Electron. Commun. Probab. **29** (2024), article no. 13, 1–8.
https://doi.org/10.1214/24-ECP583
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

A total variation version of Breuer–Major Central Limit Theorem under $\mathbb{D}^{1,2}$ assumption*

Jürgen Angst† Federico Dalmao‡ Guillaume Poly†

Abstract

In this note, we establish a qualitative total variation version of Breuer–Major Central Limit Theorem for a sequence of the type $\frac{1}{\sqrt{n}}\sum_{1\leq k\leq n} f(X_k)$, where $(X_k)_{k\geq 1}$ is a centered stationary Gaussian process, under the hypothesis that the function f has Hermite rank $d \geq 1$ and belongs to the Malliavin space $\mathbb{D}^{1,2}$. This result in particular extends the recent works of [NNP21], where a quantitative version of this result was obtained under the assumption that the function f has Hermite rank $d = 2$ and belongs to the Malliavin space $\mathbb{D}^{1,4}$. We thus weaken the $\mathbb{D}^{1,4}$ integrability assumption to $\mathbb{D}^{1,2}$ and remove the restriction on the Hermite rank of the base function. While our method is still based on Malliavin calculus, we exploit a particular instance of Malliavin gradient called the sharp operator, which reduces the desired convergence in total variation to the convergence in distribution of a bidimensional Breuer–Major type sequence.

Keywords: Breuer–Major CLT; total variation distance; Malliavin calculus; Stein's equation.
MSC2020 subject classifications: 60F05; 60G10; 60H07.
Submitted to ECP on September 12, 2023, final version accepted on March 4, 2024.

1 Framework and main result

Let us consider $X = (X_n)_{n\geq 1}$ a real-valued centered stationary Gaussian sequence with unit variance, defined on an abstract probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\rho : \mathbb{N} \to \mathbb{R}$ be the associated correlation function, in other words $\rho(|k-\ell|) = \mathbb{E}[X_k X_\ell]$, for all $k, \ell \geq 1$. We will also classically denote by $\mathcal{N}(0, \sigma^2)$ the law of a centered normal variable with variance σ^2. Set $\gamma(dx) := (2\pi)^{-1/2} e^{-x^2/2} dx$ the standard Gaussian measure on the real line and $\gamma_d = \otimes_{k=1}^{d} \gamma$ its analogue in $\mathbb{R}^d$. We then denote by $(H_m)_{m\geq 0}$ the family of Hermite polynomials which are orthogonal with respect to γ, namely $H_0 \equiv 1$ and

$$H_m(x) := (-1)^m e^{\frac{x^2}{2}} \frac{d^m}{dx^m} e^{-\frac{x^2}{2}}, \quad m \geq 1.$$

We denote by $L^2(\mathbb{R}, \gamma)$ the space of square integrable real functions with respect to the Gaussian measure. Recall that a real function $f \in L^2(\mathbb{R}, \gamma)$ is said to have Hermite rank $d \geq 0$ if it can be decomposed as a sum of the form

$$f(x) = \sum_{m=d}^{+\infty} c_m H_m(x), \quad c_d \neq 0.$$

*This work was supported by the ANR grant UNIRANDOM, ANR-17-CE40-0008.
†Univ Rennes, CNRS, IRMAR – UMR 6625, F-35000 Rennes, France.
E-mail: jurgen.angst@univ-rennes1.fr, guillaume.poly@univ-rennes1.fr
‡DMEL, Cenur LN, Universidad de la República, Salto, Uruguay. E-mail: fdalmao@unorte.edu.uy

For integers $k, p \geq 1$, we further denote by $\mathbb{D}^{k,p}(\mathbb{R},\gamma)$ the Malliavin–Sobolev space consisting of the completion of the family of polynomial functions $q : \mathbb{R} \to \mathbb{R}$ with respect to the norm

$$||q||_{k,p} := \left| \int_{\mathbb{R}} \left(|q(x)|^p + \sum_{\ell=1}^{k} |q^{(\ell)}(x)|^p \right) \gamma(dx) \right|^{1/p},$$

where $q^{(\ell)}$ is the ℓ-th derivative of q. Given a real function f, let us finally set

$$S_n(f) := \frac{1}{\sqrt{n}} \sum_{k=1}^{n} f(X_k).$$

In this framework, the celebrated Central Limit Theorem (CLT) by Breuer and Major gives sufficient conditions on ρ and f so that the sequence $S_n(f)$ satisfies a CLT.

