Probabilistic Approximations

University of Helsinki
Department of mathematics and statistics
Spring 2015

The aim of this mini-course is to provide an introduction to the combination of two probabilistic techniques. First the Stein's method (1972). This is a collection of probabilistic techniques which allow to compare probability distributions by means of the properties of differential operators (for more information, see [7]). Second the Malliavin calculus (1973). It's an infinite dimensional differential calculus (for a detailed text, see the book [15]). Interestingly, the aforementioned techniques can be sweetly combined in order to provide CLTs for non-linear functionals of an infinite dimensional *isonormal Gaussian process*. As a substantial result, we will proof an astonishing discovery (this is of one the main objectives of the course) by Nualart-Peccati (2005)[1] known nowadays as the *fourth moment theorem*, stating that, for a sequence F_n of random variables living in a fixed *Wiener chaos* such that $\mathbb{E}(F_n^2) \to 1$, the sequence F_n converges in distribution towards a standard Gaussian distribution if and only if $\mathbb{E}(F_n^4) \to 3 (= \mathbb{E}(N^4)$, where $N \sim \mathscr{N}(0,1))$. This new and efficient methodology, i.e. combining the Malliavin calculus together with the Stein's method, in literatures, is called the *Malliavin-Stein* approach. For an exposition of this fertile line of research, one can consult the following constantly updated webpage:

`https://sites.google.com/site/malliavinstein/home`

for many applications of Malliavin-Stein approach, as well as for asymptotic results that are somehow connected with the fourth moment theorem. Moreover, the monograph [13] provides a quite detailed introduction to the topics that will be discussed in the course.
The plan of the course is the following. Lecture 1 : Stein's method, Gaussian measure, stochastic integration and chaotic decompositions, Malliavin calculus. Lecture 2 : combination of the Stein's method with the Malliavin calculus and CLTs on the Wiener chaos. Lecture 3 : applications, new directions (powerful Markov triplet approach [1, 2]) and generalizations (non-Gaussian target distributions [14, 3]) as well as some important open problems if time permits.

Part I
Gaussian approximation

1 Introduction

Typical example. Take $W = \{W_t, t \geq 0\}$ a standard BM started from zero. This means that W is a centered Gaussian process such that $W_0 = 0$, W has continuous paths, and $\mathbb{E}(W_s W_t) = s \wedge t$ for every $t, s \geq 0$. A result by Jeulin (1979) says:

$$\int_0^1 \frac{W_t^2}{t^2} dt = \infty \quad \text{a.s.} \tag{1}$$

(Note that this is a property at around 0). Also, notice that for all $\varepsilon > 0$ we have

$$B_\varepsilon = \int_\varepsilon^t \frac{W_t^2}{t^2} dt < \infty$$

Remark 1. Define a new process $\hat{W}$ by $\hat{W}_0 = 0$ and $\hat{W}_u = u W_{1/u}$ for $u > 0$. It can be easily shown that $\hat{W}$ is a standard Brownian motion, and using the change of variable $u = 1/t$, it now follows that

[1] Nualart, D., Peccati, G. (2005) Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* Volume 33, Number 1, 177-193.

the property (1) is equivalent to the following statement:

$$\int_1^\infty \frac{W_t^2}{t^2} dt = \infty \qquad \text{a.s.}$$

By direct computations, one can show that (check it!)

$$\mathbb{E}(B_\varepsilon) = -\log\varepsilon, \quad \mathrm{Var}B_\varepsilon \approx \sqrt{-4\log\varepsilon}, \quad \text{as}\, \varepsilon \to 0.$$

By setting

$$\tilde{B}_\varepsilon = \frac{B_\varepsilon + \log\varepsilon}{\sqrt{-4\log\varepsilon}}, \qquad \varepsilon \in (0,1)$$

one can ask the following natural question:

Problem 1. *Prove that, as $\varepsilon \to 0$, we have $\tilde{B}_\varepsilon \overset{law}{\to} \mathscr{N}(0,1)$.*

Later on, we will present two different solutions to the above problem. One, using the classical *method of moments/cumulant*, and second, using the techniques introduced in this course. It will turn out that using the second approach, we are not only able to give a fruitful solution to the above problem but also we can provide the following quantitative bound: there exist constants C_1 and C_2 such that

$$C_1(\sqrt{-\log\varepsilon})^{-1} \le d_{\mathrm{Kol}}(\tilde{B}_\varepsilon, \mathscr{N}(0,1)) \le C_2(\sqrt{-\log\varepsilon})^{-1}.$$

2 Elements of Stein's method

The typical route is the following (a) Stein's lemma, then (b) develop a heuristic, followed by (c) an equation whose solutions (and properties thereof) will lead to bounds.

2.1 Moments/Cumulants

During the lectures, the notion of *cumulant* is sometimes used. Recall that, given a random variable Y with finite moments of all orders, i.e. $\mathbb{E}|Y|^r < \infty$ for all $r \ge 1$, and with characteristic function $\varphi_Y(t) := \mathbb{E}(e^{itY})$, $t \in \mathbb{R}$, one define the sequence of cumulants of Y, noted as $\{\kappa_r(Y) : r \ge 1\}$, as

$$\kappa_r(Y) = (-i)^r \frac{d^r}{dt^r} \log\varphi_Y(t)\Big|_{t=0}, \qquad r \ge 1.$$

For instance,

$$\begin{aligned}
\kappa_1(Y) &= \mathbb{E}(Y)\\
\kappa_2(Y) &= \mathrm{Var}(Y)\\
\kappa_3(Y) &= \mathbb{E}(Y^3) - 3\mathbb{E}(Y^2)\mathbb{E}(Y) + 2\mathbb{E}(Y)^3\\
\kappa_4(Y) &= \mathbb{E}(Y^4) - 4\mathbb{E}(Y)\mathbb{E}(Y^3) - 3\mathbb{E}(Y^2)^2 + 12\mathbb{E}(Y)^2\mathbb{E}(Y^2) - 6\mathbb{E}(Y)^4.
\end{aligned}$$

In particular, if $\mathbb{E}(Y) = 0$, then $\kappa_3(Y) = \mathbb{E}(Y^3)$ and $\kappa_4(Y) = \mathbb{E}(Y^4) - 3\mathbb{E}(Y^2)^2$. Recall that for a standard Gaussian random variable $N \sim \mathscr{N}(0,1)$, we have $\log\varphi_N(t) = -t^2/2$, and therefore $\kappa_1(N) = \mathbb{E}(N) = 0$, $\kappa_2(N) = \mathrm{Var}(N) = 1$, and $\kappa_r(N) = 0$ **for all** $r \ge 3$.

Remark 2. The following relation shows that moments can be recursively defined in terms of cumulants (and vice-versa): fix $r = 1, 2, \cdots$ and assume that $\mathbb{E}|Y|^{r+1} < \infty$, then

$$\mathbb{E}(Y^{r+1}) = \sum_{s=0}^{r} \binom{s}{r} \kappa_{s+1}(Y)\mathbb{E}(Y^{r-s}). \tag{2}$$

The reader is referred to [17, Chapter 3] for a proof of relation (2), as well as, for a self-contained presentation of more properties of cumulants and for several combinatorial characterizations.

Exercise 1. Let $N \sim \mathscr{N}(0,1)$. (a) Show that the moments sequence $\{m_r(N) := \mathbb{E}(N^r) : r \geq 1\}$ of N satisfies in the following recursion formula

$$m_{r+1}(N) = r m_{r-1}(N), \quad r \geq 1. \tag{3}$$

(b) Using induction and part (a) to prove that

$$m_r(N) = \begin{cases} (2k-1)!! & \text{if } r = 2k \\ 0 & \text{otherwise.} \end{cases}$$

where the notation double factorial $(2k-1)!! = (2k-1) \times (2k-3) \times \cdots \times 3 \times 1$.

The following lemma is a fundamental key to provide CLTs using the method of moments/cumulants.

Lemma 1. *The law of the random variable $N \sim \mathscr{N}(0,1)$ is determined by its moments/cumulants, i.e if X be a random variable such that $\mathbb{E}(X^r) = \mathbb{E}(N^r)$ [or equivalently $\kappa_r(X) = \kappa_r(N)$] for all $r \geq 1$, then $X \overset{law}{=} N$.*

Proof. Let $\text{law}(N) = \gamma$ and $\text{law}(X) = \nu$. Then, it is enough to show that their Fourier transforms are the same: $\int_{\mathbb{R}} e^{itx}\gamma(dx) = \int_{\mathbb{R}} e^{itx}\nu(dx)$, for every $t \in \mathbb{R}$. Since $m_r(N) = m_r(X)$ for all $r \geq 1$, using Taylor's formula, triangle inequality, the following elementary inequality

$$\Big| e^{itx} - \sum_{k=0}^{r} \frac{(itx)^k}{k!} \Big| \leq \frac{|tx|^{r+1}}{(r+1)!}$$

and Cauchy-Schwarz inequality to write

$$\begin{aligned} \Big| \int_{\mathbb{R}} e^{itx}\gamma(dx) - \int_{\mathbb{R}} e^{itx}\nu(dx) \Big| &\leq \int_{\mathbb{R}} \Big| e^{itx} - \sum_{k=0}^{r} \frac{(itx)^k}{k!} \Big| \gamma(dx) \\ &+ \int_{\mathbb{R}} \Big| e^{itx} - \sum_{k=0}^{r} \frac{(itx)^k}{k!} \Big| \nu(dx) \\ &\leq \Big(\int_{\mathbb{R}} \frac{|tx|^{2r+2}}{(r+1)!^2} \gamma(dx) \Big)^{\frac{1}{2}} \\ &+ \Big(\int_{\mathbb{R}} \frac{|tx|^{2r+2}}{(r+1)!^2} \nu(dx) \Big)^{\frac{1}{2}} \\ &= 2\sqrt{\frac{|t|^{2r+2} m_{2r+2}(N)}{(r+1)!^2}}, \end{aligned}$$

for every $r \geq 1$. Now, using Stirling formula $r! \sim \sqrt{2\pi r}(\frac{r}{e})^r$ as $r \to \infty$, and Lemma 2.1, one can infer that

$$\lim_{r \to \infty} \frac{|t|^{2r+2} m_{2r+2}(N)}{(r+1)!^2} = 0.$$

□

b) The following lemma known as *Stein's lemma* provides a useful characterization of one-dimensional standard Gaussian distributions.

Lemma 2. (Stein's lemma) *For a real-valued random variable Y we have $Y \sim \mathscr{N}(0,1)$ if, and only if for every $f : \mathbb{R} \to \mathbb{R}$ such that $E|f'(N)| < \infty$, we have*

$$\mathbb{E}(f'(Y) - Y f(Y)) = 0. \tag{4}$$

Proof. The sufficient condition is trivial. For the other way, note that for all polynomials the relation (4) works. But this means that

$$\mathbb{E}(Y^{r+1}) = r\mathbb{E}(Y^{r-1}).$$

Now, use Exercise 2.1 and Lemma 1. Another way is to take f complex exponential and therefore determine the characteristic function of Y (do it!). □

Theorem 1. (The method of moments/cumulants) *Let F be a real-valued random variable whose law is determined by its moments/cumulants. Assume that $\{F_n\}_{n\geq 1}$ be a sequence of random variables in which each F_n has all moments/cumulants such that $\mathbb{E}(F_n^r) \to \mathbb{E}(F^r)$, for every $r \geq 1$. Then F_n converges in distribution towards F.*

b) Heuristic. Suppose that for "many" functions f we have

$$\mathbb{E}(f'(Y) - Yf(Y)) \approx 0.$$

Can we conclude that Y is close – in some sense – to N? This is, *a priori*, not clear since there are many ways to characterize N and not all lead to a nice theory of probabilistic approximation. We will consider a very strong measure of closeness in terms of the total variation (TV) distance.

2.2 Distances between probability measures

(i) *The Kolmogorov distance*: Let F and G be two $\mathbb{R}^d$, $(d \geq 1)$ valued random variables. Let

$$\mathscr{H}_{\text{Kol}} = \{h : \mathbb{R}^d \to \mathbb{R} \,:\, h(x_1, \cdots, x_d) = \Pi_{k=1}^d \mathbf{1}_{(-\infty, z_k]}(x_k), \text{ for some } z_1, \cdots, z_d \in \mathbb{R}\}.$$

The Kolmogorov distance between the laws of random variables F and G, noted as $d_{\text{Kol}}(F, G)$, define as

$$\begin{aligned} d_{\text{Kol}}(F,G) &= \sup_{h\in\mathscr{H}_{\text{Kol}}} \Big|\mathbb{E}(h(F)) - \mathbb{E}(h(G))\Big| \\ &= \sup_{z_1,\cdots,z_d\in\mathbb{R}} \Big|\mathbb{P}(F \in (-\infty, z_1] \times \cdots \times (-\infty, z_d]) \\ &\qquad - \mathbb{P}(G \in (-\infty, z_1] \times \cdots \times (-\infty, z_d])\Big|. \end{aligned}$$

In particular $(d = 1)$: $d_{\text{Kol}}(F, G) = \sup_{z\in\mathbb{R}} \Big|\mathbb{P}(F \leq z) - \mathbb{P}(G \leq z)\Big|$. Note that always $d_{\text{Kol}}(F, G) \leq d_{\text{TV}}(F, G)$.
(ii) *The total variation distance*:

$$\mathscr{H}_{\text{TV}} = \{h : \mathbb{R}^d \to \mathbb{R} \,:\, h = \mathbf{1}_B \text{ for some } B \in \mathcal{B}(\mathbb{R}^d)\}.$$

$$\begin{aligned} d_{\text{TV}}(F,G) &= \sup_{h\in\mathscr{H}_{\text{TV}}} \Big|\mathbb{E}(h(F)) - \mathbb{E}(h(G))\Big| \\ &= \sup_{B\in\mathcal{B}(\mathbb{R}^d)} \Big|\mathbb{P}(F \in B) - \mathbb{P}(G \in B)\Big|. \end{aligned}$$

(iii) *The Wasserstein distance*:

$$\mathscr{H}_{\text{W}} = \{h : \mathbb{R}^d \to \mathbb{R} : \|h\|_{\text{lip}} \leq 1\}, \quad \|h\|_{\text{lip}} := \sup_{x\neq y\in\mathbb{R}^d} \frac{|h(x) - h(y)|}{\|x - y\|_{\mathbb{R}^d}}.$$

$$d_{\text{W}}(F,G) = \sup_{h\in\mathscr{H}_{\text{W}}} \Big|\mathbb{E}(h(F)) - \mathbb{E}(h(G))\Big|.$$

Exercise 2. Let $d \geq 1$. Show that the topologies induced by three distance d_{Kol}, d_{TV} and d_{W} on the set of probability measures on $\mathbb{R}^d$ are strictly stronger than the topology of the convergence in distribution, i.e.

$$d_{\text{Kol,TV,W}}(F_n, F) \to 0 \Longrightarrow F_n \overset{\text{law}}{\to} F.$$

Remark 3. The *Fortet–Mourier distance* (or *bounded Wasserstein distance*: $d_{\text{FM}}(F,G) = \sup_{h \in \mathscr{H}_{\text{FM}}} \Big| \mathbb{E}(h(F)) - \mathbb{E}(h(G)) \Big|$, where $\mathscr{H}_{\text{FM}} = \{h : \mathbb{R}^d \to \mathbb{R} \, : \, \|h\|_\infty + \|h\|_{\text{lip}} \leq 1\}$. The d_{FM} distance metrizes the convergence in distribution, i.e.

$$d_{\text{FM}}(F_n, F) \to 0 \Longleftrightarrow F_n \overset{\text{law}}{\to} F.$$

c) Stein's equation for normal approximation. Let $N \sim \mathscr{N}(0,1)$. Consider a function $h : \mathbb{R} \to [0,1]$ so that $\mathbb{E}|h(N)| < \infty$. The Stein equation associated to the test function h is

$$f'(x) - xf(x) = h(x) - \mathbb{E}h(N) \tag{5}$$

which is taken to hold at all $x \in \mathbb{R}$. A solution is a function f_h whose derivative is a.e. defined and for which there exists a version which satisfies (5). In particular we always speak of f' in the weak sense. For a moment, assume that f_h is a solution of (5). Then, by taking expectation of both sides (5) (together with plugging in $x = Y$, where Y is a real-valued random variable):

$$\mathbb{E}h(Y) - \mathbb{E}h(N) = \mathbb{E}(f'(Y) - Yf(Y)).$$

Therefore, for any integrable random variable Y:

$$\sup_{h \in \mathscr{H}_{\text{TV}}} \Big| \mathbb{E}h(Y) - \mathbb{E}h(N) \Big| = \sup_{f_h, h \in \mathscr{H}_{\text{TV}}} \Big| \mathbb{E}(f_h'(Y) - Yf_h(Y)) \Big|. \tag{6}$$

Note that the expression in the right hand side in above **does not** involved the target random variable N at all!

