Density convergence in the Breuer-Major theorem for Gaussian stationary sequences

Yaozhong Hu, David Nualart, Samy Tindel, Fangjun Xu

DENSITY CONVERGENCE IN THE BREUER-MAJOR THEOREM FOR GAUSSIAN STATIONARY SEQUENCES

YAOZHONG HU, DAVID NUALART, SAMY TINDEL, AND FANGJUN XU

ABSTRACT. Consider a Gaussian stationary sequence with unit variance $X = \{X_k; k \in \mathbb{N} \cup \{0\}\}$. Assume that the central limit theorem holds for a weighted sum of the form $V_n = n^{-1/2} \sum_{k=0}^{n-1} f(X_k)$, where f designates a finite sum of Hermite polynomials. Then we prove that the uniform convergence of the density of V_n towards the standard Gaussian density also holds true, under a mild additional assumption involving the causal representation of X.

1. Introduction

Let $X = \{X_k;\ k \in \mathbb{N} \cup \{0\}\}$ be a centered Gaussian stationary sequence with unit variance. For all $v \in \mathbb{Z}$, we set $\rho(v) = \mathbf{E}[X_0 X_{|v|}]$. Therefore $\rho(0) = 1$ and $|\rho(v)| \leq 1$ for all v. Let γ be the standard Gaussian probability measure and $f \in L^2(\gamma)$ be a fixed deterministic function such that $\mathbf{E}[f(X_1)] = 0$. We expand f in the orthonormal basis of Hermite polynomials $\{H_k;\ k \geq 0\}$, which are more specifically defined in Section 2.2. In particular, if f has Hermite rank $d \geq 1$, it admits the following Hermite expansion:

$$f(x) = \sum_{j=d}^{\infty} a_j H_j(x),$$

with $a_d \neq 0$. Define $V_n = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f(X_k)$. Then the celebrated Breuer-Major Theorem (see [3] or Theorem 7.2.4 in [13]) can be written as follows:

Theorem 1.1. *Suppose that $\sum_{v \in \mathbb{Z}} |\rho(v)|^d < \infty$ and set $\sigma^2 = \sum_{j=d}^{\infty} j! a_j^2 \sum_{v \in \mathbb{Z}} \rho(v)^j$, which is assumed to be in $(0, \infty)$. Then the convergence:*

$$V_n \xrightarrow{Law} \mathcal{N}(0, \sigma^2) \tag{1}$$

holds true as n tends to infinity.

We shall be in fact interested in a particular case of Theorem 1.1 for finite linear combinations of Hermite polynomials, which is stated here for convenience:

Corollary 1.2. *Consider $2 \leq d \leq q < \infty$ and a family of real numbers $\{a_j;\ j = d, \ldots, q\}$. Let H_j be the jth order Hermite polynomial, and assume that $\sigma^2 \in (0, \infty)$, where $\sigma^2 \equiv$*

Y. Hu is partially supported by a grant from the Simons Foundation #209206.
D. Nualart is supported by the NSF grant DMS1208625.
F. Xu is supported by the 111 Project(B14019).
This project has been carried out while S. Tindel was on sabbatical at the University of Kansas. He wishes to express his gratitude to this institution for its warm hospitality.

$\sum_{j=d}^q j! a_j^2 \sum_{v\in\mathbb{Z}} \rho(v)^j$. *Set*

$$V_n^{d,q} = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \sum_{j=d}^{q} a_j \, H_j(X_k). \tag{2}$$

Then $V_n^{d,q} \xrightarrow{Law} \mathcal{N}(0,\sigma^2)$ *as* n *tends to infinity. In particular, we have:*

$$\lim_{n\to 0} \mathbf{E}\left[\left(V_n^{d,q}\right)^4\right] = 3\sigma^4. \tag{3}$$

Remark 1.3. The relation between Gaussian convergence in law for sequences in a fixed Wiener chaos and behavior of the fourth moment has been extensively studied since the seminal paper [18]. We will need only a small part of the information available on the topic, such as relation (3).

Due to its importance, Breuer-Major theorem has been extended and refined in several directions. Important generalizations can be found in Arcones [1] (multidimensional case), Chambers and Slud [6] and Giraitis and Surgailis [7]. A proof of Theorem 1.1 using a combination of Stein's method with Malliavin calculus was given by Nourdin, Peccati and Podolskij in [14], where one can find explicit bounds in the total variation and Wasserstein distances. We refer the reader to the monograph by Nourdin and Peccati [12] for a more detailed account on this topic.

We shall mainly be concerned here by convergences of densities, and here again the relationship between fourth moment behavior and various type of convergences of random variables in a fixed Wiener chaos have been thoroughly studied in the recent past. The interested reader is referred to [13] for further details, but we will use here the following recent criterion (see [11, Corollary 1.2] and [10, Corollary 4.6]).

Theorem 1.4. *Let* $\{F_n;\, n \in \mathbb{N}\}$ *be a sequence of random variables belonging to a fixed chaos* $\mathcal{H}_q$ *with* $q \geq 2$. *Suppose* $\mathbf{E}[F_n^2] = 1$ *and* $\lim_{n\to\infty} \mathbf{E}[F_n^4] = 3$. *Let* p_{F_n} *be the density of the random variable* F_n *and let* $\phi(x) = (2\pi)^{-1/2}\exp(-|x|^2/2)$ *be the density of the standard Gaussian distribution on* $\mathbb{R}$.