Theorem 1 (Theorem 1 in [BM83]). If the function f belongs to $L^2(\mathbb{R},\gamma)$ with Hermite rank $d \geq 1$ and if $\rho \in \ell^d(\mathbb{N})$, i.e. $\sum_{\mathbb{N}} |\rho(k)|^d < +\infty$, then the sequence $(S_n(f))_{n\geq 1}$ converges in distribution as n goes to infinity to a normal distribution $\mathcal{N}(0,\sigma^2)$, where the limit variance is given by

$$\sigma^2 := \sum_{m=d}^{\infty} m! c_m^2 \sum_{k\in\mathbb{Z}} \rho(k)^m,$$

with $(c_m)_{m\geq d}$ being the coefficients appearing in the Hermite expansion of f.

Recently, under mild additional assumptions, a series of articles has reinforced the above convergence in distribution into a convergence in total variation, with polynomial quantitative bounds, see e.g. [KN19, NPY19, NZ21, NNP21]. Recall that the total variation distance between the distributions of two real random variables X and Y is given by

$$d_{\mathrm{TV}}(X,Y) := \sup_{A\in\mathcal{B}(\mathbb{R})} |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|,$$

where the supremum runs over $\mathcal{B}(\mathbb{R})$, the Borel sigma field on the real line. To the best of our knowledge, the best statement so far in this direction is the following

Theorem 2 (Theorem 1.2 in [NNP21]). Assume that $f \in L^2(\mathbb{R},\gamma)$ has Hermite rank $d = 2$ and that it belongs to $\mathbb{D}^{1,4}(\mathbb{R},\gamma)$. Suppose that $\rho \in \ell^d(\mathbb{N})$ and that the variance σ^2 of Theorem 1 is positive. Then, there exists a constant $C > 0$ independent of n such that

$$d_{\mathrm{TV}}\left(\frac{S_n(f)}{\sqrt{\mathrm{var}(S_n(f))}}, \mathcal{N}(0,1) \right) \leq \frac{C}{\sqrt{n}} \left[\left(\sum_{|k|\leq n} |\rho(k)| \right)^{\frac{1}{2}} + \left(\sum_{|k|\leq n} |\rho(k)|^{\frac{4}{3}} \right)^{\frac{3}{2}} \right].$$

The goal of this note is to establish that the convergence in total variation in fact holds as soon as the function f is in the Malliavin–Sobolev space $\mathbb{D}^{1,2}(\mathbb{R},\gamma)$ and has Hermite rank $d \geq 1$.

Theorem 3. Suppose that $f \in \mathbb{D}^{1,2}(\mathbb{R},\gamma)$ has Hermite rank $d \geq 1$. Suppose moreover that $\rho \in \ell^d(\mathbb{N})$ and that the variance σ^2 of Theorem 1 is positive. Then, as n goes to infinity

$$d_{\mathrm{TV}}\left(\frac{S_n(f)}{\sqrt{\mathrm{var}(S_n(f))}}, \mathcal{N}(0,1) \right) \xrightarrow[n\to+\infty]{} 0.$$

Note that, for the sake of simplicity, we only consider here a real Gaussian sequence $(X_n)_{n\geq 1}$ and a real function f but our method is robust and would yield, under similar covariance and rank assumptions, a convergence in total variation for a properly renomalized sequence of the type $\sum_{k=1}^{n} f(X_k^1,\ldots,X_k^d)$ associated with a sequence of Gaussian vectors $(X_n)_{n\geq 1}$ with values in $\mathbb{R}^d$ and a function f in the corresponding Malliavin–Sobolev space $\mathbb{D}^{1,2}(\mathbb{R}^d,\gamma_d)$.