Proposition 1. *For every $c \in \mathbb{R}$, set*

$$f_{c,h}(x) = ce^{x^2/2} + e^{x^2/2} \int_{-\infty}^{x} (h(u) - \mathbb{E}h(N)) e^{-u^2/2} du;$$

Then, $f_{c,h}$ is a solution of the Stein's equation. Moreover, the unique solution satisfying in $\lim_{n\to\infty} e^{-\frac{x^2}{2}} f(x) = 0$ *is given by $f_h = f_{0,h}$, i.e. $c = 0$.*

Proof. Note that, the Stein's equation can be written as

$$e^{\frac{x^2}{2}} \frac{d}{dx} \Big(e^{-\frac{x^2}{2}} f(x) \Big) = h(x) - \mathbb{E}h(N).$$

Now, take integral of both sides. For the second part, using dominated convergence theorem (DCT) we have

$$\lim_{|x| \to \infty} \int_{-\infty}^{x} (h(u) - \mathbb{E}h(N)) e^{-u^2/2} du = 0.$$

Recall that $\mathbb{E}|h(N)| < \infty$. $\square$

The gist of the method is that it will transform the study of a non-smooth object (the TV distance) in terms of smooth objects (the solutions $f_{0,h}$). This happens through the following lemma.

Lemma 3. *Let $h : \mathbb{R} \to [0,1]$. Then the solution f_h of the Stein's equation associated to h satisfying in*

$$\|f_h\|_\infty \leq \sqrt{\pi/2} \quad \text{and} \quad \|f_h'\|_\infty \leq 2$$

(the Stein's magic factors).

Note that these bounds are uniform over the whole family h. An immediate consequence of Lemma 3 is the following.

Corollary 1. *Let Y be a real-valued random variable such that $\mathbb{E}|Y| < \infty$. Then*

$$d_{TV}(Y, N) \leq \sup_{f \in \mathscr{F}_{TV}} \Big| \mathbb{E}(f'(Y)) - \mathbb{E}(Yf(Y)) \Big|,$$

where $\mathscr{F}_{TV} = \{f : \|f\|_\infty \leq \sqrt{\pi/2} \quad \text{and} \quad \|f'\|_\infty \leq 2\}$.

Let to stress that we have explicitly transformed the non-smooth problem, the lhs of (6), into a smooth one, the rhs of (6). Moreover, the bounds in Lemma 3 are independent of the target Gaussian random variable N, and just depends on (a functional of) a nice Y! Once we have this, is it true that the rhs is easier to evaluate than the lhs? This is Stein's intuition and it turns out to be true in many different and complicated cases. There are several techniques for working out this quantity : exchangeable pairs (Stein, 1972); dependency graphs; zero-bias transforms (Goldstein - Reinert 1995). Here we are going to develop tools to evaluate the Stein bound when Y is a (sufficiently regular) functional of a infinite-dimensional Gaussian field (e.g. the Brownian motion, the fractional Brownian motion,...). The answer for this is through *Malliavin calculus.*

Proof. (of Lemma 3) First remark that $|h(u) - \mathbb{E}h(N)| \leq 1$. Then we easily get (with Φ the Gaussian CDF)

$$\begin{aligned} |f_h(x)| &\leq e^{x^2/2} \min\{\Phi(x), 1 - \Phi(x)\} \\ &= e^{x^2/2} \int_{|x|}^{\infty} e^{-y^2/2} dy := S(x). \end{aligned}$$

A direct computation shows that S attains its maximum at $x = 0$, and also $S(0) = \sqrt{\pi/2}$. Hence, $|f_h| \leq S(0) = \sqrt{\pi/2}$, and the first claim follows.

For the second bound, we can simply write out f_h', to get

$$\begin{aligned} |f_h'(x)| &= \Big| h(x) - \mathbb{E}h(N) + xe^{x^2/2} \int_{-\infty}^{x} (h(u) - \mathbb{E}h(N)) e^{-u^2/2} du \Big| \\ &\leq 1 + |x| e^{x^2/2} \int_{|x|}^{\infty} e^{-y^2/2} dy = 2. \end{aligned}$$

□

Exercise 3. Stein's bound for the Kolmogorov distance (a) For every $z \in \mathbb{R}$, write $f_z = f_{\mathbf{1}_{(-\infty,z]}}$, that is, f_z is the solution of the Stein's equation associated to the indicator function $h = \mathbf{1}_{(-\infty,z]}$. Also, Φ stands for the cumulative distribution function of a $\mathscr{N}(0,1)$ random variable. Show that

$$f_z(x) = \begin{cases} \sqrt{2\pi} e^{\frac{x^2}{2}} \Phi(x)[1 - \Phi(z)] & \text{if } x \leq z, \\ \sqrt{2\pi} e^{\frac{x^2}{2}} \Phi(z)[1 - \Phi(x)] & \text{if } x \geq z. \end{cases}$$

(b) Prove that, for every $x \in \mathbb{R}$, $f_z(x) = f_{-z}(-x)$ (this implies that, in the estimates below, one can assume that $z \geq 0$ without loss of generality).

(c) Compute the derivative $\frac{d}{dx}[xf_z(x)]$, and deduce that the mapping $x \mapsto xf_z(x)$ is increasing.

(d) Show that $\lim_{x \to -\infty} xf_z(x) = \Phi(z) - 1$ and also that $\lim_{x \to +\infty} xf_z(x) = \Phi(z)$.

(e) Use part (a) to prove that

$$f_z'(x) = \begin{cases} [\sqrt{2\pi} x e^{\frac{x^2}{2}} \Phi(x) + 1][1 - \Phi(z)] & \text{if } x < z, \\ [\sqrt{2\pi} x e^{\frac{x^2}{2}} (1 - \Phi(x)) - 1]\Phi(z) & \text{if } x > z. \end{cases}$$

(f) Use part (e) in order to prove that

$$0 < f_z'(x) \leq zf_z(x) + 1 - \Phi(z) < 1, \quad \text{if } x < z,$$

and

$$-1 < zf_z(x) - \Phi(z) \leq f_z'(x) < 0, \quad \text{if}\, x > z$$

to deduce that $\|f_z'\|_\infty \leq 1$.
(g) Use part (f) to show that $x \mapsto f_z(x)$ attains its maximum in $x = z$. Compute $f_z(z)$ and prove that $f_z(z) \leq \frac{\sqrt{2\pi}}{4}$ for every $z \in \mathbb{R}$, to complete a proof of the following theorem.

Theorem 2. *Let $z \in \mathbb{R}$. Then the function f_z is such that $\|f_z\|_\infty \leq \frac{\sqrt{2\pi}}{4}$ and $\|f_z'\|_\infty \leq 1$. Therefore, for $N \sim \mathscr{N}(0,1)$, and for any integrable random variable F,*

$$d_{Kol}(F, N) \leq \sup_{f \in \mathscr{F}_{Kol}} |\mathbb{E}[f'(F)] - \mathbb{E}[Ff(F)]|,$$

where $\mathscr{F}_{Kol} = \{f : \|f\|_\infty \leq \frac{\sqrt{2\pi}}{4}, \|f'\|_\infty \leq 1\}$.

Part II
Gaussian measures and chaos

Take $(\Omega, \mathcal{F}, \mathbb{P})$ an underlying probability space.

2.3 Definition and first properties

We first define Gaussian measures.

Definition 1. *Take $(A, \mathcal{A}, \mu)$ a measure space (Polish, i.e. metric, separable and complete) with μ positive, σ-finite and non-atomic measure. A Gaussian measure over $(A, \mathcal{A})$ with control μ is a centered Gaussian family*

$$G = \{G(B); \quad \mu(B) < \infty\}$$

such that

$$\mathbb{E}\left(G(B)G(C)\right) = \mu(B \cap C), \quad \mu(B) < \infty \,and\, \mu(C) < \infty.$$

A couple of remarks : (i) If $A = \mathbb{R}^+$ and μ is the Lebesgue measure then $W_t = G[0,t]$ is, up to continuity, a Brownian motion (because then $\mathbb{E}W_tW_s = \min(t,s)$); (ii) one can prove that if $\{B_i\}_{i \geq 1}$ is a sequence of **disjoint** sets such that $\mu(\bigcup_i B_i) < \infty$ then

$$G(\bigcup_i B_i) = \sum_i G(B_i)$$

with convergence in $L^2(\Omega)$.

Proposition 2. *G, in fact, exists.*

Proof. Take $\{e_i\}_{i \geq 1}$ an ONB of $L^2(\mu)$. Then for all $f \in L^2(\mu)$ we have $f = \sum \langle f, e_i \rangle e_i$ with the $L^2(\mu)$ scalar product. Next take $\{\xi_i\}_{i \geq 1}$ a sequence of i.i.d. $\mathscr{N}(0,1)$ random variables and construct

$$X(f) = \sum_{ige1} \xi_i \langle f, e_i \rangle.$$

Then $\{X(f); f \in L^2(\mu)\}$ is a centered Gaussian family such that $\mathbb{E}(X(f)X(g)) = \langle f, g \rangle$ (easy exercise through Parseval's identity, check it!). We are then ready to conclude, since

$$G(B) = X(\mathbf{1}_B), \quad \mu(B) < \infty$$

is a Gaussian measure with control measure μ. □

A final remark is that GM are not probability measures! More precisely:

Proposition 3. *The mapping*

$$B \mapsto G(B)(\omega)$$

is not a signed measure for a fixed ω.

Proof. Take a Borel set B with $\mu(B) < \infty$. Since μ is non-atomic, we observe that

$$\int_A \int_A \mathbf{1}_{B\times B}(x,y)\mathbf{1}_{x=y}\mu(dx)\mu(dy) = \text{Diag}_\mu(B) = 0.$$

But, on the other hand side, one can easily show that

$$\int_A \int_A \mathbf{1}_{B\times B}(x,y)\mathbf{1}_{x=y}G(dx)G(dy) := \text{Diag}_G(B) = \mu(B).$$

(Note that the integration wrt G is shaky but will be proven rigorously later on). Indeed here a standard way to construct $\text{Diag}_G(B)$ is through

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} G(B_i^{(n)})G(B_i^{(n)}) = \lim_n \sum G(B_i^{(n)})^2$$

where $\{B_i^{(n)}, i = 1, \ldots, k_n\}$ is a sequence of partitions of B such that $\sup \mu(B_i^{(n)}) \to 0$, and one can show that (check it!), for any partition,

$$\mathbb{E}\big(\sum G(B_i^{(n)})^2 - \mu(B)\big)^2 \to 0.$$

In other words G charges, in a nontrivial way, the diagonal and hence cannot be a signed measure. □

Remark 4. In the case, when $A = \mathbb{R}^+$, μ is the Lebesgue measure, the statement $B \mapsto G(B)(\omega)$ is a singed measure on a set of positive probability will imply that the mapping $t \mapsto W_t := G[0,t]$ is of bounded variation on a set of positive probability in which is a contradiction with the following well-known fact that

$$\sum_{0\le t_i \le t} (W_{t_i^{(n)}} - W_{t_{i-1}^{(n)}})^2 \to t.$$

2.4 Single integrals

We want for any $f \in L^2(\mu)$ to define an object of the type

$$I_1(f) := \int_A f(x)G(dx)$$

To this end, we introduce a collection of "simple integrands"

$$\mathcal{E}(\mu) = \{f(x) = \sum_{j=1}^N c_j \mathbf{1}_{B_j},\, \mu(B_j) < \infty\}$$

which has the density property $\bar{\mathcal{E}}(\mu) = L^2(\mu)$. Then we can define, for any simple integrand f

$$I_1(f) := \sum_{j=1}^N c_j G(B_j).$$

With this in hand it is easy to show that, for all simple f, g we have

$$\mathbb{E}(I_1(f)I_1(g)) = \langle f, g\rangle_{L^2(\mu)}, \quad \mathbb{E}I_1(f) = 0.$$

Now, it is straightforward to extend to all functions $f \in L^2(\mu)$. Let $f \in L^2(\mu)$, then there exists a sequence of simple functions $\{f_n\}_{n\geq 1}$ such that $\|f_n - f\| \to 0$ and $\{I_1(f_n)\}$ is Cauchy in $L^2(\mathbb{P})$. One therefore sets

$$I_1(f) = \lim I_1(f_n)$$

the limit being taken in $L^2(\mathbb{P})$ and being independent of the choice of the sequence.

Remark 5. For all $f, g \in L^2(\mu)$ we have

$$\mathbb{E} I_1(f) = 0 \quad \text{and} \quad \mathbb{E}\left(I_1(f) I_1(g)\right) = \langle f, g\rangle_{L^2(\mu)}.$$

Definition 2 (Wiener, 1938)**.** The space $\mathcal{H}_1 = \{I_1(f) : f \in L^2(\mu)\}$ is the *first Wiener chaos of* G. Note that, since everything is obtained through centered Gaussian random variables, so $\mathcal{H}_1$ is a centered Gaussian family. It therefore cannot be suffice to describe all square integrable functionals measurable wrt G.