(i) *Suppose that for some* $\epsilon > 0$,

$$\sup_n \mathbf{E}\left[\|DF_n\|^{-4-\varepsilon}\right] < \infty.$$

Then, there exists a constant c *such that for all* $n \geq 1$,

$$\sup_{x\in\mathbb{R}} |p_{F_n}(x) - \phi(x)| \leq c\sqrt{\mathbf{E}[F_n^4] - 3}.$$

(ii) *Suppose that for all* $p \geq 1$,

$$\sup_n \mathbf{E}\left[\|DF_n\|^{-p}\right] < \infty.$$

Then, for any $m \geq 0$, *there exists a constant* c_m *such that for all* $n \geq 1$,

$$\sup_{x\in\mathbb{R}} |p_{F_n}^{(m)}(x) - \phi^{(m)}(x)| \leq c_m\sqrt{\mathbf{E}[F_n^4] - 3}.$$

The goal of the current paper is to apply the criterion given by Theorem 1.4 in order to get convergence of density in the landmark of Breuer-Major theorem. In order to do this we need a uniform estimate on the negative moments of the Malliavin derivative of the sequence, and this is the contents of our main result.

Theorem 1.5. *Let X be a Gaussian stationary sequence whose spectral density f_ρ satisfies $f_\rho \in L^{1/2}([-\pi,\pi])$ and $\log(f_\rho) \in L^1([-\pi,\pi])$ (see the hypothesis 2.1 and examples in the next section). Let $V_n^{d,q}$ be the random variable defined by (2), and assume the hypothesis of Corollary 1.2 to be satisfied. Then for any $p \geq 1$, there exists n_0 such that*

$$\sup_{n\geq n_0} \mathbf{E}\left[\|DV_n^{d,q}\|^{-p}\right] < \infty. \tag{4}$$

In the case of a fixed Wiener chaos we can obtain the following consequence.

Corollary 1.6. *Under the conditions of Theorem 1.5, if $q = d$, and we define $F_n = V_n^{d,d}/\sigma_n$, where $\sigma_n^2 = \mathbf{E}[(V_n^{d,d})^2]$, then, for all $m \geq 0$ there exists an n_0 (depending on m) such that*

$$\sup_{n\geq n_0} \sup_{x\in\mathbb{R}} |p_{F_n}^{(m)}(x) - \phi^{(m)}(x)| \leq c_m \sqrt{\mathbf{E}[F_n^4] - 3}.$$

In the case $q \neq d$, Theorem 1.4 cannot be applied. In the reference [10] one can find results on the uniform convergence of density for general random variables similar to those stated in Theorem 1.4, but they require a uniform lower bound for the negative moments of the random variable $|\langle DF_n, -DL^{-1}F_n\rangle_{\mathfrak{H}}|$, and our approach does not seem to work in this case because it is not clear how to express $\langle DF_n, -DL^{-1}F_n\rangle_{\mathfrak{H}}$ as a sum of squares. Nevertheless, condition (4) allows us to derive the uniform convergence of the densities and their derivatives from a general result proved below (see Proposition 2.6) although in this case we have no information about the rate of convergence.

Corollary 1.7. *Under the conditions of Theorem 1.5, if we define $F_n = V_n^{d,q}/\sigma_n$, where $\sigma_n^2 = \mathbf{E}[(V_n^{d,q})^2]$, then, for all $m \geq 0$ we have*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |p_{F_n}^{(m)}(x) - \phi^{(m)}x)| = 0.$$

Notice that a particular case of Theorem 1.5 has been established in [11], for $q = 2$ and $X_k = B_{k+1} - B_k$ for a fractional Brownian motion B with Hurst parameter $H \in (0,1)$. The proof of the existence of negative moments for $\|DF_n\|$ there is based on the Volterra representation of B, which leads to long computations. In comparison our current Theorem 1.5 is more general, since it is valid for a wide class of Gaussian stationary sequences. Its proof is also significantly simplified. These are achieved by the introduction of two new ingredients in the proof, namely:

- A general formula to compute conditional expectations for random variables of the form $H_q(X_k)$.
- Related to the previous item, we heavily resort to the causal representation of X_k, which is particularly convenient in order to compute conditional expectations.

Here is how our paper is structured: we give some preliminary results concerning Gaussian stationary sequences and related Malliavin calculus in Section 2. We then prove our main Theorem 1.5 in Section 3.

2. Preliminaries

This section is devoted to some preliminaries on causal or moving average representations for Gaussian stationary sequences, as well as Malliavin calculus tools which will be used in the sequel.

2.1. Moving average representation.

The classical results on time series presented here are borrowed from [2, 4], to which we refer for further details. Start from our Gaussian stationary sequence $\{X_k;\, k \in \mathbb{N} \cup \{0\}\}$ with covariance function ρ. We will work under the following assumptions:

Hypothesis 2.1. *We suppose that ρ admits a spectral density called f_ρ, defined by*

$$f_\rho(\lambda) = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} \rho(k)\, e^{\imath k\lambda}, \qquad \lambda \in [-\pi, \pi]. \tag{5}$$

We further assume that

$$f_\rho \in L^{1/2}([-\pi,\pi]) \quad \text{and} \quad \log(f_\rho) \in L^1([-\pi,\pi]). \tag{6}$$

Notice that the condition $\log(f_\rho) \in L^1([-\pi,\pi])$ is referred to as purely nondeterministic *property in the literature.*

The interest of dealing with purely nondeterministic sequences is that they admit a so-called causal representation which is particularly convenient for conditional expectation computations. Let us state a result in this direction, which is taken from [4, Chapter 5].