Overall strategy of the proof and novelty

The detailed proof of Theorem 3 is the object of the next section and the rest of the paper. Let us sketch here the overall strategy of the proof and compare it to existing approaches. As done in the above mentioned references, in order to establish the CLT in total variation, we use the global Malliavin–Stein approach, namely we will show that there exists a constant ν such that

$$\limsup_{n\to+\infty} \sup_{\phi} \left|\mathbb{E}\left[S_n(f)\phi(S_n(f))\right] - \nu\, \mathbb{E}\left[\phi'(S_n(f))\right]\right| = 0,$$

where the supremum is taken over C^1 function ϕ with bounded derivatives. Indeed, this will show both the Gaussianity of the limit and the convergence in total variation. To do so, also as classically done, we use an integration by parts formula to give an alternative expression for $\mathbb{E}\left[S_n(f)\phi(S_n(f))\right] = \mathbb{E}\left[\phi'(S_n(f))\Gamma[F_n, -\mathcal{L}^{-1}S_n(f)]\right]$, where Γ is the square field operator and $\mathcal{L}^{-1}$ is the pseudo-inverse Ornstein–Uhlenbeck operator, see Section 2.1 below.

It is well known that several Malliavin gradients can be associated to the same $\Gamma-$calculus. The main novelty of our approach and its efficiency then lie in the choice of a particular gradient, the so-called sharp operator, whose definition is recalled in Section 2.2, to express and control the quantity $\Gamma[S_n(f), -\mathcal{L}^{-1}S_n(f)]$. Given a random $F \in \mathbb{D}^{1,2}$, the key formula (2.2) below indeed allows to relate the Fourier transform of the gradient $^\sharp F$ to the Laplace transform of $\Gamma(F,F)$. With this tool at hand, one deduces that the convergence in law (hence in probability) of $\Gamma[S_n(f), -\mathcal{L}^{-1}S_n(f)]$ towards the constant ν is equivalent to the convergence in distribution towards a constant vector of the two-dimensional vector $(^\sharp S_n(f), {}^\sharp\mathcal{L}^{-1}S_n(f))$. But again, thanks to our particular choice of gradient and with no further assumption on the regularity/integrability of f or its Hermite rank, the latter convergence is an immediate consequence of the standard two-dimensional Breuer–Major Theorem, see Section 2.3.

The end of the proof then consists in showing that the convergence in probability of $\Gamma[S_n(f), -\mathcal{L}^{-1}S_n(f)]$ towards ν can be reinforced to a convergence in L^1, which can be done by elementary uniform integrability estimates, via hypercontractivity arguments, see Section 2.4.

In comparison with the recent references [NZ21, NNP21], our approach only provides a qualitative CLT in total variation. Indeed, under the sole $\mathbb{D}^{1,2}$ assumption, one cannot use Malliavin derivatives of order two or higher, nor use any Hölder type inequality. These last tools are precisely the ones used in the above references to provide quantitative bounds on remainders, but at the cost of requiring more regularity and integrability. Otherwise, our direct approach and choice of gradient show that the Hermite rank of the base function plays no significant role in the CLT.

2 Proof of the main result

As mentioned just above, the setting of the proof of Theorem 3 is the one of Malliavin–Stein calculus. Note that for each fixed $n \geq 1$, the quantity of interest $S_n(f)$ involves only a finite number of Gaussian coefficients. So let us sketch the framework of Malliavin–Stein method in the finite dimensional setting, and we refer to [Nua09] or [NP12] for a more general introduction.