2.5 Multiple integrals

Let $p \geq 2$. Consider $L^2(\mu^p)$ the space of square integrable functions of p arguments on the space $(A^p, \mathcal{A}^p, \mu^p)$. We then define the simple functions

$$\begin{aligned}\mathcal{E}(\mu^p) =& \text{ Simple integrands} \\ =& \left\{ f = \sum_{i_1,\cdots,i_p=1}^{n} a_{i_1\cdots i_p} \mathbf{1}_{B_{i_1}} \otimes \cdots \otimes \mathbf{1}_{B_{i_p}} : B_{i_k} \cap B_{i_l} = \emptyset \,\forall\, k \neq l \text{ and } \mu(B_{i_j}) < \infty \right\},\end{aligned}$$

and **more importantly** the coefficients $a_{i_1\cdots i_p} = 0$ if any of two indices $i_1, \cdots, i_p$ are equal, i.e. $\exists\, k \neq l$ such that $i_k = i_l$.

For any $f \in \mathcal{E}(\mu^p)$ then we define the *multiple Wiener-Itô integral of order* p *of* $f \in \mathcal{E}(\mu^p)$ *w.r.t.* G through

$$I_p(f) = \sum_{i_1,\cdots,i_p=1}^{n} a_{i_1\cdots i_p} G(B_1)\ldots G(B_d). \tag{7}$$

Now, the main properties are gathered in the following exercise.

Exercise 4. (a) Show that for $f \in \mathcal{E}(\mu^p)$, the definition of $I_p(f)$ does not depend on a particular representation of f.
(b) $I_p : \mathcal{E}(\mu^p) \to L^2(\Omega, \mathbb{P})$ is a linear map.
(c) $I_p(f) = I_p(\tilde{f})$, where $\tilde{f}$ is the *symmetrization* of f, i.e.

$$\tilde{f}(x_1, \cdots, x_p) = 1/p! \sum_{\sigma} f(x_{\sigma(1)}, \cdots, x_{\sigma(p)})$$

where the sum runs over all permutations σ of $\{1, \cdots, p\}$.

(d) For all $f \in \mathcal{E}(\mu^p)$ and $g \in \mathcal{E}(\mu^q)$ we have

$$\mathbb{E}(I_p(f)) = 0 \tag{centered}$$

and

$$\mathbb{E}(I_p(f) I_q(g)) = \begin{cases} 0 & \text{if}\, p \neq q, \\ p!\langle \tilde{f}, \tilde{g}\rangle_{L^2(\mu^p)} \end{cases} \tag{orthogonality / isometry}$$

(e) $\overline{\mathcal{E}_{\text{sym}}(\mu^p)}^{\|\,\|_{L^2(\mu^p)}} = L^2_{\text{sym}}(\mu^p)$, where here "sym" stands for symmetrized functions. Hence, deduce that the mapping I_p can be continuously extended to $L^2_{\text{sym}}(\mu^p)$.

Now let $f \in L^2_{\text{sym}}(\mu^p)$; then there exists a sequence $\{f_n\} \subset \mathcal{E}_{\text{sym}}(\mu^p)$ such that $f_n \to f$ in $L^2(\mu^p)$. Therefore, we define

$$I_p(f) := \lim_{n\to\infty} I_p(f_n) \tag{8}$$

the limit being taken in $L^2(\mathbb{P})$ and being independent of the chosen sequence.

Definition 3. *The* Wiener chaos of order p associated with G*, denoted by* $\mathcal{H}_p$*, is defined as*

$$\mathcal{H}_p := v.s.\Big\{I_p(f)\,; \quad f \in L^2_{sym}(\mu^p)\Big\}.$$

Moreover, the three properties (centered, isometry and orthogonality) extend to the whole class $\mathcal{H}_p$*. Note that we write* $I_0(c) = c; c \in \mathbb{R}$

Remark 6. *If* $A = [0,T]$*,* μ *is Lebesgue and* $W_t = G([0,t])$ *is a Brownian motion, then for symmetric* $f \in L^2([0,T]^p)$

$$I_p(f) = p! \int_0^T dW_{t_1} \int_0^{t_1} dW_{t_2} \ldots \int_0^{t_{p-1}} dW_{t_p} f(t_1,\ldots,t_p).$$

The random variables in Definition are fundamental: they allow for example to write any G-square integrable random variable as an infinite series (this will be treated later on). The following remark provides some crucial properties of random variables living in a Wiener chaos.

Remark 7. (a) Shigekawa [2] proves that if $F = \sum_{p=0}^M I_p(f_p)$ then the law of F has a density w.r.t. the Lebesgue measure.
(b) (Nelson, 1968) $\mathcal{H}_p$ is *hypercontractive*, i.e. $\forall q > 0$ there exists $C_{p,q} > 0$ such that for all $F \in \mathcal{H}_p$ we have

$$\mathbb{E}(|F|^q)^{1/q} \le C_{p,q}\, \mathbb{E}(F^2)^{1/2}. \tag{9}$$

In particular all these L^p topologies are equivalent on the Wiener chaoses.

2.6 Multiplication formulae & Chaotic expansion

The problem. What is $I_p(f) \times I_q(g)$?

Definition 4 (Contraction)**.** *For* $f \in L^2_{sym}(\mu^p)$ *and* $g \in L^2_{sym}(\mu^q)$ *for* $p,q \ge 1$ *we define for all* $r = 0,\ldots,\min(p,q)$

$$\begin{aligned} &f \otimes g(x_1, x_2, \ldots, , x_{p+q-2r}) \\ &= \int_{A^r} f(\underline{a_r}, x_1, \ldots, x_{p-r}) g(\underline{a_r}, x_{p-r+1}, \ldots, x_{p+q-2r}) \mu(da_1, \cdots, da_r). \end{aligned}$$

Here $\underline{a_r} = (a_1, \cdots, a_r)$.

Example 1. *If* $p = q = r$ *then* $f \otimes_p g = \langle f, g\rangle_{L^2(\mu^p)}$.

Example 2. *If* $r = 0$ *then*

$$f \otimes_0 g(x_1,\ldots,x_{p+q}) = f \otimes g = f(x_1,\ldots,x_p) g(x_{p+1},\ldots,x_{p+q}).$$

Example 3. *If* $p = q = 2$ *and* $r = 1$ *then*

$$f \otimes_1 g(x,y) = \int_A \mu(da) f(a,x) g(a,y).$$

[2] Shigekawa, I. *Derivatives of Wiener functionals and absolute continuity of induced measures.* J. Math. Kyoto Univ. 20 (1980), no. 2, 263-289.

Remark 8.

$$\|f \otimes_r g\|^2_{L^2(\mu^{p+q-2r})} \leq \|f\|^2_{L^2(\mu^p)} \|g\|^2_{L^2(\mu^q)} < \infty.$$

Note that if $p = q = r$ this is just the CS inequality.

Remark 9. *In general $f \otimes_r q$ is not symmetric, so we define the symmetrization*

$$\widetilde{f \otimes_r g}(x_1, \ldots, x_{p+q-2r}) = \sum_{\sigma \in \mathcal{S}_{p+q-2r}} \frac{f \otimes_r g(x_{\sigma(1)}, \ldots, x_{\sigma(p+q-2r)})}{(p+q-2r)!}.$$

Note that, in general,

$$\|\tilde{f}\|_{L^2(\mu^p)} \leq \|f\|_{L^2(\mu^p)},$$

i.e. symmetrization shrinks.

We are now ready to state a fundamental result.

Theorem 3 (Multiplication formulae)**.** *Take $f \in L^2_{sym}(\mu^p)$ and $g \in L^2_{sym}(\mu^q)$. Then*

$$I_p(f) \times I_q(g) = \sum_{r=0}^{\min(p,q)} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}\left(\widetilde{f \otimes_r g}\right).$$

We will proceed to a heuristic proof of this result. For a detailed proof consult [15].

Proof. First note how we have, at each step, been very careful to avoid the diagonals. Hence by our construction $I_p(f)$ can be seen as

$$I_p(f) = \int_A \cdots \int_A f(x_1, \ldots, x_p) \mathbf{1}_{\{x_i \neq x_j, i \neq j\}} G(dx_1) \ldots G(dx_p).$$

Then we have (as through Fubini)

$$\begin{aligned} &I_p(f) I_q(g) \\ &= \int_{A^{p+q}} \underbrace{f(x_1, \ldots, x_p)}_{\text{no diag}} \underbrace{g(y_1, \ldots, y_q)}_{\text{no diag}} G(dx_1) \ldots G(dx_p) G(dy_1) \ldots G(dy_p). \end{aligned}$$

While there are no diagonals in the first and second blocks, there are all possible mixed diagonals in the joint writing. Hence we need to take into account all these diagonals (whence the combinatorial coefficients in the statement, which count all possible diagonal sets of size r) and then "integrate out", in other words we get

$$\begin{aligned} &I_p(f) I_q(g) \\ &= \sum_{r=0}^{\min(p,q)} r! \binom{p}{r} \binom{q}{r} \int_{A^{p-r}} \int_{A^{q-r}} \zeta G(dx_1) \ldots G(dx_{p-r}) G(dy_1) \ldots G(dy_{q-r}), \end{aligned}$$

with

$$\zeta = \left(\int_{A^r} f(\underline{a_r}, x_1, \ldots, x_{p-r}) g(\underline{a_r}, x_{p-r+1}, \ldots, x_{p+q-2r}) \mu^r(d\underline{a_r}) \right).$$

Since $\mathrm{Diag}_G(da) = \mu(da)$, we get the proof.

□

2.7 Hermite polynomials and chaos

Definition 5. *We define the Hermite polynomials as the family of polynomials* $\{H_n; n \geq 0\}$ *such that* $H_0 \equiv 1$ *and, for all* $n \geq 1$*,*

$$H_n(x) = (-1)^n e^{x^2/2} d_x^n (e^{-x^2/2}).$$

The following exercise gather the important properties of the Hermite polynomials.

Exercise 5. Define the **divergence operator** δ on the space $\text{Dom}(\delta) \subset L^2(\mathbb{R}, \gamma)$ as $\delta\varphi(x) := -\varphi'(x) + x\varphi = -e^{\frac{x^2}{2}} \frac{d}{dx}(e^{-\frac{x^2}{2}} \varphi(x))$. Let $p \geq 0$ be an integer. We define the pth **Hermite Polynomial** as $H_0 = 1$ and $H_p = \delta^p 1$, where here $\delta^p = \delta \circ \cdots \circ \delta$, p times.
(a) Show that $d\delta - \delta d =$ Identity, where $d = \frac{d}{dx}$, and moreover $\delta H_p = H_{p+1}$, $dH_p = pH_{p-1}$ and $(\delta + d)H_p = xH_p$.
(b) for any $p, q \geq 0$ show that

$$\int_R H_p(x) H_q(x) \gamma(dx) = \delta_{p,q} p!,$$

where here $\delta_{p,q}$ stands for the Kronecker delta.
(c) Show that the family $\{\frac{1}{\sqrt{p!}} H_p : p \geq 0\}$ is an orthonormal basis of $L^2(\mathbb{R}, \gamma)$.
(d) Define the **Ornstein-Uhlenbeck generator** $L\varphi(x) = -x\varphi'(x) + \varphi''(x)$. Show that $LH_p = -pH_p$.

In other words the multiple integrals are infinite dimensional versions of the Hermite polynomials.

Proposition 4. *For all* $h \in L^2(\mu)$ *such that* $\|h\|_{L^2(\mu)} = 1$ *we define*

$$h^{\otimes p}(x_1, \ldots, x_p) = \prod_{i=1}^{p} h(x_i) \in L^2_{sym}(\mu^p).$$

Then, for all $p \geq 1$*, we have*

$$I_p(h^{\otimes p}) = H_p(I_1(h)).$$

This is sometimes called the Wick product of order p *of* $I_1(h)$*.*

Proof. Trivial for $p = 1$. Proceed by induction and choose $p \geq 1$. Then note that

$$I_p(h^{\otimes p}) I_1(h) = I_{p+1}(h^{\otimes p+1}) + pI_{p-1}(h^{\otimes p})$$

Whence, using the recursion,

$$I_{p+1}(h^{\otimes p+1}) = H_p(I_1(h)) I_1(h) - pH_{p-1}(I_1(h))$$

Using the previous exercise we also know that

$$d_x H_p(x) = pH_{p-1}(x)$$

and thus

$$\begin{aligned} I_{p+1}(h^{\otimes p+1}) &= H_p(I_1(h)) I_1(h) - H_p(I_1(h)) \\ &= \delta H_p(I_1(h)) \\ &= H_{p+1}(I_1(h)). \end{aligned}$$

□

Theorem 4. *[Chaotic representation] For all* $F \in L^2(\sigma(G))$ *there exists a unique* $\{f_q; q \geq 1\}$ *such that* $f_q \in L^2_{sym}(\mu^q)$ *and we have*

$$F = \mathbb{E}(F) + \sum_{q=1}^{\infty} I_q(f_q) \tag{10}$$

(the equality is in $L^2(\mathbb{P})$*). In particular*

$$\mathbb{E}(F^2) = \mathbb{E}(F)^2 + \sum_{q=1}^{\infty} q! \|f_q\|^2_{L^2(\mu^q)}.$$

Proof. We start with a few facts.

- Fact 1 : random variables of the type $I_1(h)$ with $\|h\|_{L^2(\mu)} = 1$ generate $\sigma(G)$.
- Fact 2 : for all λ the function $e^{i\lambda I_1(h)}$ can be approximated in $L^2(\mathbb{P})$ by complex linear combinations (through Taylor) of powers $I_1(h)^m$, $m \geq 1$.
- Fact 3 : If $X \in L^2(\sigma(G))$ is such that $\mathbb{E}(XI_1(h)^m) = 0$ for all h, m, then $\mathbb{E}(Xe^{i\lambda I_1(h)}) = 0$ for all λ, h implies that $X = 0$ almost surely.

As a consequence

$$\overline{v.s.}^{L^2(\sigma(G))}\{I_1(h)^m \, ; \, \|h\|_{L^2(\mu)} = 1 \,\text{and}\;\; m \geq 1\} = L^2(\sigma(G)).$$

Hence, all we need to do is to show the theorem for random variables of the type $I_1(h)^m$, i.e. we need to show that every $F = I_1(h)^m$ admits a representation (10). But we already now that there exist $C_{q,m}$, some real constants, such that

$$\begin{aligned} I_1(h)^m &= \sum_{q=0}^{m} C_{q,m} H_q(I_1(h)), \\ &= \sum_{q=0}^{m} C_{q,m} I_q(h^{\otimes q}). \end{aligned}$$

□

3 Elements of Malliavin Calculus

We work within the framework of a Gaussian measure G with the control measure μ having the suitable properties. We associate to all $F \in L^2(\sigma(G))$ an expansion $F = \mathbb{E}(F) + \sum_{q\geq 1} I_q(f_q)$.

