frac{3}{2}} (k-u)^{H-\frac{3}{2}} du. \end{aligned}$$

The term $\frac{(j+iM-1)(k+iM-1)}{(u+iM)^2}$ is lower bounded by 1. Therefore

$$\begin{aligned} \tilde{\beta}_{j,k}^M &\geq c_H^2 \int_1^{k-1} (j-u)^{H-\frac{3}{2}} (k-u)^{H-\frac{3}{2}} du = c_H^2 \int_1^{k-1} (x+j-k)^{H-\frac{3}{2}} x^{H-\frac{3}{2}} dx \\ &\geq c_H^2 \int_1^{k-1} (x+j-k)^{2H-3} dx = C ((j-k+1)^{2H-2} - (j-1)^{2H-2}). \end{aligned}$$

By similar arguments we obtain, assuming again $j \geq k \geq 2$,

$$\begin{aligned}\tilde{\alpha}_{j,k}^M &\geq \int_{j+2}^{M+1} \Delta K_2(s-j) \Delta K_2(s-k) ds \\ &\geq (H - \frac{1}{2}) \int_{j+2}^{M+1} (s-j)^{H-\frac{3}{2}} (s-k)^{H-\frac{3}{2}} ds = (H - \frac{1}{2}) \int_2^{M+1-j} x^{H-\frac{3}{2}} (x+j-k)^{H-\frac{3}{2}} dx \\ &\geq (H - \frac{1}{2}) \int_2^{M+1-j} (x+j-k)^{2H-3} dx = C ((j-k+2)^{2H-2} - (M+1-k)^{2H-2}).\end{aligned}$$

Therefore,

$$\begin{aligned}\Psi_M &\geq \frac{1}{M} \sum_{2 \leq k \leq j \leq M} \tilde{\alpha}_{j,k}^M \tilde{\beta}_{j,k}^M \\ &\geq \frac{C}{M} \sum_{k=2}^M \sum_{\ell=0}^{M-k} ((\ell+1)^{2H-2} - (\ell+k-1)^{2H-2}) ((\ell+2)^{2H-2} - (M+1-k)^{2H-2}).\end{aligned}$$

For the first term we obtain

$$\frac{1}{M} \sum_{k=2}^M \sum_{\ell=0}^{M-k} (\ell+1)^{2H-2} (\ell+2)^{2H-2} \geq \frac{1}{M} \sum_{k=2}^M \sum_{\ell=0}^{M-k} (\ell+2)^{4H-4} \geq \frac{M-1}{M} 2^{4H-4},$$

which converges to a positive constant as M tends to infinity. It is easy to check that the other terms in the above expression converge to zero. For instance,

$$\begin{aligned}\frac{1}{M} \sum_{k=2}^M \sum_{\ell=0}^{M-k} (\ell+2)^{2H-2} (\ell+k-1)^{2H-2} &= \frac{1}{M} \sum_{\ell=0}^{M-2} (\ell+2)^{2H-2} \sum_{k=2}^{M-\ell} (\ell+k-1)^{2H-2} \\ &\leq \frac{1}{M} \left(\sum_{\ell=1}^M \ell^{2H-2} \right)^2\end{aligned}$$

and this last quantity behaves as M^{4H-3} , which converges to zero because $H < \frac{3}{4}$. A similar analysis can be done for the other terms.

Step 2: Case $H < \frac{1}{2}$. In this case, see [30, (5.23)], we have that

$$\frac{\partial K_1}{\partial t}(t, s) = c_H (H - \frac{1}{2}) (t/s)^{H-\frac{1}{2}} (t-s)^{H-\frac{3}{2}} \quad (\text{with } c_H = \sqrt{\frac{2H}{(1-2H)\beta(1-2H, H+1/2)}})$$

is negative. Therefore $\Delta K_1(j, s)$ is negative if $s < j-1$ and positive if $j-1 \leq s \leq j$. Also,

$$\frac{\partial K_2}{\partial t}(u) = (H - \frac{1}{2}) u^{H-\frac{3}{2}}$$

and $\Delta K_2(u)$ is negative if $u \geq 1$ and positive if $u < 1$. Then, it suffices to show that the negative terms do not contribute to the limit, and once we get rid of these negative

terms, we can get a lower bound as in the case $H > \frac{1}{2}$. When $j = k$, the integrands in the definition of $\tilde{\beta}_{j,k}^M$ and $\tilde{\alpha}_{j,k}^M$ are nonnegative. On the other hand, for $j \geq k + 1$, we can write

$$\tilde{\alpha}_{j,k}^M = \tilde{\alpha}_{j,k,1}^M + \tilde{\alpha}_{j,k,2}^M,$$

where

$$\tilde{\alpha}_{j,k,1}^M = \int_j^{j+1} (s-j)^{H-\frac{1}{2}} \left[(s-k)^{H-\frac{1}{2}} - (s-k-1)^{H-\frac{1}{2}} \right] ds \leq 0,$$

and

$$\tilde{\alpha}_{j,k,2}^M = \int_{j+1}^{M+1} \Delta K_2(s-j) \Delta K_2(s-k) ds \geq 0.$$