2.1 A glimpse of Malliavin calculus

Let us fix an integer $n \geq 1$ and let us place ourselves in the product probability space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \gamma_n)$ with $\gamma_n := \otimes_{k=1}^n \gamma$, the n-dimensional standard Gaussian distribution on $\mathbb{R}^n$. Consider the classical *Ornstein–Ulhenbeck* operator $\mathcal{L}_n := \Delta - \vec{x}\cdot\nabla$ which is symmetric with respect to γ_n. We have then the standard decomposition of the L^2-space in Wiener chaoses, namely

$$\begin{aligned} L^2(\gamma_n) &= \bigoplus_{k=0}^{\infty} \mathrm{Ker}\left(\mathcal{L}_n + k\mathrm{I}\right), \quad \text{with} \\ \mathrm{Ker}\left(\mathcal{L}_n + k\mathrm{I}\right) &= \mathrm{Vect}\left(\prod_{i=0}^{n} H_{k_i}(x_i) \Big| \sum_{i=0}^{n} k_i = k\right) := \underbrace{\mathcal{W}_k.}_{k\text{-th Wiener chaos}} \end{aligned}$$

The square field or “carré du champ” operator Γ_n is then defined as the bilinear operator $\Gamma_n := [\cdot,\cdot] = \nabla\cdot\nabla$. As a glimpse of the power of Malliavin–Stein approach in view of establishing total variation estimates, recall that if $F \in \mathrm{Ker}\left(\mathcal{L}_n + k\mathrm{I}\right)$ is such that $\mathbb{E}[F^2]=1$, then for some constant C_k only depending on k, the total variation distance between the variable F and a standard Gaussian can be upper bounded by

$$d_{TV}\left(F, \mathcal{N}(0,1)\right) \leq C_k \sqrt{\mathrm{var}\left(\Gamma\left[F,F\right]\right)}.$$

Via the notion of isonormal Gaussian process, the finite dimensional framework for Malliavin–Stein method sketched above can in fact be extended to the infinite dimensional setting giving rise to an Ornstein–Uhlenbeck operator $\mathcal{L}$ and an associated “carré du champ” Γ, see e.g. Chapter 2 in [NP12].

2.2 The sharp gradient

A detailed introduction to the sharp gradient can be found in Section 4.1 of the reference [AP20]. We only recall here the basics which will be useful to our purpose. Let us assume that $(N_k)_{k\geq 1}$ is an i.i.d. sequence of standard Gaussian variables on $(\Omega,\mathcal{F},\mathbb{P})$ which generate the first Wiener chaos. Without loss of generality, we shall assume that $\mathcal{F}=\sigma(N_k,\ k\geq 1)$. We will also need a copy $(\hat{\Omega},\hat{\mathcal{F}},\hat{\mathbb{P}})$ of this probability space as well as $(\hat{N}_i)_{i\geq 1}$ a corresponding i.i.d. sequence of standard Gaussian variables such that $\hat{\mathcal{F}}=\sigma(\hat{N}_k,\ k\geq 1)$. We will denote by $\hat{\mathbb{E}}$ the expectation with respect to the measure $\hat{\mathbb{P}}$. For any integer $m\geq 1$ and any function Φ in the space $\mathcal{C}_b^1(\mathbb{R}^m,\mathbb{R})$ of continuously differentiable functions with a bounded gradient, we then set

$$^\sharp\Phi(N_1,\cdots,N_m) := \sum_{i=1}^{m} \partial_i\Phi(N_1,\cdots,N_m)\hat{N}_i. \tag{2.1}$$

In Sections 4.1.1 and 4.1.2 of [AP20], it is shown that this *gradient* is closable and extends to the Malliavin space $\mathbb{D}^{1,2}$, where

$$\mathbb{D}^{1,2} := \left\{F\in\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{P}),\ \mathbb{E}[F^2]+\mathbb{E}\left[(^\sharp F)^2\right] < +\infty\right\}.$$