3.1 The derivative operator D

We take

$$\text{dom}(D) := \left\{ F \in L^2(\sigma(G)) \, : \, \sum_{q=1}^{\infty} qq! \|f_q\|^2_{L^2(\mu^q)} < \infty \right\}$$

For $F \in \text{dom}(D)$ we define

$$D_t F = \sum_{q=1}^{\infty} q I_{q-1}(f_q(t, \bullet)), \, t \in A$$

where $\bullet$ indicates that we integrate over the $(q-1)$ remaining variables. We can then see that

$$\begin{aligned} \mathbb{E}\left[\int_A (D_t F)^2 \mu(dt) \right] &= \int_A \mu(dt) E\left[\left(\sum_{q=1}^{\infty} q I_{q-1}(f_q(t, \bullet)) \right)^2 \right] \\ &= \int_A \mu(dt) \sum_{q=1}^{\infty} q^2 (q-1) \|f_q(t, \bullet)\|^2 \\ &= \sum_{q=1}^{\infty} qq! \|f_q\|^2_{L^2(\mu^q)} < \infty. \end{aligned}$$

First note that for $f_q = h^{\otimes q}$ with $\|h\| = 1$, then $I_q(f_q) = H_q(I_1(h))$ and

$$\begin{aligned} D_t I_q(f_q) &= q I_{q-1}(f_q(t, \bullet)) = q I_{q-1}(h^{\otimes q-1}) h(t) = q H_{q-1}(I_1(h)) h(t) \\ &= H_q'(I_1(h)) h(t). \end{aligned} \tag{11}$$

In particular $D_t I_1(h) = h(t)$. Using this fact together with several approximation arguments one can prove the following *chain rules*.

Proposition 5. *[Chain rule 1]Let* $h_1, \ldots, h_d \in L^2(\mu)$ *and take* $f : \mathbb{R}^d \to \mathbb{R} \in C_b^1$*. Now define* $F = f(I_1(h_1), \ldots, I_1(h_d))$*. Then* $F \in \mathrm{dom}(D)$*, and*

$$D_t F = \sum_{j=1}^{d} \partial_{x_j} f(I_1(h_1), \ldots, I_1(h_d)) h_j(t).$$

The requirement that the functions be C_b^1 (differentiable with bounded derivatives) is too stringent, and can be replaced by polynomial tail behavior.

Proposition 6. *[Chain rule 2] Take* $F \in \mathrm{dom}(D)$ *and* $f : \mathbb{R} \to \mathbb{R} \in C_b^1$*. Then*

$$D_t f(F) = f'(F) D_t F.$$

Note that nowhere do we suppose that F have a density; we could end up sometimes with random variables defined a.e. and for which $f'(F)$ is only defined almost everywhere. Assuming that F has a density one can go a step further.

Proposition 7. *[Chain rule 3] Take* $F \in \mathrm{dom}(D)$ *and* $f : \mathbb{R} \to \mathbb{R}$ *Lipschitz (in particular absolutely continuous and a.e. differentiable). Suppose moreover that* F *has a density (wrt Lebesgue measure). Then*

$$D_t f(F) = f'(F) D_t F.$$

Working upwards we can also show the last chain rule, which will in particular allow us to work with polynomials (and hence compute moments).

Proposition 8. *[Chain rule 4] If* $F = \sum_{j=0}^{M} I_j(f_j)$ *be a finite sum of multiple integrals (in particular having density), then*

$$D_t p(F) = p'(F) D_t F$$

for every polynomial $p : \mathbb{R} \to \mathbb{R}$*.*

3.2 Generator of the Ornstein-Uhlenbeck semigroup L

We take

$$\mathrm{dom}(L) := \left\{ F \in L^2(\sigma(G)) \,:\, \sum_{q=1}^{\infty} q^2 q! \|f_q\|^2_{L^2(\mu^q)} < \infty \right\}.$$

For all $F \in \mathrm{dom}(L)$ we define

$$LF = -\sum_{q=1}^{\infty} q I_q(f_q).$$

For every $F \in L^2(\sigma(G))$ we also define

$$L^{-1}F = -\sum_{q=1}^{\infty} \frac{1}{q} I_q(f_q).$$

This is a pseudo-inverse of the operator L, because

$$LL^{-1}F = F - \mathbb{E}(F).$$

Note that $L^{-1}(F) \in \mathrm{dom}(D)$ and $\mathrm{dom}(L)$ always, because this is just the chaotic expansion of a r.v. whose kernels are I_q/q which can be safely multiplied by q and q^2.

Proposition 9. [Malliavin integration by parts] *Assume that* $F, G \in L^2(\sigma(G))$ *with* $\mathbb{E}(F) = 0$ *and* $G \in \mathrm{dom}(D)$*. Then*

$$\mathbb{E}(FG) = \mathbb{E}(\langle DG, -DL^{-1}F \rangle_{L^2(\mu)}).$$

Proof. Again by density arguments we just prove it for $F = I_q(f)$ and $G = I_p(g)$. But then

$$\mathbb{E}(FG) = \delta_{p,q} q! \langle f, g \rangle_{L^2(\mu^q)}$$

and

$$F_t G = p I_{p-1}(g(t, \bullet)).$$

Also we have

$$L^{-1} F = -\frac{1}{q} I_q(f) \text{ and } - D_t L^{-1} F = I_{q-1}(f(t, \bullet))$$

so that, taking expectations, we get

$$\begin{aligned} \mathbb{E}(\langle DG, -DL^{-1}F \rangle_{L^2(\mu)}) &= \int_A \mu(dt) \mathbb{E}[p I_{p-1}(g(t, \bullet)) I_{q-1}(f(t, \bullet))] \\ &= \delta_{p,q} p \int_A \mu(dt)(p-1)! \int_{A^{p-1}} g(t, \bar{x}_{p-1}) f(t, \bar{x}_{p-1}) d\mu^{p-1} \\ &= \delta_{p,q} p! \langle f, g \rangle_{L^2(\mu^p)}. \end{aligned}$$

□

Corollary 2. *Assume that $\mathbb{E}(F) = 0$ for $F \in \mathrm{dom}(D)$. Also assume that f is such that the chain rule applies. Then*

$$\begin{aligned} \mathbb{E}(F f(F)) &= \mathbb{E}(\langle DF, -DL^{-1}F \rangle_{L^2(\mu)}) \\ &= \mathbb{E}(f'(F) \langle DF, -DL^{-1}F \rangle_{L^2(\mu)}) \end{aligned}$$

Corollary 3. *Take $F = I_q(f)$ and $n \geq 1$. Then*

$$\begin{aligned} \mathbb{E}(F^{n+1}) = \mathbb{E}(F F^n) &= n \mathbb{E}(F^{n-1} \langle DF, -DL^{-1}F \rangle_{L^2(\mu)}) \\ &= \frac{n}{q} \mathbb{E}(F^{n-1} \|DF\|^2_{L^2(\mu)}), \end{aligned}$$

where $L^{-1}F = -\frac{1}{q}F$. In particular, we have

$$\mathbb{E}(F^4) = \frac{3}{q} \mathbb{E}\left(F^2 \|DF\|^2_{L^2(\mu)} \right).$$

Part III
Stein meets Malliavin

Via the Stein's approach, we have already seen that for any integrable random variable $F \in L^1(\mathbb{P})$ we have

$$d_{TV}(F, \mathscr{N}(0,1)) \leq \sup_{f \in \mathscr{F}_{\mathrm{TV}}} \left| E(F f(F) - f'(F)) \right|$$

with $\mathscr{F}_{\mathrm{TV}} = \{\|f\| \leq \sqrt{\pi/2}, \|f'\| \leq 2\}$. As was noted before, the supremum is annoying. The following theorem shows that, in the Gaussian framework, things are extremely favorable. From hereon, we assume that G is a Gaussian random measure over $(A, \mathcal{A}, \mu)$, and all random variables are measurable functionals of G.

Theorem 5. *Let $F \in \mathrm{dom}(D)$ with $\mathbb{E}(F) = 0$. Assume that F has a density (to use Proposition 7). Then*

$$d_{TV}(F, \mathscr{N}(0,1)) \leq 2 \mathbb{E} \left| 1 - \langle DF, -DL^{-1}F \rangle_{L^2(\mu)} \right|.$$

We have thus reduces the problem of controlling the TV distance to the computation of an expectation!

Proof. For every $f \in \mathscr{F}_{\text{TV}}$, and using Proposition 7 we have

$$\begin{aligned}\left|\mathbb{E}(Ff(F) - f'(F))\right| &= \left|\mathbb{E}(f'(F)(\langle DF, -DL^{-1}F\rangle_{L^2(\mu)} - 1))\right| \\ &\le 2\mathbb{E}\left|(\langle DF, -DL^{-1}F\rangle_{L^2(\mu)} - 1)\right|.\end{aligned}$$

Note that to obtain the last inequality in above, we use the fact that $\|f'\|_\infty \le 2$ for every $f \in \mathscr{F}_{\text{TV}}$. ☐

Corollary 4. *If $F = I_1(h)$, i.e. $F \sim \mathscr{N}(0, \|h\|^2)$ then*

$$\langle DF, -DL^{-1}F\rangle_{L^2(\mu)} = \|h\|^2.$$

Hence

$$d_{TV}(I_1(h), \mathscr{N}(0,1)) \le 2\left|1 - \|h\|^2\right|.$$

Corollary 5. *If $F = I_p(h)$ for some $f \in L^2_{sym}(\mu^p)$, then $\langle DF, -DL^{-1}F\rangle_{L^2(\mu)} = \frac{1}{p}\|DF\|^2$, and hence*

$$\begin{aligned}d_{TV}(F, \mathscr{N}(0,1)) &\le 2\mathbb{E}\left|1 - \frac{1}{p}\|DF\|^2\right| \\ &\le 2\sqrt{\mathbb{E}\left(1 - \frac{1}{p}\|DF\|^2\right)^2}.\end{aligned}$$

Therefore, for a sequence $\{F_n = I_p(h_n)\}_{n\ge1}$ of multiple integrals of a fixed order $p \ge 2$ such that $\mathbb{E}(F_n)^2 \to 1$, we have

$$\|DF_n\|^2_{L^2(\mu)} \overset{L^2(\mathbb{P})}{\longrightarrow} p \quad \Longrightarrow \quad F_n \overset{law}{\to} \mathscr{N}(0,1).$$

Corollary 6. *If $F = I_p(h)$ for some $f \in L^2_{sym}(\mu^p)$, and $\mathbb{E}(F^2) = \sigma^2$. Then*

$$d_{TV}(F, \mathscr{N}(0,\sigma^2)) \le \frac{2}{\sigma^2}\sqrt{\mathbb{E}\left(\sigma^2 - \frac{1}{p}\|DF\|^2\right)^2}. \tag{12}$$

Hence,

$$d_{TV}(F, \mathscr{N}(0,1)) \le 2|1-\sigma^2| + \frac{2}{\sigma^2}\sqrt{Var\left(\frac{1}{p}\|DF\|^2\right)}. \tag{13}$$

Moreover, for a sequence $\{F_n = I_p(h_n)\}_{n\ge1}$ of multiple integrals of a fixed order $p \ge 2$ such that $\mathbb{E}(F_n^2) \to \sigma^2 > 0$, we have

$$\|DF_n\|^2_{L^2(\mu)} \overset{L^2(\mathbb{P})}{\longrightarrow} p \times \sigma^2 \quad \Longrightarrow \quad F_n \overset{law}{\to} \mathscr{N}(0,\sigma^2).$$

Proof. For the claim (12), use the facts that if $N \sim \mathscr{N}(0,\sigma^2)$, then $\frac{N}{\sigma} \sim \mathscr{N}(0,1)$ together with

$$d_{\text{TV}}(F, \mathscr{N}(0,\sigma^2)) = d_{\text{TV}}(\frac{F}{\sigma}, \frac{N}{\sigma}).$$

Now, just left to apply Corollary 5. For the claim (13), use the triangular inequality, Corollary 4 and the relation (12). Note that when $\mathbb{E}(F^2) = \sigma^2$, then $\mathbb{E}(\frac{1}{p}\|DF\|) = \mathbb{E}(F^2)$, by using the Malliavin integration by part formula. ☐

Now, we are ready to state that approximating random variables in a fixed Wiener chaos by Gaussian is a nontrivial enterprise.

Theorem 6 (Nourdin, Peccati (2009) [3]). *Let $p \ge 2$ and $f \in L^2_{sym}(\mu^p) \ne 0$. Take $F = I_p(f)$. Then*

$$\begin{aligned}d_{TV}(F, \mathscr{N}(0,1)) &\le 2\left|1 - \mathbb{E}(F^2)\right| + 2\sqrt{Var\left(\frac{1}{p}\|DF\|^2\right)} \\ &\le 2\left|1 - \mathbb{E}(F^2)\right| + 2\sqrt{\frac{p-1}{3p}}\sqrt{\mathbb{E}(F^4) - 3\mathbb{E}(F^2)^2}.\end{aligned} \tag{14}$$

Note that $\sqrt{\frac{p-1}{3p}} \le \frac{2}{\sqrt{3}}$ and is thus independent of p.

[3] Nourdin, I., Peccati, G. (2009) Stein's method on Wiener chaos. *Probab. Theory Related Fields*, 45, no. 1-2, 75-118

In particular in order to have a CLT in a fixed Wiener chaos it suffices to control the fourth moments (whereas before this discovery, one had to show convergence of all moments!).