Similarly

$$\tilde{\beta}_{j,k}^M = \tilde{\beta}_{j,k,1}^M + \tilde{\beta}_{j,k,2}^M,$$

where

$$\tilde{\beta}_{j,k,1}^M = \int_{k-1}^k K_1(k+iM, u+iM) \Delta K_1(j+iM, u+iM) du \leq 0,$$

and

$$\tilde{\beta}_{j,k,2}^M = \int_0^{k-1} \Delta K_1(k+iM, u+iM) \Delta K_1(j+iM, u+iM) du \geq 0.$$

In this way we obtain the decomposition

$$\begin{aligned} \Psi_M &= \frac{1}{M} \sum_{j=1}^M \tilde{\alpha}_{j,j}^M \tilde{\beta}_{j,j}^M + \frac{2}{M} \sum_{1 \leq k \leq j-1 \leq M-1} (\tilde{\alpha}_{j,k,1}^M + \tilde{\alpha}_{j,k,2}^M) (\tilde{\beta}_{j,k,1}^M + \tilde{\beta}_{j,k,2}^M) \\ &= \Psi_M^1 + \Psi_M^2 + \Psi_M^3, \end{aligned}$$

where

$$\Psi_M^1 = \frac{1}{M} \sum_{j=1}^M \tilde{\alpha}_{j,j}^M \tilde{\beta}_{j,j}^M + \frac{2}{M} \sum_{1 \leq k \leq j-1 \leq M-1} (\tilde{\alpha}_{j,k,1}^M \tilde{\beta}_{j,k,1}^M + \tilde{\alpha}_{j,k,2}^M \tilde{\beta}_{j,k,2}^M),$$

$$\Psi_M^2 = \frac{2}{M} \sum_{1 \leq k \leq j-1 \leq M-1} \tilde{\alpha}_{j,k,1}^M \tilde{\beta}_{j,k,2}^M,$$

and

$$\Psi_M^3 = \frac{2}{M} \sum_{1 \leq k \leq j-1 \leq M-1} \tilde{\alpha}_{j,k,2}^M \tilde{\beta}_{j,k,1}^M.$$

The term Ψ_M^1 is nonnegative and it can be bounded below as follows

$$\begin{aligned} \Psi_M^1 &\geq \frac{2}{M} \sum_{2 \leq k \leq j-1 \leq M-1} \left(\int_{j+2}^{M+1} \Delta K_2(s-j) \Delta K_2(s-k) ds \right) \\ &\quad \times \left(\int_1^{k-1} \Delta K_1(j+iM, u+iM) \Delta K_1(k+iM, u+iM) du \right). \end{aligned}$$

By the same arguments as in the case $H > \frac{1}{2}$ we can show that $\liminf_{M \rightarrow \infty} \Psi_M^1 > 0$. Therefore, it suffices to show that

$$\lim_{M \rightarrow \infty} \Psi_M^2 = 0, \quad (4.41)$$

and

$$\lim_{M \rightarrow \infty} \Psi_M^3 = 0. \quad (4.42)$$

These limits are based on the following estimates. On one hand, $\tilde{\beta}_{j,k,2}^M$ and $\tilde{\alpha}_{j,k,2}^M$ are uniformly bounded:

$$\begin{aligned} \tilde{\beta}_{j,k,2}^M &= \int_{iM}^{iM+k-1} \Delta K_1(k+iM, u) \Delta K_1(j+iM, u) du \\ &\leq \sqrt{\int_{iM}^{iM+k-1} \Delta K_1(k+iM, u)^2 du} \sqrt{\int_{iM}^{iM+k-1} \Delta K_1(j+iM, u)^2 du} \\ &\leq \sqrt{\int_0^{k+iM} \Delta K_1(k+iM, u)^2 du} \sqrt{\int_0^{j+iM} \Delta K_1(j+iM, u)^2 du} \\ &= \sqrt{E(|B_{k+iM} - B_{k+iM-1}|^2)} \sqrt{E(|B_{j+iM} - B_{j+iM-1}|^2)} = 1, \end{aligned}$$