The last space $\mathbb{D}^{1,2}$ is naturally the infinite dimensional version of the Malliavin–Sobolev space $\mathbb{D}^{1,2}(\mathbb{R},\gamma)$ introduced in Section 1 in the one-dimensional setting. In particular, Proposition 8 in the latter reference shows that

$$\forall F\in\mathbb{D}^{1,2},\ \forall\phi\in\mathcal{C}_b^1(\mathbb{R},\mathbb{R})\ :\ ^\sharp\phi(F)=\phi'(F)^\sharp F.$$

Given $F\in\mathbb{D}^{1,2}$, taking first the expectation $\hat{\mathbb{E}}$ with respect $\hat{\mathbb{P}}$ and using Fubini inversion of sums yields the following key relation, for all $\xi\in\mathbb{R}$

$$\mathbb{E}\left(\exp\left(-\frac{\xi^2}{2}\Gamma[F,F]\right)\right) = \hat{\mathbb{E}}\mathbb{E}\left(\exp\left(i\xi^\sharp F\right)\right). \tag{2.2}$$

By essence, via their Laplace/Fourier transforms, this key equation allows to relate the asymptotic behavior in distribution (or in probability if the limit is constant) of the carré du champ $\Gamma[F,F]$ with the one of the sharp gradient $^\sharp F$.

Finally, let us remark that by definition, the image $({}^\sharp X_k)_{k\geq 1}$ of our initial stationary sequence $(X_k)_{k\geq 1}$ by the sharp gradient is an independent copy of $(X_k)_{k\geq 1}$. We will write $({}^\sharp X_k)_{k\geq 1} = (\hat{X}_k)_{k\geq 1}$ in the sequel.

2.3 Convergence in probability via a two dimensional CLT

Let us suppose that f satisfies the assumptions of Theorem 3, namely $f \in \mathbb{D}^{1,2}(\mathbb{R},\gamma)$ with Hermite rank $d \geq 1$, so that it can be decomposed as $f = \sum_{m=d}^{\infty} c_m H_m$ in $L^2(\mathbb{R},\gamma)$. Let $\mathcal{L}^{-1}$ denote the pseudo-inverse of the Ornstein–Uhlenbeck operator and consider the pre-image

$$g(x) := -\mathcal{L}^{-1}[f](x) = \sum_{m=d}^{\infty} \frac{c_m}{m} H_m(x).$$

To simplify the expressions in the sequel, we set

$$F_n := S_n(f) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} f(X_k), \ \text{ and } \ G_n := S_n(g) = -\mathcal{L}^{-1} F_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} g(X_k).$$

Now, take $(s,t,\xi) \in \mathbb{R}^3$ and let us apply the above key relation (2.2) with the random variable $tF_n + sG_n$, we get

$$\mathbb{E}\left[\exp\left(-\frac{\xi^2}{2}\Gamma[tF_n + sG_n, tF_n + sG_n]\right)\right] = \hat{\mathbb{E}}\mathbb{E}\left[\exp\left(i\xi\left(t\,{}^\sharp F_n + s\,{}^\sharp G_n\right)\right)\right]. \tag{2.3}$$

On the one hand, by bilinearity of the carré du champ operator, we have

$$\Gamma[tF_n + sG_n, tF_n + sG_n] = t^2\Gamma[F_n,F_n] + s^2\Gamma[G_n,G_n] + 2ts\Gamma[F_n, -\mathcal{L}^{-1}F_n]. \tag{2.4}$$