Corollary 7 (Nualart, Peccati (2005))**.** *Assume that $F_n = I_p(f_n)$ for $p \geq 2$ such that $\mathbb{E}(F_n^2) \to 1$. Then*

$$F_n \to \mathscr{N}(0,1)$$

in TV distance (and in particular, in distribution) if and only if

$$\mathbb{E}(F_n^4) \to 3 = \mathbb{E}(\mathscr{N}(0,1)^4).$$

Remark 10. *The "If" part of Corollary 7 is a consequence of* (14)*. On the other hand if $\mathbb{E}(F_n^2) \to 1$ and $F_n \to \mathscr{N}(0,1)$ in distribution, then for all $r \geq 2$ the sequence $\{\mathbb{E}(|F_n|^r)\}$ is bounded by hypercontractivity and thus, for all $r \geq 3$, we have*

$$\mathbb{E}(F_n^r) \to \mathbb{E}(\mathscr{N}(0,1)^r).$$

Proof of Theorem 6. We aim to prove that

$$\mathrm{Var}(\frac{1}{p}\|DF\|^2) \leq \frac{p-1}{3p}\left\{\mathbb{E}(F^4) - 3\mathbb{E}(F^2)^2\right\}.$$

We are going to use the formula

$$\mathbb{E}(F^4) = \frac{3}{p}\mathbb{E}(F^2\|DF\|^2_{L^2(\mu)}).$$

The whole proof relies on the derivation of the chaotic decompositions of the rv's of interest. Now, for $F = I_p(f)$ recall that $D_tF = pI_{p-1}(f(t,\cdot))$ to write (using product formula)

$$\begin{aligned}
\frac{1}{p}\|DF\|^2 &= p\int_A \mu(dt)(I_{p-1}(f(t,\cdot)))^2 \\
&= p\int_A \mu(dt)\sum_{r=0}^{p-1} r!\binom{p-1}{r}^2 I_{2(p-1)-2r}(f(t,\cdot)\widetilde{\otimes_r} f(t,\cdot)) \\
&= p\sum_{r=0}^{p-1} r!\binom{p-1}{r}^2 I_{2p-2(r+1)}(f\widetilde{\otimes_{r+1}} f)
\end{aligned}$$

where to obtain this we have used a stochastic integral version of Fubini's theorem. Pursuing with a change of summation variables we deduce

$$\begin{aligned}
\frac{1}{p}\|DF\|^2 &= p\sum_{r=1}^{p}(r-1)!\binom{p-1}{r-1}^2 I_{2p-2r}(f\widetilde{\otimes_r} f) \\
&= p!\|f\|^2 + p\sum_{r=1}^{p-1}(r-1)!\binom{p-1}{r-1}^2 I_{2p-2r}(f\widetilde{\otimes_r} f).
\end{aligned}$$

Note that $p!\|f\|^2 = \mathbb{E}(F^2)$, and we get

$$\frac{1}{p}\|DF\|^2 = \mathbb{E}(F^2) + p\sum_{r=1}^{p-1}(r-1)!\binom{p-1}{r-1}^2 I_{2p-2r}(f\widetilde{\otimes_r} f).$$

Since we have already shown that $\frac{1}{p}\mathbb{E}(\|DF\|^2) = \mathbb{E}(F^2)$ we deduce, by orthogonality of Wiener chaoses, our first estimate

$$\begin{aligned}
\mathrm{Var}\left(\frac{1}{p}\|DF\|^2\right) &= \sum_{r=1}^{p-1} p^2(r-1)!^2\binom{p-1}{r-1}^4(2p-2r)!\|f\widetilde{\otimes_r} f\|^2 \\
&= \frac{1}{p^2}\sum_{r=1}^{p-1} r^2 r!^2\binom{p}{r}^4(2p-2r)!\|f\widetilde{\otimes_r} f\|^2.
\end{aligned} \tag{15}$$

Now we also have

$$F^2 = \sum_{r=0}^{p} r! \binom{p}{r}^2 I_{2p-2r}(f \widetilde{\otimes}_r f) = p!\|f\|^2 + \sum_{r=0}^{p-1} r! \binom{p}{r}^2 I_{2p-2r}(f \widetilde{\otimes}_r f). \tag{16}$$

We can now compute the fourth moment

$$\begin{aligned} \mathbb{E}(F^4) &= 3\mathbb{E}(F^2 \times \frac{1}{p}\|DF\|^2) \\ &= 3\mathbb{E}(F^2)^2 + 3\sum_{r=1}^{p-1} pr!(r-1)! \binom{p}{r}^2 \binom{p-1}{r-1}^2 (2p-2r)! \|f \widetilde{\otimes}_r f\|^2. \end{aligned} \tag{17}$$

Therefore,

$$\begin{aligned} \mathbb{E}(F^4) - 3\mathbb{E}(F^2)^2 &= 3p \sum_{r=1}^{p-1} r!(r-1)! \binom{p}{r}^2 \binom{p-1}{r-1}^2 (2p-2r)! \|f \widetilde{\otimes}_r f\|^2 \\ &= \frac{3}{p} \sum_{r=1}^{p-1} r r!^2 \binom{p}{r}^4 (2p-2r)! \|f \widetilde{\otimes}_r f\|^2. \end{aligned} \tag{18}$$

Comparing the sums in (15) and (18) we recover the desired inequality

$$\operatorname{Var}\left(\frac{1}{p}\|DF\|^2\right) \leq \frac{p-1}{3p} \left\{ \mathbb{E}(F^4) - 3\mathbb{E}(F^2)^2 \right\}. \tag{19}$$

Note that the following estimate is also in order:

$$\frac{p-1}{3p} \left\{ \mathbb{E}(F^4) - 3\mathbb{E}(F^2)^2 \right\} \leq (p-1)\operatorname{Var}\left(\frac{1}{p}\|DF\|^2\right).$$

□

Remark 11. If one instead using the clever Malliavin relation

$$\mathbb{E}(F^4) = \frac{3}{p}\mathbb{E}(F^2\|DF\|^2_{L^2(\mu)})$$

by expanding F^4 on Wiener chaoses to compute $\mathbb{E}(F^4)$ what we will end up with (take into account 16)

$$\begin{aligned} \mathbb{E}(F^4) = \mathbb{E}(F^2 \times F^2) &= \sum_{r=0}^{p} r!^2 \binom{p}{r}^4 (2p-2r)! \|f \widetilde{\otimes}_r f\|^2 \\ &= p!^2\|f\|^4 + (2p)!\|f \widetilde{\otimes}_0 f\|^2 + \sum_{r=1}^{p-1} r!^2 \binom{p}{r}^4 (2p-2r)! \|f \widetilde{\otimes}_r f\|^2 \\ &= \mathbb{E}(F^2)^2 + (2p)!\|f \widetilde{\otimes}_0 f\|^2 + \sum_{r=1}^{p-1} r!^2 \binom{p}{r}^4 (2p-2r)! \|f \widetilde{\otimes}_r f\|^2. \end{aligned}$$

Hence, we have the presence of the norm of the zero-contraction $\|f \widetilde{\otimes}_0 f\|^2$, which in fact never appears in the expansion $\operatorname{Var}\left(\frac{1}{p}\|DF\|^2\right)$. Therefore, using this approach, one needs to represent the norm of the zero-contraction $\|f \widetilde{\otimes}_0 f\|^2$ in terms of the norm of other non-zero contraction to be able to do comparison. Hopefully, this can be done and is the message of the next exercise. Here, we highlight that the appearance of norms (inner products) of zero-contractions involving the kernel f is in fact the main obstacle in front towards generalization of the Malliavin-Stein approach for non-Gaussian approximation using product formula as the main tools. A typical example is when the target distribution is of the form $N_1 \times N_2$ and $N_1, N_2 \sim \mathscr{N}(0,1)$ are independent.

Exercise 6. ***(a)*** *Show that*

$$(2p)!\|\widetilde{f \otimes_0 f}\|^2 = 2(p!)^2\|f\|^4 + p!^2 \sum_{r=1}^{p-1} \binom{p}{r}^2 \|f \otimes_r f\|^2.$$

(b) *Use part (a) to show that*

$$\mathbb{E}(F^4) - 3\mathbb{E}(F^2)^2 = p!^2 \sum_{r=1}^{p-1} \binom{p}{r} \left\{ \|f \otimes_r f\|^2 + \binom{2p-2r}{p-r} \|\widetilde{f \otimes_r f}\|^2 \right\}.$$

Remark 12. *The computations in the proof show that there exists a constant $c := c(p) > 0$ only depending on p for which*

$$\begin{aligned} d_{TV}(I_p(f), \mathscr{N}(0,1)) &\le c(p) \max_{r=1,\dots,p-1} \left\{ \|\widetilde{f \otimes_r f}\|_{L^2(\mu^{2p-2r})} \right\} \\ &\le c(p) \max_{r=1,\dots,p-1} \left\{ \|f \otimes_r f\|_{L^2(\mu^{2p-2r})} \right\} \end{aligned} \tag{20}$$

The estimate in (20) *is that which is most used in practical situations since it is easier to estimate contractions rather than moments.*

3.3 The multidimensional case

Let $d \ge 2$, and fix d natural numbers $1 \le p_1 \le p_2 \le \cdots \le p_d$. Consider a sequence of d-dimensional random vectors of the form

$$F_n = (F_n^1, \cdots, F_n^d) = (I_{p_1}(f_n^1), \cdots, I_{p_d}(f_n^d)). \tag{21}$$

Our aim in this section is to prove the following multidimensional version of the fourth moment theorem due to Peccati–Tudor (2005).

Theorem 7. *Let F_n be a sequence of d-dimensional random vectors of the form* (22) *such that*

$$\lim_{n\to\infty} \mathbb{E}(F_n^i \times F_n^j) = \delta_{ij}, \qquad 1 \le i, j \le d.$$

Then the following statements are equivalent.

(a) $F_n^i \overset{law}{\to} \mathscr{N}(0,1)$, *for all* $1 \le i \le d$.

(b) $\mathbb{E}(F_n^i)^4 \to 3$, *for all* $1 \le i \le d$.

(c) $\|DF_n^i\|^2 \overset{L^2(\mathbb{P})}{\to} p_i$, *for all* $1 \le i \le d$.

(d) $\|f_n^i \otimes_r f_n^i\|_{L^2(\mu^{2p_i-2r})} \to 0$, *for all* $1 \le i \le d$, *and* $r = 1, \cdots, p_i - 1$.

(e) $F_n \overset{law}{\to} \mathscr{N}_d(0, I_d)$.

In other words, Theorem 7 tells us that the component-wise convergence to Gaussian distributions implies the joint convergence of the vector to the multidimensional Gaussian. There are different ways to prove Theorem 7. Here, we follow the path was developed by Nualart & Ortiz-Latorre. Their strategy mainly consists of showing that the characteristic function of any adherence value in distribution satisfies in the same ordinary differential equation as the characteristic function of the d-dimensional Gaussian random variable. The advantage of their approach compare to multidimensional Stein's method (which is more involved compare to one dimensional version) is its simplicity and as drawback this approach is not quantitative, and hence one can not provide any rate of convergence. The interested reader can consult [13, Chapter 6] for a proof of Theorem 7 using multidimensional Stein's method.

The token of the main part of Theorem 7 can be decoded using the following lemma in which the Malliavin derivative matrix

$$\Gamma_n = (\Gamma_n^{i,j})_{1\le i,j\le d} = \big(\langle DF_n^i, DF_n^j\rangle_{L^2(\mu)}\big)_{1\le i,j\le d}$$

plays an essential role. The Malliavin derivative matrix Γ is in the core of the studies of regularities of laws of random vectors (see [15, Chapter 2]). In the next lemma, we will use again heavily the specific structure of the underlying random variables.

Lemma 4. *Let*

$$F_n = (F_n^1, \cdots, F_n^d) = (I_{p_1}(f_n^1), \cdots, I_{p_d}(f_n^d)) \tag{22}$$

such that for every $1 \le i, j \le d$, $\mathbb{E}(F_n^i \times F_n^j) \to \delta_{i,j}$. *Then*

$$\|DF_n^i\|^2 \overset{L^2(\mathbb{P})}{\to} p_i \quad \Longrightarrow \quad \Gamma_n^{i,j} = \langle DF_n^i, DF_n^j\rangle_{L^2(\mu)} \overset{L^2(\mathbb{P})}{\to} \sqrt{p_i p_j}\delta_{i,j}.$$

Proof. We need to show that for any $i < j$ (and so $p_i \le p_j$) we have

$$\lim_{n\to\infty} \mathbb{E}\left(\langle DF_n^i, DF_n^j\rangle^2_{L^2(\mu)}\right) = 0.$$

Using exercise 2, part (a), we know that

$$\begin{aligned}\mathbb{E}\left(\langle DF_n^i, DF_n^j\rangle^2\right)_{L^2(\mu)} &= \sum_{r=1}^{p_i} \frac{(p_i!p_j!)^2}{((p_i-r)!(p_j-r)!(r-1)!)^2}\|f_n^i\widetilde{\otimes}_r f_n^j\|^2 \\ &\le \sum_{r=1}^{p_i} \frac{(p_i!p_j!)^2}{((p_i-r)!(p_j-r)!(r-1)!)^2}\|f_n^i\otimes_r f_n^j\|^2\end{aligned}$$

So, we are left to show that $\|f_n^i \otimes_r f_n^j\|^2 \to 0$ for all $1 \le r \le p_i$. Using the very definition of the contraction, Fubini's theorem, and Cauchy–Schwarz inequality, we can write

$$\begin{aligned}\|f_n^i \otimes_r f_n^j\|^2 &= \langle f_n^i \otimes_r f_n^j, f_n^i \otimes_r f_n^j, \rangle_{L^2(\mu^{p_i+p_j-2r})} \\ &= \langle f_n^i \otimes_{p_i-r} f_n^i, f_n^j \otimes_{p_j-r} f_r^j\rangle_{L^2(\mu^{2r})} \\ &\le \|f_n^i \otimes_{p_i-r} f_n^i\| \times \|f_n^j \otimes_{p_j-r} f_n^j\|.\end{aligned} \tag{23}$$

Case (a): if $r = p_i = p_j$, then $\|f_n^i \otimes_r f_n^j\|^2 = \left(\mathbb{E}(F_n^i \times F_n^j)\right)^2 \to 0$ by assumption. Case (b): if $1 \le r \le p_i - 1$, then assumption $\|DF_n^i\|^2 \overset{L^2(\mathbb{P})}{\to} p_i$ implies that $F_n^i \overset{\text{law}}{\to} \mathscr{N}(0,1)$ and therefore $\|f_n^i \otimes_r f_n^i\|^2 \to 0$ for all $1 \le r \le p_i - 1$. Hence, the right hand side inequality (23) tends to zero. Case (c): if $r = p_i < p_j$. In this case, the right hand side of (23) takes the form

$$\begin{aligned}\|f_n^i \otimes_{p_i-r} f_n^i\| \times \|f_n^j \otimes_{p_j-r} f_n^j\| &= \|f_n^i\|^2 \times \|f_n^j \otimes_{p_j-r} f_n^j\| \\ &= \mathbb{E}(F_n^i)^2 \times \|f_n^j \otimes_{p_j-r} f_n^j\| \to 0,\end{aligned}$$

Because $\mathbb{E}(F_n^i)^2 \to 1$ and so bounded and $\|f_n^j \otimes_{p_j-r} f_n^j\| \to 0$. □

Proof. Proof of Theorem 7. It is enough to prove the implication $(c) \Rightarrow (e)$. Since $\mathbb{E}(F_n^i \times F_n^j) \to \delta_{i,j}$, for $i = j$, this implies that $\sup_{n\ge1} \mathbb{E}(F_n^i)^2 < +\infty$. Therefore, the sequence $\{F_n\}_{n\ge1}$ is tight, and so it is enough to show that the limit of any convergence in distribution subsequence $\{F_{n_k}\}_{n\ge1}$ is in fact $\mathscr{N}_d(0, I_d)$. To this end, assume that $F_{n_k} \overset{\text{law}}{\to} F_\infty$, as $k \to \infty$ for some random vector $F_\infty = (F_\infty^1, \cdots, F_\infty^d)$. By our assumptions, first we have that $F_\infty^i \in L^2(\Omega)$ for all $1 \le i \le d$, and moreover $\mathbb{E}(F_\infty^i \times F_\infty^j) = 0$ if $i \ne j$. Now, let's denote the characteristic function $\varphi_n(t) = \mathbb{E}(e^{i\langle t, F_n\rangle_{\mathbb{R}^d}})$ for $t \in \mathbb{R}^d$. Then

$\varphi_{n_k}(t) \to \varphi(t)$ for any t, where φ_∞ is the characteristic function of F_∞. Note that the fact that $F_\infty^j \in L^2(\Omega)$ implies that the partial derivatives $\frac{\partial}{\partial t_j}\varphi_\infty = i\mathbb{E}(F_\infty^j e^{i\langle t,F_\infty\rangle_{\mathbb{R}^d}})$ are well defined. Now, continuous mapping theorem tells us that

$$F_{n_k} e^{i\langle t,F_{n_k}\rangle_{\mathbb{R}^d}} \overset{\text{law}}{\to} F_\infty^j e^{i\langle t,F_\infty\rangle_{\mathbb{R}^d}}. \tag{24}$$