and

$$\begin{aligned} \tilde{\alpha}_{j,k,2}^M &= \int_{j+1}^{M+1} [(s-j)^{H-\frac{1}{2}} - (s-j-1)^{H-\frac{1}{2}}][(s-k)^{H-\frac{1}{2}} - (s-k-1)^{H-\frac{1}{2}}] ds \\ &= \int_1^{M+1-j} [x^{H-\frac{1}{2}} - (x-1)^{H-\frac{1}{2}}][(x+j-k)^{H-\frac{1}{2}} - (x+j-k-1)^{H-\frac{1}{2}}] dx \\ &\leq C + \int_2^{M+1-j} [x^{H-\frac{1}{2}} - (x-1)^{H-\frac{1}{2}}][(x+j-k)^{H-\frac{1}{2}} - (x+j-k-1)^{H-\frac{1}{2}}] dx \\ &\leq C + C_H \int_2^{M+1-j} (x-1)^{H-\frac{3}{2}} (x+j-k-1)^{H-\frac{3}{2}} dx \\ &\leq C + C_H \int_2^{M+1-j} (x-1)^{2H-3} dx \leq C'. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} |\tilde{\alpha}_{j,k,1}^M| &\leq \int_0^1 x^{H-\frac{1}{2}} \left[(j-k+x-1)^{H-\frac{1}{2}} - (j-k+x)^{H-\frac{1}{2}} \right] dx \\ &= \frac{1}{\frac{3}{2}-H} \int_0^1 x^{H-\frac{1}{2}} \left(\int_{j-k-1}^{j-k} (u+x)^{H-\frac{3}{2}} du \right) dx \leq \frac{1}{(\frac{3}{2}-H)(\frac{1}{2}+H)} (j-k-1)^{H-\frac{3}{2}}, \end{aligned}$$

and

$$\begin{aligned} |\tilde{\beta}_{j,k,1}^M| &\leq \int_{k-1}^k K_1(k+iM, u+iM) |\Delta K_1(j+iM, u+iM)| du \\ &= c_H \int_{k-1}^k K_1(k+iM, u+iM) \left(\int_{j+iM-1}^{j+iM} \left(\frac{x}{u+iM} \right)^{H-\frac{1}{2}} (x-u-iM)^{H-\frac{3}{2}} dx \right) du \\ &\leq c_H \left(\frac{j+iM-1}{k+iM} \right)^{H-\frac{1}{2}} \int_{k-1}^k K_1(k+iM, u+iM) \left(\int_{j-1}^j (x-u)^{H-\frac{3}{2}} dx \right) du \\ &\leq C(j-k-1)^{H-\frac{3}{2}}, \end{aligned}$$

because $\frac{j+iM-1}{k+iM} \geq 1$, the integral $\int_{j-1}^j (x-u)^{H-\frac{3}{2}} dx$ is bounded by $(j-k-1)^{H-\frac{3}{2}}$ uniformly in $u \in [k-1, k]$, and

$$\sup_{k,M} \int_{k-1}^k K_1(k+iM, u+iM) du < \infty,$$

as it can be easily checked from the expression of K_1 . Finally,

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{1 \leq k \leq j-2 \leq M-2} (j-k-1)^{H-\frac{3}{2}} = 0,$$

which implies both (4.41) and (4.42).

References

- [1] O. Arizmendi: Convergence of the fourth moment and infinite divisibility. Preprint, 2013
- [2] E. Azmoodeh, S. Campese and G. Poly: Fourth Moment Theorems for Markov Diffusion Generators. *J. Funct. Anal.*, to appear, 2013
- [3] J.-M. Bardet and D. Surgailis: Moment bounds and central limit theorems for Gaussian subordinated arrays. *J. Multi. Anal.*, to appear
- [4] A. R. Barron: Entropy and the central limit theorem. *Ann. Probab.*, 14(1):336–342, 1986.