On the other hand, the right hand side of Equation (2.3) is simply the characteristic function under $\mathbb{P}\otimes\hat{\mathbb{P}}$ of the variable $t^\sharp F_n + s^\sharp G_n$ and one remarks that

$$t^\sharp F_n + s^\sharp G_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left(tf'(X_k) + sg'(X_k)\right)\hat{X}_k = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \Psi_{s,t}(X_k, \hat{X}_k)$$

is another “Breuer–Major type” sequence with respect to the $\mathbb{R}^2-$ valued centered stationary Gaussian process $(\hat{X}_k, X_k)_{k\geq 1}$ and function $\Psi_{s,t} : \mathbb{R}^2 \to \mathbb{R}$

$$(x,y) \mapsto \Psi_{s,t}(x,y) := \left(tf'(x) + sg'(x)\right) y.$$

Since f is in $\mathbb{D}^{1,2}(\mathbb{R},\gamma)$, its derivative f' is in $L^2(\mathbb{R},\gamma)$ and so is g', so that $\Psi_{s,t}$ is in $L^2(\mathbb{R}^2,\gamma_2)$ with rank $k \geq 1$ for non vanishing (s,t). Moreover $(\hat{X}_k)_{k\geq 1}$ is an independent copy of $(X_k)_{k\geq 1}$ so that their cross correlations vanish, therefore the multivariate counterpart of the classical Breuer–Major Theorem applies, see Theorem 4 of [Arc94]. As a consequence, for any $(t,s) \in \mathbb{R}^2$, $t^\sharp F_n + s^\sharp G_n$ converge in distribution as n goes to infinity to an explicit centered Gaussian variable.

As a result, the bidimensional sequence $({}^\sharp F_n, {}^\sharp G_n)$ converges in distribution, under $\mathbb{P}\otimes\hat{\mathbb{P}}$, towards a bidimensional centered Gaussian vector with a symmetric semi-positive covariance matrix Σ. Therefore, from Equations (2.3) and (2.4) and via the characterization of convergence in distribution in terms of Fourier transform, there exist real numbers λ, μ, ν (depending on the limit covariance matrix Σ) such that for any $(s,t,\xi) \in \mathbb{R}^3$, as n goes to infinity, we have

$$\mathbb{E}\left[e^{-\frac{\xi^2 t^2}{2}\Gamma[F_n,F_n] - \frac{\xi^2 s^2}{2}\Gamma[G_n,G_n] - \xi^2 ts\Gamma[F_n,-\mathcal{L}^{-1}F_n]}\right] \xrightarrow[n\to\infty]{} e^{-\frac{\xi^2}{2}\left(\lambda t^2 + \mu s^2 + 2\nu ts\right)}.$$

Since the above convergence is valid for any $\xi \in \mathbb{R}$, this shows in particular that for any fixed $(s,t) \in \mathbb{R}^2$, the sequence $\Gamma[tF_n + sG_n, tF_n + sG_n]$ converges in distribution (and thus in probability) towards the constant variable $\left(\lambda t^2 + \mu s^2 + 2\nu ts\right)$. Choosing $s = t = 1$, we thus get that $\Gamma[F_n + G_n, F_n + G_n]$ converges in probability towards $(\lambda + \mu + 2\nu)$. Choosing $s = 0$ and $t = 1$, then $t = 0$ and $s = 1$, one deduce in the same manner that $\Gamma[F_n, F_n]$ and $\Gamma[G_n, G_n]$ both converge in probability towards λ and μ respectively. Finally, by Equation (2.4), one can conclude that the cross term

$$\Gamma[F_n, G_n] = \Gamma(F_n, -\mathcal{L}^{-1}F_n) = \hat{\mathbb{E}}\left[{}^\sharp F_n {}^\sharp G_n\right]$$

also converges in probability towards the constant limit variable ν.