Proposition 2.2. *Let X be a Gaussian stationary sequence satisfying Hypothesis 2.1. Then for each $k \in \mathbb{N} \cup \{0\}$ the random variable X_k can be decomposed as*

$$X_k = \sum_{j\geq 0} \psi_j\, w_{k-j}, \tag{7}$$

where $(w_k)_{k\in\mathbb{Z}}$ is a discrete Gaussian white noise and the coefficients ψ_j are deterministic. With a slight abuse of notation, extend the sequence ψ to $(\psi_j)_{j\in\mathbb{Z}}$ by setting $\psi_{-j} = 0$ for $j \geq 0$. Then one can choose ψ such that it enjoys the following properties:

(i) *The sequence ψ admits a spectral density f_ψ such that $f_\psi = \frac{f_\rho^{1/2}}{2\pi}$.*

(ii) *In particular, $\psi_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_\rho^{1/2}(\lambda)\, d\lambda$ and $\psi_0 > 0$.*

(iii) *For all $k_1, k_2 \in \mathbb{N}$ we have $\rho(k_1 - k_2) = \sum_{l=-\infty}^{k_1\wedge k_2} \psi_{k_1-l}\, \psi_{k_2-l}$.*

Proof. As mentioned above, those results are classic and borrowed from [4], see in particular formula (5.7.9) therein. Let us justify briefly the fact that $\psi_0 > 0$. Indeed, the purely nondeterministic assumption imposes $\log(f_\rho) \in L^1([-\pi,\pi])$, and thus $f_\rho > 0$ almost everywhere. This easily yields

$$\psi_0 = \int_{-\pi}^{\pi} f_\rho^{1/2}(\lambda)\, d\lambda > 0.$$

Item (iii) is obtained by writing $\mathbf{E}[X_{k_1} X_{k_2}]$ with the expression of X_k given by (7).

□

Let us now turn to examples for which our standing assumptions of Hypothesis 2.1 are met. The following proposition provides two typical and classical cases for which a spectral density exists and satisfies some integrability properties.

Proposition 2.3. *Let ρ be the covariance function of X. We have the following statements.*
(i) *If $\rho \in \ell^1$, then the spectral density f_ρ exists and is a nonnegative L^2 function defined on $[-\pi, \pi]$. Condition (5) is thus fulfilled.*
(ii) *If $\lim_{k\to\infty} |k|^\alpha \rho(k) = c_\rho$ for some $\alpha \in (0, 1)$ and some positive constant c_ρ, then the spectral density exists, is strictly positive almost everywhere and satisfies $\lim_{\lambda\to 0} |\lambda|^{1-\alpha} f_\rho(\lambda) = c_f$. In particular, condition (5) is satisfied.*

We now give two specific and important examples which satisfy Hypothesis 2.1.

Example 2.4. The so-called autoregressive fractionally integrated moving-average (ARFIMA) processes are introduced in [8] and [9]. Denote by B the one lag backward operator ($BX_k = X_{k-1}$). Let $\phi(z)$ and $\theta(z)$ be two polynomials which have no common zeros and such that the zeros of ϕ lie outside the closed unit disk $\{z\,,\ |z| \le 1\}$. Suppose that X_k is given by

$$\phi(B)X_k = (\mathrm{Id} - B)^{-d}\theta(B)w_k\,, \tag{8}$$

where $-1 < d < 1/2$, and where the operator $(\mathrm{Id} - B)^{-d}$ is defined by:

$$(\mathrm{Id} - B)^{-d} = \sum_{j=1}^{\infty} \eta_j B^j \quad \text{with} \quad \eta_j = \frac{\Gamma(d + j)}{\Gamma(j + 1)\Gamma(d)}\,.$$

Also notice that in (8) the sequence $(w_k)_{k\in\mathbb{Z}}$ is a discrete Gaussian white noise. It is well-known (see [19], Theorem 3.4 and equation (3.19)) that under the above conditions, $\{X_k, k \in \mathbb{N}\}$ admits a spectral density whose exact expression is:

$$f(\lambda) = \frac{1}{2\pi}\left[2\sin\frac{\lambda}{2}\right]^{-2d} \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}.$$

It is thus readily checked that the conditions (5) and (6) are satisfied, and hence X_k has a causal representation.

Example 2.5. Our second example is the fractional Gaussian noise. Let $\{B_t, t \ge 0\}$ be a fractional Brownian motion of Hurst parameter $H \in (0, 1)$. Then $\{X_k = B_{k+1} - B_k\,, k \in \mathbb{N} \cup \{0\}\}$ is a stationary Gaussian process with correlation

$$\rho(k) = \frac{1}{2}\left[(k + 1)^{2H} - 2k^{2H} - (k - 1)^{2H}\right]\,.$$

Its spectral density (see e.g. [2], equation (2.17)) is:

$$f(\lambda) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \rho(|k|)e^{i\lambda} = 2c_f(1 - \cos(\lambda)) \sum_{j=-\infty}^{\infty} |2\pi j + \lambda|^{-2H-1}\,, \quad \lambda \in [-\pi, \pi]\,,$$

where $c_f = (2\pi)^{-1}\sin(\pi H)\Gamma(2H + 1)$. If $H \le 1/2$, it is clear that $\sum_{k=-\infty}^{\infty} |\rho(|k|)| < \infty$. This implies

$$\sup_{\lambda\in[-\pi,\pi]} |f(\lambda)| < \infty\,.$$

Thus $f \in L^{1/2}$. If $1/2 < H < 1$, then

$$0 \le f(\lambda) \;\le\; 2c_f(1-\cos(\lambda))|\lambda|^{-2H-1} + 2c_f \sum_{j\neq 0} |2\pi j + \lambda|^{-2H-1}\,, \quad \lambda \in [-\pi, \pi]\,.$$

The first term is in L^1 since $H < 1$. When $j \neq 0$, $\int_{-\pi}^{\pi} |2\pi j + \lambda|^{-2H-1} d\lambda \le C j^{-2H}$ for some positive constant C. Thus $\int_{-\pi}^{\pi} \sum_{j\neq 0} |2\pi j + \lambda|^{-2H-1}\, d\lambda < \infty$, owing to the fact that $H > 1/2$. Therefore, we have $f \in L^1$ and hence $f \in L^{1/2}$. Summarizing we have $f \in L^{1/2}$ for all $H \in (0,1)$. This also implies $\log^+ f(\lambda) \in L^1$. To see $\log^- f(\lambda) \in L^1$, we notice that

$$f(\lambda) \ge 2c_f(1-\cos(\lambda))|\lambda|^{-2H-1}\,.$$

So $\log^- f(\lambda) \le C + \left|\log\left[(1-\cos(\lambda))|\lambda|^{-2H-1}\right]\right|$ which is in L^1. In conclusion, the sequence X satisfies Hypothesis 2.1.