Note that the sequence in the left hand side of (24) is bounded in $L^2(\Omega)$ and so uniformly integrable. Hence, for all $1 \le j \le d$, and $t \in \mathbb{R}^d$, as $k \to \infty$:

$$\frac{\partial}{\partial t_j}\varphi_{n_k}(t) = i\mathbb{E}(F_{n_k}^j e^{i\langle t,F_{n_k}\rangle_{\mathbb{R}^d}}) \to \frac{\partial}{\partial t_j}\varphi_\infty = i\mathbb{E}(F_\infty^j e^{i\langle t,F_\infty\rangle_{\mathbb{R}^d}}). \tag{25}$$

On the other hand side, using integration by part formula: (note that $\mathbb{E}(F_{n_k}^j) = 0$)

$$\begin{aligned}
\mathbb{E}\left(F_{n_k}^j e^{i\langle t,F_{n_k}\rangle_{\mathbb{R}^d}}\right) &= \mathbb{E}\left(\langle De^{i\langle t,F_{n_k}\rangle_{\mathbb{R}^d}}, -DL^{-1}F_{n_k}^j\rangle\right) \\
&= -\frac{i}{p_j}\sum_{l=1}^d t_l\,\mathbb{E}\left(e^{i\langle t,F_{n_k}\rangle_{\mathbb{R}^d}}\mathbb{E}(\langle DF_{n_k}^l, DF_{n_k}^j\rangle)\right) \\
&= -\frac{i}{p_j}\sum_{l=1}^d t_l\,\mathbb{E}\left(e^{i\langle t,F_{n_k}\rangle_{\mathbb{R}^d}}\Gamma_{n_k}^{l,j}\right).
\end{aligned}$$

Therefore

$$\frac{\partial}{\partial t_j}\varphi_{n_k}(t) = -\frac{1}{p_j}\sum_{l=1}^d t_l\,\mathbb{E}\left(e^{i\langle t,F_{n_k}\rangle_{\mathbb{R}^d}}\Gamma_{n_k}^{l,j}\right).$$

Hence, Lemma 4 implies that the characteristic function φ_∞ satisfies in equation

$$\frac{\partial}{\partial t_j}\varphi_\infty(t) = -t_j\varphi_\infty(t),$$

for all $j = 1, \cdots, d$ and $t \in \mathbb{R}^d$. Therefore, the only possibility for φ_∞ is to be the characteristic function of $\mathscr{N}_d(0, I_d)$. □

We finish this section with the following very general result.

Theorem 8. *Let $\{F_n\}_{n\ge1}$ be a square-integrable sequence with the following chaos decompositions: for every $n \ge 1$*

$$F_n = \sum_{p=1}^\infty I_p(f_{n,p}). \tag{26}$$

In addition, assume the following:

(a) for all $p \ge 1$, we have $p!\|f_{n,p}\|^2 \to \sigma_p^2$.

(b) $\sum_{p\ge1}\sigma_p^2 < +\infty$.

(c) for all $p \ge 2$ and every $r = 1, \cdots, p-1$, we have $\|f_{n,p} \otimes_r f_{n,p}\| \to 0$, as $n \to \infty$.

(d)

$$\lim_{N\to\infty}\sup_{n\ge1}\sum_{p=N+1}^\infty p!\,\|f_{n,p}\|^2 = 0.$$

Then we have $F_n \overset{law}{\to} \mathscr{N}(0,\sigma^2)$.

Proof. For all $N \geq 1$, set

$$F_{n,N} = \sum_{p=1}^{N} I_p(f_{n,p})$$
$$G_N \sim \mathcal{N}(0, \sigma_1^2 + \cdots + \sigma_N^2)$$
$$G \sim \mathcal{N}(0, \sigma^2).$$

Therefore, for any $t \in \mathbb{R}$:

$$\begin{aligned}\left|\mathbb{E}(e^{itF_n}) - \mathbb{E}(e^{itG})\right| &\leq \left|\mathbb{E}(e^{itF_n}) - \mathbb{E}(e^{itF_{n,N}})\right| \\ &+ \left|\mathbb{E}(e^{itF_{n,N}}) - \mathbb{E}(e^{itG_N})\right| \\ &+ \left|\mathbb{E}(e^{itG_N}) - \mathbb{E}(e^{itG})\right| := a_{n,N} + b_{n,N} + c_N.\end{aligned}$$

Note that

$$c_N = \left|e^{-\frac{t^2}{2}(\sigma_1^2 + \cdots + \sigma_N^2)} - e^{-\frac{t^2}{2}\sigma^2}\right| \leq \frac{t^2}{2}\left|\sigma^2 - \sum_{i=1}^{N} \sigma_i^2\right| \to 0,$$

as $N \to \infty$, because of assumption (b). Moreover,

$$\begin{aligned}\sup_{n\geq 1} a_{n,N} &= \sup_{n\geq 1}\left|\mathbb{E}(e^{itF_n}) - \mathbb{E}(e^{itF_{n,N}})\right| \\ &\leq |t| \sup_{n\geq 1} \mathbb{E}|F_n - F_{n,N}| \leq |t| \sqrt{\sup_{n\geq 1} \mathbb{E}(F_n - F_{n,N})^2} \\ &\leq |t| \sqrt{\sup_{n\geq 1} \sum_{p\geq N+1} \sigma_p^2} \to 0,\end{aligned}$$

by assumption (d). Hence, for $\varepsilon > 0$, choose N large enough so that $\sup_{n\geq 1} a_{n,N} \leq \varepsilon/3$ and $c_N \leq \varepsilon/3$. Also, according to Peccati–Tudor multidimensional version of the fourth moment theorem, we have in fact that, as $n \to \infty$,

$$(I_1(f_{n,1}), \cdots, I_N(f_{n,N})) \overset{\text{law}}{\to} \mathcal{N}_N\big(0, \operatorname{diag}(\sigma_1^2, \cdots, \sigma_N^2)\big).$$

Therefore, $F_{n,N} = \sum_{p=1}^{N} I_p(f_{n,p}) \overset{\text{law}}{\to} \mathcal{N}(0, \sigma_1^2 + \cdots + \sigma_N^2)$. Hence, $b_{n,N} \leq \varepsilon/3$ if n is large enough. □

4 Applications and related topics

The following two subsections are borrowed mostly from [4]

4.1 Breuer–Major Theorem

Let $\{X_k\}_{k\geq 1}$ be a centered stationary Gaussian family. In this context, stationary just means that there exists $\rho : \mathbb{Z} \to \mathbb{R}$ such that $\mathbb{E}[X_k X_l] = \rho(k - l)$, $k, l \geq 1$. Assume further that $\rho(0) = 1$, that is, each X_k is $\mathcal{N}(0, 1)$ distributed.

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a measurable function satisfying

$$E[\varphi^2(X_1)] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi^2(x) e^{-x^2/2} dx < \infty. \tag{27}$$

[4]Nourdin, Ivan: Lectures on Gaussian approximations with Malliavin calculus. Séminaire de Probabilités XLV, 389, Lecture Notes in Math., 2078, Springer, 2013.

It is a well-known fact that, when it verifies (27), the function φ may be expanded in $L^2(\mathbb{R}, e^{-x^2/2}dx)$ (in a unique way) in terms of Hermite polynomials as follows:

$$\varphi(x) = \sum_{q=0}^{\infty} a_q H_q(x). \tag{28}$$

Let $d \geq 0$ be the first integer $q \geq 0$ such that $a_q \neq 0$ in (28). It is called the *Hermite rank* of φ; it will play a key role in our study. Also, let us mention the following crucial property of Hermite polynomials with respect to Gaussian elements. For any integer $p, q \geq 0$ and any jointly Gaussian random variables $U, V \sim \mathscr{N}(0,1)$, we have

$$E[H_p(U)H_q(V)] = \begin{cases} 0 & \text{if } p \neq q \\ q!E[UV]^q & \text{if } p = q. \end{cases} \tag{29}$$

In particular (choosing $p = 0$) we have that $E[H_q(X_1)] = 0$ for all $q \geq 1$, meaning that $a_0 = E[\varphi(X_1)]$ in (28). Also, combining the decomposition (28) with (29), it is straightforward to check that

$$E[\varphi^2(X_1)] = \sum_{q=0}^{\infty} q! a_q^2. \tag{30}$$

We are now in position to state the celebrated Breuer-Major theorem.

Theorem 9 (Breuer, Major, 1983; see [6]). *Let $\{X_k\}_{k\geq 1}$ and $\varphi : \mathbb{R} \to \mathbb{R}$ be as above. Assume further that $a_0 = \mathbb{E}[\varphi(X_1)] = 0$ and that $\sum_{k\in\mathbb{Z}} |\rho(k)|^d < \infty$, where ρ is the covariance function of $\{X_k\}_{k\geq 1}$ and d is the Hermite rank of φ (observe that $d \geq 1$). Then, as $n \to \infty$,*

$$V_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \varphi(X_k) \overset{\text{law}}{\to} \mathscr{N}(0, \sigma^2), \tag{31}$$

with σ^2 given by

$$\sigma^2 = \sum_{q=d}^{\infty} q! a_q^2 \sum_{k\in\mathbb{Z}} \rho(k)^q \in [0, \infty). \tag{32}$$

(The fact that $\sigma^2 \in [0, \infty)$ is part of the conclusion.)

The proof of Theorem 9 is far from being obvious. The original proof follows the method of the moments/cumulant, i.e. to show that *all* the moments of V_n converge to those of the Gaussian law $\mathscr{N}(0, \sigma^2)$. As anyone might guess, this required a high ability and a lot of combinatorics. In the proof we will offer, we use the strong approach of the fourth moment theorem. Hence, the complexity is the same as checking that the variance and the fourth moment of V_n converges to σ^2 and $3\sigma^4$ respectively, which is a drastic simplification with respect to the original proof. Before doing so, let us make some other comments.

Remark 13. 1. First, it is worthwhile noticing that Theorem 9 (strictly) contains the classical central limit theorem (CLT), which is not an evident claim at first glance. Indeed, let $\{Y_k\}_{k\geq 1}$ be a sequence of i.i.d. centered random variables with common variance $\sigma^2 > 0$, and let F_Y denote the common cumulative distribution function. Consider the pseudo-inverse F_Y^{-1} of F_Y, defined as

$$F_Y^{-1}(u) = \inf\{y \in \mathbb{R} : u \leq F_Y(y)\}, \quad u \in (0,1).$$

When $U \sim \mathscr{U}_{[0,1]}$ is uniformly distributed, it is well-known that $F_Y^{-1}(U) \overset{\text{law}}{=} Y_1$. Observe also that $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X_1} e^{-t^2/2} dt$ is $\mathscr{U}_{[0,1]}$ distributed. By combining these two facts, we get that $\varphi(X_1) \overset{\text{law}}{=} Y_1$ with

$$\varphi(x) = F_Y^{-1}\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt\right), \quad x \in \mathbb{R}.$$

Assume now that $\rho(0)=1$ and $\rho(k)=0$ for $k\neq 0$, that is, assume that the sequence $\{X_k\}_{k\geq 1}$ is composed of i.i.d. $\mathscr{N}(0,1)$ random variables. Theorem 9 yields that

$$\frac{1}{\sqrt{n}}\sum_{k=1}^n Y_k \overset{\text{law}}{=} \frac{1}{\sqrt{n}}\sum_{k=1}^n \varphi(X_k) \overset{\text{law}}{\to} \mathscr{N}\left(0,\sum_{q=d}^\infty q!a_q^2\right),$$

thereby concluding the proof of the CLT since $\sigma^2=\mathbb{E}[\varphi^2(X_1)]=\sum_{q=d}^\infty q!a_q^2$, see (30). Of course, such a proof of the CLT is like to crack a walnut with a sledgehammer. This approach has nevertheless its merits: it shows that the independence assumption in the CLT is not crucial to allow a Gaussian limit. Indeed, this is rather the summability of a series which is responsible of this fact, see also the second point of this remark.

2. Assume that $d\geq 2$ and that $\rho(k)\sim |k|^{-D}$ as $|k|\to\infty$ for some $D\in(0,\frac{1}{d})$. In this case, it may be shown that $n^{dD/2-1}\sum_{k=1}^n\varphi(X_k)$ converges in law to a non-Gaussian (non degenerated) random variable. This shows in particular that, in the case where $\sum_{k\in\mathbb{Z}}|\rho(k)|^d=\infty$, we can get a non-Gaussian limit. In other words, the summability assumption in Theorem 9 is, roughly speaking, equivalent (when $d\geq 2$) to the asymptotic normality. However, this line of results are really out of the scope of the course (for more information in this regard consult [13] and references therein).

Let us now prove Theorem 9. We first compute the limiting variance, which will justify the formula (32) we claim for σ^2. Thanks to (28) and (29), we can write

$$\begin{aligned}\mathbb{E}[V_n^2] &= \frac{1}{n}\mathbb{E}\left[\left(\sum_{q=d}^\infty a_q\sum_{k=1}^n H_q(X_k)\right)^2\right]=\frac{1}{n}\sum_{p,q=d}^\infty a_pa_q\sum_{k,l=1}^n \mathbb{E}[H_p(X_k)H_q(X_l)]\\ &= \frac{1}{n}\sum_{q=d}^\infty q!a_q^2\sum_{k,l=1}^n\rho(k-l)^q=\sum_{q=d}^\infty q!a_q^2\sum_{r\in\mathbb{Z}}\rho(r)^q\big(1-\frac{|r|}{n}\big)\mathbf{1}_{\{|r|<n\}}.\end{aligned}$$

When $q\geq d$ and $r\in\mathbb{Z}$ are fixed, we have that

$$q!a_q^2\rho(r)^q\big(1-\frac{|r|}{n}\big)\mathbf{1}_{\{|r|<n\}}\to q!a_q^2\rho(r)^q \quad \text{as } n\to\infty.$$

On the other hand, using that $|\rho(k)|=|\mathbb{E}[X_1X_{k+1}]|\leq\sqrt{\mathbb{E}[X_1^2]\mathbb{E}[X_{1+k}^2]}=1$, we have

$$q!a_q^2|\rho(r)|^q\big(1-\frac{|r|}{n}\big)\mathbf{1}_{\{|r|<n\}}\leq q!a_q^2|\rho(r)|^q\leq q!a_q^2|\rho(r)|^d,$$

with $\sum_{q=d}^\infty\sum_{r\in\mathbb{Z}}q!a_q^2|\rho(r)|^d=E[\varphi^2(X_1)]\times\sum_{r\in\mathbb{Z}}|\rho(r)|^d<\infty$, see (30). By applying the dominated convergence theorem, we deduce that $\mathbb{E}[V_n^2]\to\sigma^2$ as $n\to\infty$, with $\sigma^2\in[0,\infty)$ given by (32).

Let us next concentrate on the proof of (31). We shall do it in three steps of increasing generality (but of decreasing complexity!):

(i) when $\varphi=H_q$ has the form of a Hermite polynomial (for some $q\geq 1$);

(ii) when $\varphi=P\in\mathbb{R}[X]$ is a real polynomial;

(iii) in the general case when $\varphi\in L^2(\mathbb{R},e^{-x^2/2}dx)$.