2.4 Gaining some uniform integrability

Since our goal is to derive convergence in total variation of $F_n = S_n(f)$, the convergence in probability of the term $\Gamma[F_n, -\mathcal{L}^{-1}F_n]$ is not sufficient. Indeed, with Stein's Equation in mind, the lack of uniform integrability is a problem to deduce the following required asymptotic behavior for any $\phi \in \mathcal{C}_b^1(\mathbb{R})$, as n goes to infinity

$$\mathbb{E}\left[\phi'(F_n)\Gamma[F_n, -\mathcal{L}^{-1}F_n]\right] \approx \nu\, \mathbb{E}\left[\phi'(F_n)\right].$$

In order to bypass this problem, let us go back to the two-dimensional version of the classical Breuer–Major theorem used in the last section. With the above notations, we have

$${}^\sharp F_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n \Psi_{1,0}(X_k, \hat{X}_k), \quad {}^\sharp G_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n \Psi_{0,1}(X_k, \hat{X}_k).$$

For any integer $p \geq 1$ and $(s,t) \in \mathbb{R}^2$, let us denote by $\Psi_{s,t}^p$ the projection of $\Psi_{s,t}$ on the first $p-th$ chaoses.

Applying Theorem 4 and Equation (2.43) of [Arc94], we get that there exists a constant $C > 0$ (which depends only on the covariance structure of the underlying Gaussian process) such that

$$\sup_{n \geq 1} \mathbb{E}\hat{\mathbb{E}}\left[\left|\frac{1}{\sqrt{n}} \sum_{k=1}^n (\Psi_{s,t} - \Psi_{s,t}^p)(X_k, \hat{X}_k)\right|^2\right] \leq C \times \int_{\mathbb{R}^2} |(\Psi_{s,t} - \Psi_{s,t}^p)(x)|^2 \gamma_2(dx).$$

Since $\Psi_{s,t}$ belongs to $L^2(\mathbb{R}^2, \gamma_2)$ for all (s,t), the last term on the right hand side goes to zero as p goes to infinity. As a result, uniformly in $n \geq 1$, the two-dimensional process

$$\left({}^\sharp F_n, {}^\sharp G_n\right) = \frac{1}{\sqrt{n}} \sum_{k=1}^n \left(\Psi_{1,0}(X_k, \hat{X}_k), \Psi_{0,1}(X_k, \hat{X}_k)\right)$$

can be approximated arbitrarily closely in $L^2(\mathbb{P} \otimes \hat{\mathbb{P}})$ by the following process which is finitely expanded on the Wiener chaoses

$$Z_n^p := (Z_n^{p,1}, Z_n^{p,2}) := \frac{1}{\sqrt{n}} \sum_{k=1}^n \left(\Psi_{1,0}^p(X_k, \hat{X}_k), \Psi_{0,1}^p(X_k, \hat{X}_k)\right).$$

Therefore, choosing $p \geq 1$ large enough, uniformly in $n \geq 1$, the product ${}^\sharp F_n \times {}^\sharp G_n$ can be approximated arbitrarily closely in $L^1(\mathbb{P} \otimes \hat{\mathbb{P}})$ by $\Delta_n^p := Z_n^{p,1} \times Z_n^{p,2}$. In other words, for any $\varepsilon > 0$ and $p \geq 1$ large enough, we have

$$\sup_n \mathbb{E}\left[\left|\hat{\mathbb{E}}\left({}^\sharp F_n \times {}^\sharp G_n\right) - \hat{\mathbb{E}}\left(\Delta_n^p\right)\right|\right] \leq \sup_n \mathbb{E}\hat{\mathbb{E}}\left[\left|{}^\sharp F_n \times {}^\sharp G_n - \Delta_n^p\right|\right] < \varepsilon.$$

But mimicking the proof detailed in the previous Section 2.3 for the convergence in probability of $\Gamma[F_n, G_n]$ towards the constant variable ν, one would then similarly get here that $\hat{\mathbb{E}}[\Delta_n^p]$ converges in probability under $\mathbb{P}$ towards a constant random variable $\nu_p \in \mathbb{R}$, and by construction $\lim_{p\to+\infty} \nu_p = \nu$. The crucial point here is that both random variables Δ_n^p and $\hat{\mathbb{E}}[\Delta_n^p]$ are now finitely expanded on the Wiener chaoses under $\mathbb{P} \otimes \hat{\mathbb{P}}$ and $\mathbb{P}$ respectively. Therefore, by hypercontractivity, the convergence in probability can be freely upgraded to the convergence in L^q for every $q \geq 1$. In particular, as n goes to infinity, the sequence $\hat{\mathbb{E}}[\Delta_n^p]$ converges in L^1 to the constant variable ν_p.