2.2. **Malliavin calculus.** We start by briefly recalling some basic notation and results connected to Gaussian analysis and Malliavin calculus. The reader is referred to [13, 17] for details or missing proofs.

2.2.1. *Wiener space and generalities.* 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 $W = \{W(h) : h \in \mathfrak{H}\}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and such that $\mathbf{E}[W(h)W(g)] = \langle h, g\rangle_{\mathfrak{H}}$ for every $h, g \in \mathfrak{H}$.

In our application the underlying Gaussian family will be a discrete Gaussian white noise $(w_k)_{k\in\mathbb{Z}}$. The space $\mathfrak{H}$ is given here by $\mathfrak{H} = \ell^2(\mathbb{Z})$ (the space of square integrable sequences indexed by $\mathbb{Z}$) equipped with its natural inner product. Set $\{\varepsilon^j;\, j \in \mathbb{Z}\}$ for the canonical basis of $\ell^2(\mathbb{Z})$, that is $\varepsilon^j_k = \delta_j(k)$. We thus identify w_j with $W(\varepsilon^j)$. Assume from now on that our underlying σ-algebra $\mathcal{F}$ is generated by W.

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

$$H_q(x) = (-1)^q e^{\frac{x^2}{2}} \frac{d^q}{dx^q}\left(e^{-\frac{x^2}{2}}\right). \tag{9}$$

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

$$F = g(W(h_1), \ldots, W(h_n)),$$

where $n \ge 1$, $h_i \in \mathfrak{H}$, 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

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

By iteration, for every $m \ge 2$, we define the mth derivative $D^m F$. This is an element of $L^2(\Omega; \mathfrak{H}^{\odot m})$, where $\mathfrak{H}^{\odot m}$ designates the symmetric mth tensor product of $\mathfrak{H}$. For $m \ge 1$ and $p \ge 1$, $\mathbb{D}^{m,p}$ denote the closure of $\mathcal{S}$ with respect to the norm $\|\cdot\|_{m,p}$ defined by

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

Set $\mathbb{D}^\infty = \cap_{m,p}\mathbb{D}^{m,p}$. 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., [13, Theorem 2.7.2]), all the $\|\cdot\|_{m,p}$-norms are equivalent in any *finite* sum of Wiener chaoses.

Finally, let us recall that the Malliavin derivative D satisfies the following *chain rule*: if $\varphi : \mathbb{R}^n \to \mathbb{R}$ 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

$$D\varphi(F_1,\dots,F_n) = \sum_{i=1}^n \frac{\partial\varphi}{\partial x_i}(F_1,\dots,F_n)\, DF_i. \tag{10}$$

2.3. **Convergence of densities.** Suppose that F is a random variable in $\mathbb{D}^\infty$ such that $\mathbf{E}[\|DF\|^{-p}] < \infty$ for all $p \geq 1$. Then, we know that F has an infinitely differentiable density and there are explicit formulas for the density and its derivatives (see [17, Proposition 2.1.5]). Using this result, we can establish the following criteria for convergence of densities for random variables in a finite sum of Wiener chaos.

Proposition 2.6. *Let $\{F_n; n \in \mathbb{N}\}$ be a sequence of random variables belonging to a finite sum of Wiener chaos $\oplus_{k=1}^M \mathcal{H}_k$, which converges in law to a nonzero random variable F_∞. Suppose that, for all $p \geq 1$*

$$\sup_n \mathbf{E}[\|DF_n\|^{-p}] < \infty. \tag{11}$$

Then, for all $m \geq 0$ the derivative $p_n^{(m)}$ of the density of F_n, converges uniformly and in in $L^p(\mathbb{R})$ for all $p \geq 1$ to the corresponding derivative of the density of F_∞.

Proof. First notice that by condition (11), the random variable F_n has an infinitely differentiable density p_n, whose derivatives can be expressed by

$$p_n^{(m)}(x) = \mathbf{E}\left[\mathbf{1}_{\{F>x\}}\, G_n^{(m)}\right], \tag{12}$$

where the random variables $G_n^{(m)}$ are defined recursively by $G_n^{(0)} = \delta\left(\frac{DF_n}{\|DF_n\|_H^2}\right)$ and

$$G_n^{(m)} = -\delta\left(\frac{G_n^{(m-1)} DF_n}{\|DF_n\|_H^2}\right),$$

for any $m \geq 1$. From this formula it follows that the derivatives $p_n^{(m)}$ are uniformly bounded and also uniformly bounded in $L^p(\mathbb{R})$ for all $p \geq 1$. Indeed, by [16, Lemma 2.4] we have $\sup_n \mathbf{E}[|F_n|^r] < \infty$ for all $r \geq 1$. This uniform bound on the moments, together with the equivalence of the $\|\cdot\|_{m,p}$ norms in any finite sum of Wiener chaoses and condition (11) imply that $\sup_n \|G_n^{(m)}\|_{L^p(\Omega)} = c_{m,p} < \infty$ for all $m \geq 0$. Then, we can write from (12)

$$\sup_n \sup_x p_n^{(m)}(x) \leq \sup_n \mathbf{E}\left[|G_n^{(m)}|\right] = c_{m,1} < \infty,$$

and, using the fact that $\mathbf{E}[G_n^{(m)}] = 0$, we get:

$$\begin{aligned}\sup_n p_n^{(m)}(x) &\leq \sup_n \left(\mathbf{P}(|F_n| > |x|)\, \mathbf{E}\left[|G_n^{(m)}|^2\right]\right)^{\frac12} \\ &\leq c_{m,2} \sup_n \sqrt{\mathbf{P}(|F_n| > |x|)} \leq c_{m,2} \sup_n \mathbf{E}[|F_n|^q]^{\frac12} |x|^{-\frac{q}{2}} \leq c|x|^{-\frac{q}{2}},\end{aligned}$$

for all $q \geq 1$ and for some constant c depending on q and m.