- [5] F. Bolley and C. Villani: Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities. *Ann. Fac. Sci. Toulouse Math. (6)*, 14(3):331–352, 2005.
- [6] H. Biermé, A. Bonami, I. Nourdin and G. Peccati: Optimal Berry-Esseen rates on the Wiener space: the barrier of third and fourth cumulants. *ALEA*, 9(2): 473–500, 2012.
- [7] P. Breuer and P. Major: Central limit theorems for non-linear functionals of Gaussian fields. *J. Mult. Anal.*, 13: 425–441, 1983.
- [8] A. Carbery and J. Wright: Distributional and L^q norm inequalities for polynomials over convex bodies in \mathbb{R}^n . *Math. Research Lett.*, 8:233–248, 2001.
- [9] A. De, I. Diakonikolas and R. Servedio: Deterministic Approximate Counting for Juntas of Degree-2 Polynomial Threshold Functions. Preprint, 2013
- [10] A. De and R. Servedio: Efficient deterministic approximate counting for low-degree polynomial threshold function. Preprint, 2013
- [11] A. Deya, S. Noredine and I. Nourdin: Fourth Moment Theorem and q -Brownian Chaos. *Comm. Math. Phys.*, 321(1), 113–134, 2013
- [12] Y. Hu, F. Lu and D. Nualart: Convergence of densities of some functionals of Gaussian processes. *J. Funct. Anal.*, to appear.
- [13] Y. Hu and D. Nualart: Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Stat. Probab. Lett.*, 80(11-12), 1030–1038, 2010
- [14] O. Johnson. *Information theory and the central limit theorem*. Imperial College Press, London, 2004.
- [15] T. Kemp, I. Nourdin, G. Peccati and R. Speicher: Wigner chaos and the fourth moment. *Ann. Probab.*, 40(4), 1577–1635, 2012
- [16] R. Lachièze-Rey and G. Peccati: Fine Gaussian fluctuations on the Poisson space I: contractions, cumulants and random geometric graphs. *Electron. J. Probab.*, 18(32), 1–35, 2013
- [17] M. Ledoux. Chaos of a Markov operator and the fourth moment condition. *Ann. Probab.*, 40(6):2439–2459, 2012.
- [18] C. Ley and Y. Swan. Stein’s density approach and information inequalities. *Electron. Comm. Probab.*, 18(7):1–14, 2013.
- [19] D. Marinucci and G. Peccati: Ergodicity and Gaussianity for Spherical Random Fields. *J. Math. Phys.* 51, 043301, 2010

- [20] D. Marinucci and I. Wigman: On Nonlinear Functionals of Random Spherical Eigenfunctions. *Comm. Math. Phys.*, in press, 2013
- [21] I. Nourdin (2012). *Selected Aspects of Fractional Brownian Motion*. Springer-Verlag.
- [22] I. Nourdin and G. Peccati: Stein’s method on Wiener chaos. *Probab. Theory Related Fields*, 145:75–118, 2009.
- [23] I. Nourdin and G. Peccati: Universal Gaussian fluctuations of non-Hermitian matrix ensembles: from weak convergence to almost sure CLTs. *ALEA*, 7, 341–375, 2010
- [24] I. Nourdin and G. Peccati: *Normal Approximations using Malliavin Calculus: from Stein’s Method to the Universality*. Cambridge University Press, 2012.
- [25] I. Nourdin and G. Peccati: Poisson approximations on the free Wigner chaos. *Ann. Probab.*, 41(4), 2709–2723, 2013
- [26] I. Nourdin and G. Peccati: The optimal fourth moment theorem. *Proceedings of the AMS*, to appear, 2013
- [27] I. Nourdin, G. Peccati and Y. Swan: Entropy and the fourth moment phenomenon. *J. Funct. Anal.*, to appear.
- [28] I. Nourdin, G. Peccati and F.G. Viens: Comparison inequalities on Wiener space. *Stoch. Proc. Appl.*, to appear.
- [29] I. Nourdin and G. Poly. Convergence in total variation on Wiener chaos. *Stoch. Proc. Appl.*, to appear, 2012.
- [30] D. Nualart: *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.
- [31] D. Nualart and G. Peccati: Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.*, 33(1):177–193, 2005.
- [32] M. Reitzner and M. Schulte: Central Limit Theorems for U -statistics of Poisson Point Processes. *Ann. Probab.*, to appear
- [33] I. Shigekawa: Derivatives of Wiener functionals and absolute continuity of induced measures. *J. Math. Kyoto Univ.*, 20(2), 263–289, 1980.
- [34] R. Shimizu: On Fisher’s amount of information for location family. In *A Modern Course on Statistical Distributions in Scientific Work*, pages 305–312. Springer, 1975.
- [35] F.G. Viens: Stein’s lemma, Malliavin calculus, and tail bounds, with application to polymer fluctuation exponent. *Stoch. Proc. Appl.*, 119, 3671–3698, 2009