We first show that (ii) implies (iii). That is, let us assume that Theorem 9 is shown for polynomial functions φ, and let us show that it holds true for any function $\varphi\in L^2(\mathbb{R},e^{-x^2/2}dx)$. We proceed by approximation. Let $N\geq 1$ be a (large) integer (to be chosen later) and write

$$V_n=\frac{1}{\sqrt{n}}\sum_{q=d}^N a_q\sum_{k=1}^n H_q(X_k)+\frac{1}{\sqrt{n}}\sum_{q=N+1}^\infty a_q\sum_{k=1}^n H_q(X_k)=:V_{n,N}+R_{n,N}.$$

Similar computations as above lead to

$$\sup_{n\geq 1} \mathbb{E}[R_{n,N}^2] \leq \sum_{q=N+1}^{\infty} q! a_q^2 \times \sum_{r\in\mathbb{Z}} |\rho(r)|^d \to 0 \text{ as } N \to \infty. \tag{33}$$

(Recall from (30) that $\mathbb{E}[\varphi^2(X_1)] = \sum_{q=d}^{\infty} q! a_q^2 < \infty$.) On the other hand, using (ii) we have that, for fixed N and as $n \to \infty$,

$$V_{n,N} \overset{\text{law}}{\to} \mathscr{N}\left(0, \sum_{q=d}^{N} q! a_q^2 \sum_{k\in\mathbb{Z}} \rho(k)^q\right). \tag{34}$$

It is then a routine exercise (details are left to the reader) to deduce from (33)-(34) that $V_n = V_{n,N} + R_{n,N} \overset{\text{law}}{\to} \mathscr{N}(0,\sigma^2)$ as $n \to \infty$, that is, that (iii) holds true.

Next, let us prove (i), that is, (31) when $\varphi = H_q$ is the qth Hermite polynomial. The space

$$\mathcal{H} := \overline{\text{span}\{X_1, X_2, \ldots\}}^{L^2(\Omega)}$$

being a real separable Hilbert space, it is isometrically isomorphic to either $\mathbb{R}^N$ (with $N \geq 1$) or $L^2(\mathbb{R}_+)$. Let us assume that $\mathcal{H} \simeq L^2(\mathbb{R}_+)$, the case where $\mathcal{H} \simeq \mathbb{R}^N$ being easier to handle. Let $\Phi : \mathcal{H} \to L^2(\mathbb{R}_+)$ be an isometry. Set $e_k = \Phi(X_k)$ for each $k \geq 1$. We have

$$\rho(k-l) = \mathbb{E}[X_k X_l] = \int_0^\infty e_k(x) e_l(x) dx, \quad k,l \geq 1 \tag{35}$$

If $B = (B_t)_{t\geq 0}$ denotes a standard Brownian motion, we deduce that

$$\{X_k\}_{k\geq 1} \overset{\text{law}}{=} \left\{\int_0^\infty e_k(t) dB_t\right\}_{k\geq 1},$$

these two families being indeed centered, Gaussian and having the same covariance structure (by construction of the e_k's). On the other hand, for any function $e \in L^2(\mathbb{R}_+)$ such that $\|e\|_{L^2(\mathbb{R}_+)} = 1$, we have

$$\begin{aligned} H_q\left(\int_0^\infty e(t) dB_t\right) &= I_q^B(e^{\otimes q}) \\ &= q! \int_0^\infty dB_{t_1} e(t_1) \int_0^{t_1} dB_{t_2} e(t_2) \ldots \int_0^{t_{q-1}} dB_{t_q} e(t_q). \end{aligned} \tag{36}$$

Let us go back to the proof of (i), that is, to the proof of (31) for $\varphi = H_q$. Recall that the sequence $\{e_k\}$ has be chosen for (35) to hold. Using (36), we can write $V_n = I_q^B(f_n)$, with

$$f_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} e_k^{\otimes q}.$$

We already showed that $E[V_n^2] \to \sigma^2$ as $n \to \infty$. So, according to fourth moment theorem, to get (i) it remains to check that $\|f_n \otimes_r f_n\|_{L^2(\mathbb{R}_+^{2q-2r})} \to 0$ for any $r = 1, \ldots, q-1$. We have

$$\begin{aligned} f_n \otimes_r f_n &= \frac{1}{n} \sum_{k,l=1}^{n} e_k^{\otimes q} \otimes_r e_l^{\otimes q} = \frac{1}{n} \sum_{k,l=1}^{n} \langle e_k, e_l \rangle_{L^2(\mathbb{R}_+)}^r e_k^{\otimes q-r} \otimes e_l^{\otimes q-r} \\ &= \frac{1}{n} \sum_{k,l=1}^{n} \rho(k-l)^r e_k^{\otimes q-r} \otimes e_l^{\otimes q-r}, \end{aligned}$$

implying in turn

$$
\begin{aligned}
&\|f_n \otimes_r f_n\|^2_{L^2(\mathbb{R}_+^{2q-2r})} \\
&= \frac{1}{n^2} \sum_{i,j,k,l=1}^{n} \rho(i-j)^r \rho(k-l)^r \langle e_i^{\otimes q-r} \otimes e_j^{\otimes q-r}, e_k^{\otimes q-r} \otimes e_l^{\otimes q-r} \rangle_{L^2(\mathbb{R}_+^{2q-2r})} \\
&= \frac{1}{n^2} \sum_{i,j,k,l=1}^{n} \rho(i-j)^r \rho(k-l)^r \rho(i-k)^{q-r} \rho(j-l)^{q-r}.
\end{aligned}
$$

Observe that $|\rho(k-l)|^r |\rho(i-k)|^{q-r} \le |\rho(k-l)|^q + |\rho(i-k)|^q$. This, together with other obvious manipulations, leads to the bound

$$
\begin{aligned}
\|f_n \otimes_r f_n\|^2_{L^2(\mathbb{R}_+^{2q-2r})} &\le \frac{2}{n} \sum_{k\in\mathbb{Z}} |\rho(k)|^q \sum_{|i|<n} |\rho(i)|^r \sum_{|j|<n} |\rho(j)|^{q-r} \\
&\le \frac{2}{n} \sum_{k\in\mathbb{Z}} |\rho(k)|^d \sum_{|i|<n} |\rho(i)|^r \sum_{|j|<n} |\rho(j)|^{q-r} \\
&= 2\sum_{k\in\mathbb{Z}} |\rho(k)|^d \times n^{-\frac{q-r}{q}} \sum_{|i|<n} |\rho(i)|^r \times n^{-\frac{r}{q}} \sum_{|j|<n} |\rho(j)|^{q-r}.
\end{aligned}
$$

Thus, to get that $\|f_n \otimes_r f_n\|_{L^2(\mathbb{R}_+^{2q-2r})} \to 0$ for any $r = 1, \ldots, q-1$, it suffices to show that

$$
s_n(r) := n^{-\frac{q-r}{q}} \sum_{|i|<n} |\rho(i)|^r \to 0 \;\text{ for any } r = 1, \ldots, q-1.
$$

Let $r = 1, \ldots, q-1$. Fix $\delta \in (0,1)$ (to be chosen later) and let us decompose $s_n(r)$ into

$$
s_n(r) = n^{-\frac{q-r}{q}} \sum_{|i|<[n\delta]} |\rho(i)|^r + n^{-\frac{q-r}{q}} \sum_{[n\delta]\le|i|<n} |\rho(i)|^r =: s_{1,n}(\delta,r) + s_{2,n}(\delta,r).
$$

Using Hölder inequality, we get that

$$
s_{1,n}(\delta,r) \le n^{-\frac{q-r}{r}} \left(\sum_{|i|<[n\delta]} |\rho(i)|^q \right)^{r/q} (1+2[n\delta])^{\frac{q-r}{q}} \le \text{cst} \times \delta^{1-r/q},
$$

as well as

$$
s_{2,n}(\delta,r) \le n^{-\frac{q-r}{r}} \left(\sum_{[n\delta]\le|i|<n} |\rho(i)|^q \right)^{r/q} (2n)^{\frac{q-r}{q}} \le \text{cst} \times \left(\sum_{|i|\ge[n\delta]} |\rho(i)|^q \right)^{r/q}.
$$

Since $1 - r/q > 0$, it is a routine exercise (details are left to the reader) to deduce that $s_n(r) \to 0$ as $n \to \infty$. Since this is true for any $r = 1, \ldots, q-1$, this concludes the proof of (i). It remains to show (ii), that is, convergence in law (31) whenever φ is a real polynomial. We shall use the multidimensional Peccati–Tudor theorem. Let φ have the form of a real polynomial. In particular, it admits a decomposition of the type $\varphi = \sum_{q=d}^{N} a_q H_q$ for some *finite* integer $N \ge d$. Together with (i), Peccati–Tudor Theorem yields that

$$
\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} H_d(X_k), \ldots, \frac{1}{\sqrt{n}} \sum_{k=1}^{n} H_N(X_k) \right) \overset{\text{law}}{\to} \mathscr{N}\big(0, \text{diag}(\sigma_d^2, \ldots, \sigma_N^2)\big),
$$

where $\sigma_q^2 = q! \sum_{k\in\mathbb{Z}} \rho(k)^q$, $q = d, \ldots, N$. We deduce that

$$
V_n = \frac{1}{\sqrt{n}} \sum_{q=d}^{N} a_q \sum_{k=1}^{n} H_q(X_k) \overset{\text{law}}{\to} \mathscr{N} \left(0, \sum_{q=d}^{N} a_q^2 q! \sum_{k\in\mathbb{Z}} \rho(k)^q \right),
$$

which is the desired conclusion in (ii) and conclude the proof of Theorem 9.

4.2 Quadratic variation of the fractional Brownian motion

In this section, we aim to illustrate Theorem 6 in a concrete situation. More precisely, we shall use Theorem 6 in order to derive an explicit bound for the second-order approximation of the quadratic variation of a fractional Brownian motion on $[0,1]$. Let $B^H = (B_t^H)_{t\geq 0}$ be a fractional Brownian motion with Hurst index $H \in (0,1)$. This means that B^H is a centered Gaussian process with covariance function given by

$$\mathbb{E}[B_t^H B_s^H] = \frac{1}{2}\big(t^{2H} + s^{2H} - |t-s|^{2H}\big), \quad s,t \geq 0.$$

It is easily checked that B^H is self-similar of index H and has stationary increments. Fractional Brownian motion has been successfully used in order to model a variety of natural phenomena coming from different fields, including hydrology, biology, medicine, economics or traffic networks. A natural question is thus the identification of the Hurst parameter from real data. To do so, it is popular and classical to use the quadratic variation (on, say, $[0,1]$), which is observable and given by

$$S_n = \sum_{k=0}^{n-1} (B^H_{(k+1)/n} - B^H_{k/n})^2, \quad n \geq 1.$$

For application of the fourth moment theorem in parameter estimation see also [4, 5, 10, 18].

Exercise 7. *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function so that $\mathbb{E}(f^2(N)) < +\infty$, where $N \sim \mathcal{N}(0,1)$. Prove that for fractional Brownian motion B^H of Hurst index $H \in (0,1)$, as $n \to \infty$, we have*

$$\frac{1}{n}\sum_{k=1}^{n} f(B_k^H - B_{k-1}^H) \xrightarrow{L^2} \mathbb{E}(f(N)).$$

Using, Exercise 7, One may infer that

$$n^{2H-1} S_n \overset{\text{proba}}{\to} 1 \quad \text{as } n \to \infty. \tag{37}$$

We deduce that the estimator $\widehat{H}_n$, defined as

$$\widehat{H}_n = \frac{1}{2} - \frac{\log S_n}{2 \log n},$$

satisfies $\widehat{H}_n \overset{\text{proba}}{\to} H$ as $n \to \infty$. To study the asymptotic normality, consider

$$F_n = \frac{n^{2H}}{\sigma_n} \sum_{k=0}^{n-1} \big[(B^H_{(k+1)/n} - B^H_{k/n})^2 - n^{-2H}\big] \overset{\text{(law)}}{=} \frac{1}{\sigma_n}\sum_{k=0}^{n-1}\big[(B^H_{k+1} - B^H_k)^2 - 1\big],$$

where $\sigma_n > 0$ is so that $\mathbb{E}[F_n^2] = 1$. We then have the following result.

Theorem 10. *Let $N \sim \mathcal{N}(0,1)$ and assume that $H \leq 3/4$. Then, $\lim_{n\to\infty} \sigma_n^2/n = 2\sum_{r\in\mathbb{Z}} \rho^2(r)$ if $H \in (0, \frac{3}{4})$, with $\rho : \mathbb{Z} \to \mathbb{R}$ given by*

$$\rho(r) = \frac{1}{2}\big(|r+1|^{2H} + |r-1|^{2H} - 2|r|^{2H}\big), \tag{38}$$

and $\lim_{n\to\infty} \sigma_n^2/(n \log n) = \frac{9}{16}$ if $H = \frac{3}{4}$. Moreover, there exists a constant $c_H > 0$ (depending only on H) such that, for every $n \geq 1$,

$$d_{TV}(F_n, N) \leq c_H \times \begin{cases} \frac{1}{\sqrt{n}} & \text{if } H \in (0, \frac{5}{8}) \\ \frac{(\log n)^{3/2}}{\sqrt{n}} & \text{if } H = \frac{5}{8} \\ n^{4H-3} & \text{if } H \in (\frac{5}{8}, \frac{3}{4}) \\ \frac{1}{\log n} & \text{if } H = \frac{3}{4} \end{cases}. \tag{39}$$

As an immediate consequence of Theorem 10, provided $H < 3/4$ we obtain that

$$\sqrt{n}\big(n^{2H-1}S_n - 1\big) \overset{\text{law}}{\to} \mathscr{N}\big(0, 2\sum_{r\in\mathbb{Z}}\rho^2(r)\big) \quad \text{as } n\to\infty, \tag{40}$$

implying in turn

$$\sqrt{n}\log n\big(\widehat{H}_n - H\big) \overset{\text{law}}{\to} \mathscr{N}\big(0, \frac{1}{2}\sum_{r\in\mathbb{Z}}\rho^2(r)\big) \quad \text{as } n\to\infty. \tag{41}$$

Indeed, we can write

$$\log x = x - 1 - \int_1^x du \int_1^u \frac{dv}{v^2} \quad \text{for all } x > 0,$$

so that (by considering $x \geq 1$ and $0 < x < 1$)

$$\big|\log x + 1 - x\big| \leq \frac{(x-1)^2}{2}\left\{1 + \frac{1}{x^2}\right\} \quad \text{for all } x > 0.$$

As a result,

$$\sqrt{n}\log n\big(\widehat{H}_n - H\big) = -\frac{\sqrt{n}}{2}\log(n^{2H-1}S_n) = -\frac{\sqrt{n}}{2}(n^{2H-1}S_n - 1) + R_n$$

with

$$|R_n| \leq \frac{\big(\sqrt{n}(n^{2H-1}S_n - 1)\big)^2}{4\sqrt{n}}\left\{1 + \frac{1}{(n^{2H-1}S_n)^2}\right\}.$$

Using (37) and (40), it is clear that $R_n \overset{\text{proba}}{\to} 0$ as $n\to\infty$ and then that (41) holds true. Now, let us go back to the proof of Theorem 10. We first need the following ancillary result.