By the results of [16], the densities p_n converge in $L^1(\mathbb{R})$ to the density p_∞ of F_∞. The boundedness in $L^p(\mathbb{R})$ and the uniform bound of p_n imply that this convergence holds in $L^p(\mathbb{R})$ for any $p \geq 1$. From a compactness argument we can deduce the convergence of all the derivatives in $L^p(\mathbb{R})$ for any $p \geq 1$. Finally, the uniform convergence is also easy to establish. □

2.4. **A key lemma.** Our future computations will heavily rely on an efficient way to compute conditional expectations. Towards this aim, we state here some general results. Let us start with a decomposition for Hermite polynomials:

Lemma 2.7. *For any $q \geq 1$, let H_q be the polynomial defined by relation* (9)*. Consider $y, z \in \mathbb{R}$ and two real parameters a, b with $a^2 + b^2 = 1$. Then the following relation holds true:*

$$H_q(ay + bz) = \sum_{\ell=0}^{q} \binom{q}{\ell} a^{q-\ell} b^{\ell}\, H_{q-\ell}(y)\, H_\ell(z). \tag{13}$$

Proof. By the definition of the Hermite polynomials, we have

$$e^{aty - \frac{(at)^2}{2}} = \sum_{i=0}^{\infty} (at)^i\, H_i(y), \quad \text{and} \quad e^{tbz - \frac{(bt)^2}{2}} = \sum_{j=0}^{\infty} (bt)^j\, H_j(z). \tag{14}$$

In the same way, we also obtain:

$$e^{t(ay+bz) - t^2/2} = \sum_{q=0}^{\infty} t^q H_q(ay + bz). \tag{15}$$

Since $a^2 + b^2 = 1$, we obviously have $e^{aty - \frac{(at)^2}{2}} e^{tbz - \frac{(bt)^2}{2}} = e^{t(ay+bz) - t^2/2}$. Thus, multiplying the right hand sides of (14) we recover the right hand side of (15), namely:

$$\sum_{q=0}^{\infty} t^q H_q(ay + bz) = \sum_{i=0}^{\infty} (at)^i H_i(y) \sum_{j=0}^{\infty} (bt)^j H_j(z),$$

which easily yields the desired identity (13). □

With this preliminary result in hand, we are ready to state our result on conditional expectations:

Proposition 2.8. *Let Y and Z be two centered Gaussian random variables such that Y is measurable with respect to a σ-algebra $\mathcal{G} \subset \mathcal{F}$ and Z is independent of $\mathcal{G}$. Assume that $\mathbf{E}[Y^2] = \mathbf{E}[Z^2] = 1$. Then for any $q \geq 1$, and real parameters a, b such that $a^2 + b^2 = 1$, we have:*

$$\mathbf{E}[H_q(aY + bZ)|\mathcal{G}] = a^q H_q(Y). \tag{16}$$

Proof. Apply identity (13) in order to decompose $H_q(aY + bZ)$. Then identity (16) follows easily from the fact that Y is $\mathcal{G}$-measurable, Z is independent from $\mathcal{G}$ and Hermite polynomials have 0 mean under a centered Gaussian measure except for $H_0 \equiv 1$. □

2.5. **Carbery-Wright inequality.** In the proof of Theorem 1.5, we will make use of the following inequality due to Carbery and Wright [5, Theorem 8], which is recalled here for convenience:

Proposition 2.9. *Let $X = (X_1, \ldots, X_n)$ be a Gaussian random vector in $\mathbb{R}^n$ and $Q : \mathbb{R}^n \to \mathbb{R}$ a polynomial of degree at most m. Then there is a universal constant $c > 0$ such that:*

$$\mathbf{E}[|Q(X_1, \ldots, X_n)|]^{\frac{1}{m}} \, \mathbf{P}(|Q(X_1, \ldots, X_n)| \leq x) \leq c \, m \, x^{\frac{1}{m}}, \quad \textit{for all} \quad x > 0. \tag{17}$$

3. Proof of Theorem 1.5

In this section, we will prove our main result, which amounts to show the inequality (4). This will be done into several steps.

Step 1: Computation of the Malliavin norm. Recall that $V_n^{d,q}$ is defined by relation (2), and set for the moment $f = \sum_{j=d}^{q} a_j \, H_j$. Invoking relation (10), plus the fact that $Dw_j = \varepsilon^j$ with the notation of Section 2.2.1, we get:

$$DV_n^{d,q} = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f'(X_k) \left(\sum_{j \geq 0} \psi_j \, \varepsilon^{k-j} \right) = \frac{1}{\sqrt{n}} \sum_{l=-\infty}^{n-1} \left(\sum_{k=l^+}^{n-1} \psi_{k-l} \, f'(X_k) \right) \varepsilon^l \,, \tag{18}$$

where $l^+ = \max\{l, 0\}$. Invoking Proposition 2.3, item (iii), it is thus readily checked that:

$$\left\langle DV_n^{d,q}, \, DV_n^{d,q} \right\rangle_{\mathfrak{H}} = \frac{1}{n} \sum_{k_1, k_2 = 0}^{n-1} f'(X_{k_1}) \rho(k_1 - k_2) f'(X_{k_2}),$$

where we recall that ρ is the covariance function of the Gaussian stationary sequence $\{X_k; \, k \geq 0\}$. This is consistent with the expression found in [13, Chapter 5]. However, in order to write the above expression as sum of some squares, we will start directly from expression (18). Since $\{\varepsilon^l; \, l \in \mathbb{Z}\}$ is an orthonormal basis of $\ell^2(\mathbb{Z})$ we obtain:

$$\left\langle DV_n^{d,q}, \, DV_n^{d,q} \right\rangle_{\mathfrak{H}} = \frac{1}{n} \sum_{\ell=-\infty}^{n-1} \left(\sum_{k=\ell^+}^{n-1} \psi_{k-\ell} \, f'(X_k) \right)^2 .$$