2.5 Conclusion

We go back to Stein's Equation. Let $\phi \in \mathcal{C}_b^1(\mathbb{R})$ and $\varepsilon > 0$. Integrating by parts, for $p \geq 1$ large enough and by the results of the last section, we have

$$\begin{aligned}
&|\mathbb{E}[F_n\phi(F_n)] - \nu\,\mathbb{E}[\phi'(F_n)]| = \left|\mathbb{E}\left[\phi'(F_n)\Gamma[F_n, -\mathcal{L}^{-1}F_n]\right] - \nu\,\mathbb{E}[\phi'(F_n)]\right| \\
&= |\mathbb{E}[\phi'(F_n)\Gamma[F_n, G_n]] - \nu\,\mathbb{E}[\phi'(F_n)]| \\
&= \left|\mathbb{E}\left[\phi'(F_n)\left(\Gamma[F_n, G_n] - \hat{\mathbb{E}}[\Delta_n^p]\right)\right] + \mathbb{E}\left[\phi'(F_n)\left(\hat{\mathbb{E}}[\Delta_n^p] - \nu_p\right)\right] + (\nu_p - \nu)\mathbb{E}[\phi'(F_n)]\right| \\
&\leq ||\phi'||_\infty\varepsilon + ||\phi'||_\infty\mathbb{E}\left[\left|\hat{\mathbb{E}}[\Delta_n^p] - \nu_p\right|\right] + ||\phi'||_\infty|\nu_p - \nu|.
\end{aligned}$$

As a result, letting first n and then p go to infinity, we get that uniformly in ϕ such that $||\phi'||_\infty \leq C$

$$\limsup_{n\to+\infty} |\mathbb{E}[F_n\phi(F_n)] - \nu\,\mathbb{E}[\phi'(F_n)]| = 0.$$

One can then classically conclude using Stein's approach for the convergence in total variation.

References

[AP20] Jürgen Angst and Guillaume Poly. On the absolute continuity of random nodal volumes. *The Annals of Probability*, 48(5):2145–2175, 2020. MR4152638

[Arc94] Miguel A. Arcones. Limit theorems for nonlinear functionals of a stationary gaussian sequence of vectors. *The Annals of Probability*, 22(4):2242–2274, 1994. MR1331224

[BM83] Péter Breuer and Péter Major. Central limit theorems for non-linear functionals of gaussian fields. *Journal of Multivariate Analysis*, 13(3):425–441, 1983. MR0716933

[KN19] Sefika Kuzgun and David Nualart. Rate of convergence in the Breuer–Major theorem via chaos expansions. *Stochastic Analysis and Applications*, 37(6):1057–1091, 2019. MR4020062

[NNP21] Ivan Nourdin, David Nualart, and Giovanni Peccati. The Breuer–Major theorem in total variation: Improved rates under minimal regularity. *Stochastic Processes and their Applications*, 131:1–20, 2021. MR4151212

[NP12] Ivan Nourdin and Giovanni Peccati. *Normal approximations with Malliavin calculus*, volume 192 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2012. From Stein's method to universality. MR2962301

[NPY19] Ivan Nourdin, Giovanni Peccati, and Xiaochuan Yang. Berry–Esseen bounds in the Breuer–Major CLT and Gebelein's inequality. *Electronic Communications in Probability*, 24(none):1–12, 2019. MR3978683

[Nua09] David Nualart. *Malliavin calculus and its applications*, volume 110 of *CBMS Regional Conference Series in Mathematics*. Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2009. MR2498953

[NZ21] David Nualart and Hongjuan Zhou. Total variation estimates in the Breuer–Major theorem. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 57(2):740–777, 2021. MR4260482