Rearranging terms (namely, change $k - \ell$ to k and then $n - \ell - 1$ to m), we end up with:

$$\begin{aligned} \left\langle DV_n^{d,q}, \, DV_n^{d,q} \right\rangle_{\mathfrak{H}} \ &\geq \ \frac{1}{n} \sum_{\ell=0}^{n-1} \left(\sum_{k=0}^{n-\ell-1} f'(X_{\ell+k}) \, \psi_k \right)^2 \\ &= \ \frac{1}{n} \sum_{m=0}^{n-1} \left(\sum_{k=0}^{m} f'(X_{n-1-(m-k)}) \, \psi_k \right)^2 \equiv A_n. \end{aligned}$$

As a last preliminary step we resort to the fact that $X = \{X_k; \, k \in \mathbb{N} \cup \{0\}\}$ is a Gaussian stationary sequence, which allows to assert that A_n is identical in law to B_n with

$$B_n := \frac{1}{n} \sum_{m=0}^{n-1} \left(\sum_{k=0}^{m} f'(X_{m-k}) \, \psi_k \right)^2 = \frac{1}{n} \sum_{m=0}^{n-1} \left(\sum_{k=0}^{m} f'(X_k) \, \psi_{m-k} \right)^2 .$$

We will now bound the negative moments of B_n.

Step 2: Block decomposition. We now wish to apply the Carbery-Wright inequality (17) in order to get bounds for negative moments of B_n. However, relation (17) only applies to moments of small order, and this is why we proceed to a decomposition of B_n into smaller blocks.

Fix thus an integer $N \geq 1$ and let $M = [n/N]$ be the integer part of n/N. Then $n \geq NM$ and as a consequence,

$$B_n = \frac{1}{n} \sum_{m=0}^{n-1} \left(\sum_{k=0}^{m} f'(X_k)\, \psi_{m-k} \right)^2 \geq \frac{1}{n} \sum_{i=0}^{N-1} \sum_{m=iM}^{(i+1)M-1} \left(\sum_{k=0}^{m} f'(X_k)\, \psi_{m-k} \right)^2 .$$

For $i = 0, \ldots, N-1$, define

$$B_n^i = \frac{1}{n} \sum_{m=iM}^{(i+1)M-1} \left(\sum_{k=0}^{m} f'(X_k)\, \psi_{m-k} \right)^2$$

so that $B_n \geq \sum_{i=0}^{N-1} B_n^i$. Then it is readily checked that:

$$(B_n)^{-\frac{p}{2}} \leq \prod_{i=0}^{N-1} (B_n^i)^{-\frac{p}{2N}}. \tag{19}$$

Recall once again the representation of the sequence X in (7), and denote by $\mathcal{F}_k$ the filtration generated by $\{w_\ell : \ell < k\}$. Then starting from (19) we obtain:

$$\mathbf{E}\left[(B_n)^{-\frac{p}{2}}\right] \leq \mathbf{E}\left[\prod_{i=0}^{N-1} (B_n^i)^{-\frac{p}{2N}}\right] = \mathbf{E}\left[\mathbf{E}\left[(B_n^{N-1})^{-\frac{p}{2N}} | \mathcal{F}_{(N-1)M}\right] \prod_{i=0}^{N-2} (B_n^i)^{-\frac{p}{2N}}\right]. \tag{20}$$

Step 3: Application of Carbery-Wright. Let us go back to the particular situation of $f = \sum_{j=d}^{q} a_j\, H_j$, which means in particular that $f' = \sum_{j=d}^{q} j\, a_j\, H_{j-1}$. We are now in a position to apply a conditional version of inequality (17) to the block $(B_n^{N-1})^{-\frac{p}{2N}}$ in (20). First, we notice

$$\mathbf{E}\left[(B_n^{N-1})^{-\frac{p}{2N}} | \mathcal{F}_{(N-1)M}\right] \leq 1 + \frac{p}{2N} \int_0^1 \mathbf{P}\left(B_n^{N-1} \leq x |\, \mathcal{F}_{(N-1)M}\right) x^{-\frac{p}{2N}-1} dx\,. \tag{21}$$

Since B_n^{N-1} is a polynomial of order $m = 2(q-1)$, Carbery-Wright's inequality (17) yields:

$$\mathbf{P}\left(B_n^{N-1} \leq x | \mathcal{F}_{(N-1)M}\right) \leq \frac{c\, x^{\frac{1}{2(q-1)}}}{\left[\mathbf{E}\left(B_n^{N-1} | \mathcal{F}_{(N-1)M}\right)\right]^{\frac{1}{2(q-1)}}}. \tag{22}$$

Step 4: Estimates for the conditional expectation. We now estimate the conditional expectation $\mathbf{E}[B_n^{N-1} | \mathcal{F}_{(N-1)M}]$. We have:

$$\begin{aligned}
\mathbf{E}\left[B_n^{N-1} | \mathcal{F}_{(N-1)M}\right] &= \frac{1}{n} \sum_{m=(N-1)M}^{NM-1} \mathbf{E}\left[\left(\sum_{k=0}^{m} f'(X_k)\, \psi_{m-k}\right)^2 \Big| \mathcal{F}_{(N-1)M}\right] \\
&\geq \frac{1}{n} \sum_{m=(N-1)M}^{NM-1} A_m,
\end{aligned} \tag{23}$$

where we have set

$$A_m = \mathbf{Var}\left(\sum_{k=(N-1)M}^{m} f'(X_k)\, \psi_{m-k} \middle| \mathcal{F}_{(N-1)M} \right).$$

Furthermore, notice that

$$f'(X_k) = f'\left(\sum_{\ell=-\infty}^{k} \psi_{k-i}\, w_i \right) = f'(Y_k + Z_k),$$

where $Y_k = \sum_{i=-\infty}^{(N-1)M-1} \psi_{k-i}\, w_i$ is $\mathcal{F}_{(N-1)M}$-measurable and $Z_k = \sum_{i=(N-1)M}^{k} \psi_{k-i}\, w_i$ is independent of $\mathcal{F}_{(N-1)M}$. Recalling that $f' = \sum_{j=d}^{q} j\, a_j\, H_{j-1}$, we can thus resort to Lemmas 2.7 and 2.8. This gives:

$$H_{q-1}(X_k) - \mathbf{E}[H_{q-1}(X_k)|\mathcal{F}_{(N-1)M}] = \sum_{j=d}^{q} \sum_{\ell=1}^{j-1} j^2 a_j \binom{j-1}{\ell} \sigma_{Y_k}^{j-1-\ell} H_{j-1-\ell}(\widetilde{Y}_k) \sigma_{Z_k}^{\ell} H_\ell(\widetilde{Z}_k),$$

where $\sigma_{Y_k} = [\mathbf{Var}(Y_k)]^{1/2}$, $\sigma_{Z_k} = [\mathbf{Var}(Z_k)]^{1/2}$, $\widetilde{Y}_k = Y_k/\sigma_{Y_k}$ and $\widetilde{Z}_k = Z_k/\sigma_{Z_k}$. Therefore,

$$\begin{aligned} A_m &= \mathbf{E}\left[\left(\sum_{k=(N-1)M}^{m} \sum_{j=d}^{q} \sum_{\ell=1}^{j-1} a_{j,\ell,k}\, H_{j-1-\ell}(\widetilde{Y}_k)\, H_\ell(\widetilde{Z}_k)\, \psi_{m-k} \right)^2 \middle| \mathcal{F}_{(N-1)M} \right] \\ &= \mathbf{E}\left[\left(\sum_{\ell=1}^{q-1} \sum_{k=(N-1)M}^{m} \sum_{j=(\ell+1)\vee d}^{q} a_{j,\ell,k}\, H_{j-1-\ell}(\widetilde{Y}_k)\, H_\ell(\widetilde{Z}_k)\, \psi_{m-k} \right)^2 \middle| \mathcal{F}_{(N-1)M} \right], \end{aligned}$$

where we have set $a_{j,\ell,k} = j^2 a_j \binom{j-1}{\ell} \sigma_{Y_k}^{j-1-\ell} \sigma_{Z_k}^{\ell}$.

Recall that the random variables $\widetilde{Y}_k$ are $\mathcal{F}_{(N-1)M}$-measurable while the random variables $\widetilde{Z}_k$ are independent of $\mathcal{F}_{(N-1)M}$. By decorrelation properties of Hermite polynomials we thus get:

$$A_m = \sum_{\ell=1}^{q-1} \mathbf{E}\left[\left(\sum_{k=(N-1)M}^{m} \sum_{j=(\ell+1)\vee d}^{q} a_{j,\ell,k}\, H_{j-1-\ell}(\widetilde{Y}_k)\, H_\ell(\widetilde{Z}_k)\, \psi_{m-k} \right)^2 \middle| \mathcal{F}_{(N-1)M} \right]$$

and we trivially lower bound this quantity by taking the term corresponding to $\ell = q-1$. In this situation the sum $\sum_{j=(\ell+1)\vee d}^{q}$ is reduced to the term corresponding to $j = q$, and since $a_{q,q-1,k} = q^2 a_q \sigma_{Z_k}^{q-1}$ we obtain:

$$\begin{aligned} A_m &\geq \mathbf{E}\left[\left(\sum_{k=(N-1)M}^{m} q^2\, a_q\, \sigma_{Z_k}^{q-1}\, H_{q-1}(\widetilde{Z}_k)\, \psi_{m-k} \right)^2 \middle| \mathcal{F}_{(N-1)M} \right] \\ &= q^4\, a_q^2\; \mathbf{E}\left[\left(\sum_{k=(N-1)M}^{m} \sigma_{Z_k}^{q-1}\, H_{q-1}(\widetilde{Z}_k)\, \psi_{m-k} \right)^2 \right]. \end{aligned}$$

We now invoke the identity $\mathbf{E}[H_p(\tilde{Z}_{k_1})H_p(\tilde{Z}_{k_2})] = \frac{1}{p!}(\mathbf{E}[\tilde{Z}_{k_1}\tilde{Z}_{k_2}])^p$ in order to obtain

$$A_m \geq \frac{q^5\, a_q^2}{q!} \sum_{k_1,k_2=(N-1)M}^{m} \sigma_{Z_{k_1}}^{q-1}\, \sigma_{Z_{k_2}}^{q-1}\, \mathbf{E}\left[\widetilde{Z}_{k_1}\, \widetilde{Z}_{k_2}\right]^{q-1} \psi_{m-k_1}\psi_{m-k_2}.$$

Furthermore, it is readily checked that:

$$\mathbf{E}\left[\widetilde{Z}_{k_1}\, \widetilde{Z}_{k_2}\right] = \frac{1}{\sigma_{Z_{k_1}}\, \sigma_{Z_{k_2}}} \sum_{i=(N-1)M}^{k_1\wedge k_2} \psi_{k_1-i}\, \psi_{k_2-i},$$

and thus

$$\begin{aligned} A_m &\geq \frac{q^5\, a_q^2}{q!} \sum_{k_1,k_2=(N-1)M}^{m} \left(\sum_{i=(N-1)M}^{k_1\wedge k_2} \psi_{k_1-i}\, \psi_{k_2-i} \right)^{q-1} \psi_{m-k_1}\psi_{m-k_2} \\ &= \frac{q^5\, a_q^2}{q!} \sum_{i_1,\dots,i_{q-1}=(N-1)M}^{m} \ \sum_{k_1,k_2=\max(i_1,\dots,i_{q-1})}^{m} \psi_{m-k_1}\psi_{m-k_2} \prod_{j=1}^{q-1} \psi_{k_1-i_j}\, \psi_{k_2-i_j} \\ &= \frac{q^5\, a_q^2}{q!} \sum_{i_1,\dots,i_{q-1}=(N-1)M}^{m} \left(\sum_{k=\max(i_1,\dots,i_{q-1})}^{m} \psi_{m-k} \prod_{j=1}^{q-1} \psi_{k-i_j} \right)^2. \end{aligned}$$

Here again, this sum of squares is trivially lower bounded by taking the term corresponding to $i_1 = \cdots = i_{q-1} = m$, which yields:

$$A_m \geq c_{a,q,\psi} \quad \text{with} \quad c_{a,q,\psi} \equiv \frac{q^5\, a_q^2}{q!}\, \psi_0^{2q} > 0. \tag{24}$$

Step 5: Conclusion. In the remainder of the proof the constants $c_{a,q,\psi,N}$ and so can change from line to line without further mention. Plugging relation (24) into (23) and recalling that N is a given integer whose exact value will be fixed below, we get:

$$\mathbf{E}\left[B_n^{N-1}|\mathcal{F}_{(N-1)M}\right] \geq \frac{M\, c_{a,q,\psi}}{n} \geq c_{a,q,\psi,N} > 0,$$

as long as N stays bounded. We then insert back this inequality into (21) and (22) in order to get:

$$\mathbf{P}\left(B_n^{N-1} \leq x|\mathcal{F}_{(N-1)M}\right) \leq 1 + \frac{p\, c_{a,q,\psi,N}}{2N} \int_0^1 x^{\frac{1}{2(q-1)} - \frac{p}{2N} - 1} dx = c_{a,q,\psi,N,p} < \infty,$$

where we have chosen N such that $\frac{p}{2N} < \frac{1}{2(q-1)}$. Iterating this bound into (20), we have thus obtained:

$$\mathbf{E}\left[(B_n)^{-\frac{p}{2}}\right] \leq c_{a,q,\psi,N,p}^{N},$$

which is a finite quantity. Finally recall from Step 1 that $\mathbf{E}[(B_n)^{-\frac{p}{2}}] = \mathbf{E}[\|DV_n^{d,q}\|_{\mathfrak{H}}^{-p}]$, which finishes the proof.

References

[1] M. A. Arcones: Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.* **22** (1994), 2242–2274.

[2] J. Beran: *Statistics for long-memory processes.* Chapman and Hall, 1994.

[3] P. Breuer and P. Major: Central limit theorems for non-linear functionals of Gaussian fields. *J. Mult. Anal.* **13** (1983), 425–441.

[4] P. Brockwell and R. Davis: *Time series: theory and methods.* Second edition. Springer, 1991.

[5] A. Carbery and J. Wright: Distributional and L^q norm inequalities for polynomials over convex bodies in $\mathbb{R}^n$. *Math. Research Lett.* **8** (2001), 233–248.

[6] D. Chambers and E. Slud: Central limit theorems for non-linear functionals of stationary Gaussian processes. *Probab. Theory Related Fields* **80** (1989), no. 3, 323–346.

[7] L. Giriatis and D. Surgailis: CLT and other limit theorems for functionals of Gaussian processes. *Z. Wahrsch. Verw. Gebiete* **70** (1985), no. 2, 191–212.

[8] C. W. J. Granger and R. Joyeux: An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.* **1** (1980), no. 1, 15–29.

[9] J. R. M. Hosking: Fractional differencing. *Biometrika* **68** (1981), 165–176.

[10] Y. Hu, F. Lu and D. Nualart: Convergence of densities of some functionals of Gaussian processes. *J. Funct. Anal.* **266** (2014), 814–875.

[11] I. Nourdin and D. Nualart: Fisher Information and the Fourth Moment Theorem. Preprint. Arxiv 1312.5841.

[12] I. Nourdin and G. Peccati: Stein's method on Wiener chaos. *Probab. Theory Rel. Fields* **145** (2009), 75–118.

[13] I. Nourdin and G. Peccati: *Normal Approximations using Malliavin Calculus: from Stein's Method to the Universality*. Cambridge University Press, 2012.

[14] I. Nourdin, G. Peccati and M. Podolskij: Quantitative Breuer-Major theorems. *Stoch Process. Appl.* **121** (2011), 793–812.

[15] I. Nourdin, G. Peccati and Y. Swan: Entropy and the fourth moment phenomenon. *J. Funct. Anal.* **266** (2014), 3170–3207.

[16] I. Nourdin and G. Poly: Convergence in total variation on Wiener chaos. *Stochastic Process. Appl.* **123** (2013), no. 2, 651–674.

[17] D. Nualart: *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[18] D. Nualart and G. Peccati: Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** (2005), no. 1, 177–193.

[19] W. Palma: Long-memory time series. Theory and methods. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2007.

Yaozhong Hu
Department of Mathematics
University of Kansas
Lawrence, Kansas 66045 USA
hu@math.ku.edu

David Nualart
Department of Mathematics
University of Kansas
Lawrence, Kansas 66045 USA
nualart@math.ku.edu

Samy Tindel
Institut Elie Cartan
Université de Lorraine
54506 Vandoeuvre-lès-Nancy, France
samy.tindel@univ-lorraine.fr

Fangjun Xu
School of Finance and Statistics
East China Normal University
Shanghai, China 200241
fangjunxu@gmail.com