Lemma 5. 1. *For any $r\in\mathbb{Z}$, let $\rho(r)$ be defined by (38). If $H \neq \frac{1}{2}$, one has $\rho(r) \sim H(2H-1)|r|^{2H-2}$ as $|r|\to\infty$. If $H = \frac{1}{2}$ and $|r| \geq 1$, one has $\rho(r) = 0$. Consequently, $\sum_{r\in\mathbb{Z}}\rho^2(r) < \infty$ if and only if $H < 3/4$.*

2. *For all $\alpha > -1$, we have $\sum_{r=1}^{n-1} r^\alpha \sim \frac{n^{\alpha+1}}{\alpha+1}$ as $n\to\infty$.*

Proof. 1. The sequence ρ is symmetric, that is, one has $\rho(n) = \rho(-n)$. When $r\to\infty$,

$$\rho(r) = H(2H-1)r^{2H-2} + o(r^{2H-2}).$$

Using the usual criterion for convergence of Riemann sums, we deduce that $\sum_{r\in\mathbb{Z}}\rho^2(r) < \infty$ if and only if $4H - 4 < -1$ if and only if $H < \frac{3}{4}$.
2. For $\alpha > -1$, we have:

$$\frac{1}{n}\sum_{r=1}^{n}\left(\frac{r}{n}\right)^\alpha \to \int_0^1 x^\alpha dx = \frac{1}{\alpha+1} \quad \text{as } n\to\infty.$$

We deduce that $\sum_{r=1}^{n} r^\alpha \sim \frac{n^{\alpha+1}}{\alpha+1}$ as $n\to\infty$. □

We are now in position to prove Theorem 10.

Proof of Theorem 10. Without loss of generality, we will rather use the second expression of F_n:

$$F_n = \frac{1}{\sigma_n}\sum_{k=0}^{n-1}\big[(B^H_{k+1} - B^H_k)^2 - 1\big].$$

Consider the linear span $\mathcal{H}$ of $(B^H_k)_{k\in\mathbb{N}}$, that is, $\mathcal{H}$ is the closed linear subspace of $L^2(\Omega)$ generated by $(B^H_k)_{k\in\mathbb{N}}$. It is a real separable Hilbert space and, consequently, there exists an isometry $\Phi : \mathcal{H} \to L^2(\mathbb{R}_+)$. For any $k\in\mathbb{N}$, set $e_k = \Phi(B^H_{k+1} - B^H_k)$; we then have, for all $k, l\in\mathbb{N}$,

$$\int_0^\infty e_k(s)e_l(s)ds = \mathbb{E}[(B^H_{k+1} - B^H_k)(B^H_{l+1} - B^H_l)] = \rho(k-l) \tag{42}$$

with ρ given by (38). Therefore,

$$\{B_{k+1}^H - B_k^H : k \in \mathbb{N}\} \overset{\text{law}}{=} \left\{ \int_0^\infty e_k(s) dB_s : k \in \mathbb{N} \right\} = \left\{ I_1^B(e_k) : k \in \mathbb{N} \right\},$$

where B is a Brownian motion and $I_p^B(\cdot)$, $p \geq 1$, stands for the pth multiple Wiener-Itô integral associated to B. As a consequence we can, without loss of generality, replace F_n by

$$F_n = \frac{1}{\sigma_n} \sum_{k=0}^{n-1} \left[\left(I_1^B(e_k) \right)^2 - 1 \right].$$

Now, using the product formula for multiple stochastic integrals, we deduce that

$$F_n = I_2^B(f_n), \quad \text{with } f_n = \frac{1}{\sigma_n} \sum_{k=0}^{n-1} e_k \otimes e_k.$$

By using the same arguments as in the proof of Theorem 9, we obtain the exact value of σ_n:

$$\sigma_n^2 = 2 \sum_{k,l=0}^{n-1} \rho^2(k-l) = 2 \sum_{|r|<n} (n - |r|) \rho^2(r).$$

Assume that $H < \frac{3}{4}$ and write

$$\frac{\sigma_n^2}{n} = 2 \sum_{r \in \mathbb{Z}} \rho^2(r) \left(1 - \frac{|r|}{n} \right) \mathbf{1}_{\{|r|<n\}}.$$

Since $\sum_{r \in \mathbb{Z}} \rho^2(r) < \infty$ by Lemma 5, we obtain by dominated convergence that, when $H < \frac{3}{4}$,

$$\lim_{n \to \infty} \frac{\sigma_n^2}{n} = 2 \sum_{r \in \mathbb{Z}} \rho^2(r). \tag{43}$$

Assume now that $H = \frac{3}{4}$. We then have $\rho^2(r) \sim \frac{9}{64|r|}$ as $|r| \to \infty$, implying in turn

$$n \sum_{|r|<n} \rho^2(r) \sim \frac{9n}{64} \sum_{0<|r|<n} \frac{1}{|r|} \sim \frac{9n \log n}{32}$$

and

$$\sum_{|r|<n} |r| \rho^2(r) \sim \frac{9}{64} \sum_{|r|<n} 1 \sim \frac{9n}{32}$$

as $n \to \infty$. Hence, when $H = \frac{3}{4}$,

$$\lim_{n \to \infty} \frac{\sigma_n^2}{n \log n} = \frac{9}{16}. \tag{44}$$

On the other hand, recall that the convolution of two sequences $\{u(n)\}_{n \in \mathbb{Z}}$ and $\{v(n)\}_{n \in \mathbb{Z}}$ is the sequence $u * v$ defined as $(u * v)(j) = \sum_{n \in \mathbb{Z}} u(n) v(j-n)$, and observe that $(u * v)(l-i) = \sum_{k \in \mathbb{Z}} u(k-l) v(k-i)$ whenever $u(n) = u(-n)$ and $v(n) = v(-n)$ for all $n \in \mathbb{Z}$. Set

$$\rho_n(k) = |\rho(k)| \mathbf{1}_{\{|k| \leq n-1\}}, \quad k \in \mathbb{Z}, \, n \geq 1.$$

We then have (using (15), and noticing that $f_n \otimes_1 f_n = f_n \widetilde{\otimes}_1 f_n$),

$$\begin{aligned}
& \mathbb{E}\left[\left(1 - \frac{1}{2} \| D[I_2^B(f_n)] \|_{L^2(\mathbb{R}_+)}^2 \right)^2 \right] \\
= \; & 8 \, \| f_n \otimes_1 f_n \|_{L^2(\mathbb{R}_+^2)}^2 = \frac{8}{\sigma_n^4} \sum_{i,j,k,l=0}^{n-1} \rho(k-l) \rho(i-j) \rho(k-i) \rho(l-j) \\
\leq \; & \frac{8}{\sigma_n^4} \sum_{i,l=0}^{n-1} \sum_{j,k \in \mathbb{Z}} \rho_n(k-l) \rho_n(i-j) \rho_n(k-i) \rho_n(l-j) \\
= \; & \frac{8}{\sigma_n^4} \sum_{i,l=0}^{n-1} (\rho_n * \rho_n)(l-i)^2 \leq \frac{8n}{\sigma_n^4} \sum_{k \in \mathbb{Z}} (\rho_n * \rho_n)(k)^2 = \frac{8n}{\sigma_n^4} \| \rho_n * \rho_n \|_{\ell^2(\mathbb{Z})}^2.
\end{aligned}$$

Recall Young's inequality: if $s, p, q \geq 1$ are such that $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{s}$, then

$$\|u * v\|_{\ell^s(\mathbb{Z})} \leq \|u\|_{\ell^p(\mathbb{Z})} \|v\|_{\ell^q(\mathbb{Z})}. \tag{45}$$

Let us apply (45) with $u = v = \rho_n$, $s = 2$ and $p = \frac{4}{3}$. We get $\|\rho_n * \rho_n\|^2_{\ell^2(\mathbb{Z})} \leq \|\rho_n\|^4_{\ell^{\frac{4}{3}}(\mathbb{Z})}$, so that

$$E\left[\left(1 - \frac{1}{2}\|D[I_2^B(f_n)]\|^2_{L^2(\mathbb{R}_+)}\right)^2\right] \leq \frac{8n}{\sigma_n^4}\left(\sum_{|k|<n} |\rho(k)|^{\frac{4}{3}}\right)^3. \tag{46}$$

Recall the asymptotic behavior of $\rho(k)$ as $|k| \to \infty$ from Lemma 5(1). Hence

$$\sum_{|k|<n} |\rho(k)|^{\frac{4}{3}} = \begin{cases} O(1) & \text{if } H \in (0, \frac{5}{8}) \\ O(\log n) & \text{if } H = \frac{5}{8} \\ O(n^{(8H-5)/3}) & \text{if } H \in (\frac{5}{8}, 1). \end{cases} \tag{47}$$

Assume first that $H < \frac{3}{4}$ and recall (43). This, together with (46) and (47), imply that

$$\begin{aligned} \mathbb{E}\left[\left|1 - \frac{1}{2}\|D[I_2^B(f_n)]\|^2_{L^2(\mathbb{R}_+)}\right|\right] &\leq \sqrt{\mathbb{E}\left[\left(1 - \frac{1}{2}\|D[I_2^B(f_n)]\|^2_{L^2(\mathbb{R}_+)}\right)^2\right]} \\ &\leq c_H \times \begin{cases} \frac{1}{\sqrt{n}} & \text{if } H \in (0, \frac{5}{8}) \\ \frac{(\log n)^{3/2}}{\sqrt{n}} & \text{if } H = \frac{5}{8} \\ n^{4H-3} & \text{if } H \in (\frac{5}{8}, \frac{3}{4}) \end{cases}. \end{aligned}$$

Therefore, the desired conclusion holds for $H \in (0, \frac{3}{4})$ by applying Theorem 6. Assume now that $H = \frac{3}{4}$ and recall (44). This, together with (46) and (47), imply that

$$\begin{aligned} \mathbb{E}\left[\left|1 - \frac{1}{2}\|D[I_2^B(f_n)]\|^2_{L^2(\mathbb{R}_+)}\right|\right] &\leq \sqrt{\mathbb{E}\left[\left(1 - \frac{1}{2}\|D[I_2^B(f_n)]\|^2_{L^2(\mathbb{R}_+)}\right)^2\right]} \\ &= O(1/\log n), \end{aligned}$$

and leads to the desired conclusion for $H = \frac{3}{4}$ as well. □

Part IV
Markov triplet approach

4.3 The setup: fourth moment theorem and new moments estimates

Recall that for any $X \in \mathcal{H}_p$ (the pth Wiener chaos), we have $\mathrm{L}(X) = -pX$, where L stands for Ornstein-Uhlenbeck operator. In the language of operator theory, we can write

$$\mathcal{H}_p = \mathbf{Ker}(\mathrm{L} + p\mathbf{Id}), \qquad \forall\, p \geq 1.$$

We also define the associated *carré-du-champ* operator Γ (for $X, Y \in \mathrm{dom}(\mathrm{L})$ such that $XY \in \mathrm{dom}(\mathrm{L})$) as:

$$\Gamma[X, Y] := \frac{1}{2}\left\{\mathrm{L}(XY) - Y\mathrm{L}X - X\mathrm{L}Y\right\}.$$

Note that according to Exercise 5, part (a), in the first sheet, we have in fact that $\Gamma[X, Y] = \langle DX, DY \rangle$. In particular, $\Gamma[X, X] = \Gamma[X] = \|DX\|^2$. In below, we summarize the fundamental properties of the Ornstein-Uhlenbeck operator L.

(a) *Diffusion:* For any text function $\phi\colon \mathbb{R} \to \mathbb{R}$, and any nice X in the domain, it holds that

$$\Gamma\left[\phi(X), X\right] = \phi'(X)\Gamma[X, X].$$

Or Equivalently,

$$\mathrm{L}\left[\phi(X)\right] = \phi'(X)\mathrm{L}[X] + \phi''(X)\Gamma[X, X]. \tag{48}$$

Note that, by taking $\phi = 1$, we get $\mathrm{L}[1] = 0$ which is called the *Markov property.*

(b) *Spectral decomposition:* The operator $-\mathrm{L}$ is diagonalizable on $L^2(\mathbb{P})$ with $\mathbf{sp}(-\mathrm{L}) = \mathbb{N}$, that is to say:

$$L^2(\mathbb{P}) = \bigoplus_{i=0}^{\infty} \mathbf{Ker}(\mathrm{L} + i\mathbf{Id}).$$

(c) *Spectral stability:* For any pair of eigenfunctions (X, Y) of the operator $-\mathrm{L}$ associated with eigenvalues (p_1, p_2),

$$XY \in \bigoplus_{i \le p_1 + p_2} \mathbf{Ker}\left(\mathrm{L} + i\mathbf{Id}\right). \tag{49}$$

Remark 14. (i) Property (a) is important regarding functional calculus. Using property (a) the *Malliavin integration by parts formula* reads to: for any X, Y in dom(L) and any text function ϕ we have that

$$\mathbb{E}\left[\phi'(X)\Gamma\left[X, Y\right]\right] = -\mathbb{E}\left[\phi(X)\mathrm{L}\left[Y\right]\right] = -\mathbb{E}\left[Y\mathrm{L}\left[\phi(X)\right]\right]. \tag{50}$$

(ii) Property (b) allows to use spectral theory. Actually, we stress that our results hold true under the weaker assumption that $\mathbf{sp}(-\mathrm{L}) = \{\lambda_i\,;\, i \in \mathbb{N}\} \subset \mathbb{R}_+$ is simply discrete so that $\lambda_{mi} \le m\lambda_i$ for all $m \in \mathbb{N}$ and $i \ge 1$. However, we stick to the assumption $\mathbf{sp}(-\mathrm{L}) = \mathbb{N}$ since it encompasses the most common cases (Wiener space and Laguerre space). The reader interested in relaxing this spectral assumption can consult [1] where the spectrum is only assumed to be discrete.

(iii) Property (c) is our main assumption, which will allow us to obtain fundamental spectral inequalities. A simple induction on (49) shows that, for any $X \in \mathbf{Ker}(\mathrm{L} + p\mathbf{Id})$ and any polynomial P of degree m, we have

$$P(X) \in \bigoplus_{i \le mp} \mathbf{Ker}\left(\mathrm{L} + i\mathbf{Id}\right). \tag{51}$$

Now, we are in the position to give another (very simple) proof of the fourth moment theorem relying only on assumptions (a)-(b)-(c).

Theorem 11. *Let $X \in \mathcal{H}_p$ be an eigenfunction of the operator* L *with eigenvalue $-p$, i.e.* $\mathrm{L}X = -pX$. *Then*

$$\mathbb{E}\left[(\Gamma[X] - p)^2\right] \le \frac{p^2}{6}\left\{\mathbb{E}[X^4] - 6\mathbb{E}[X^2] + 3\right\}.$$

Hence, for a sequence $\{X_n\}_{n\ge1} \in \boldsymbol{Ker}(\mathrm{L} + p\boldsymbol{Id})$ such that $\mathbb{E}[X_n^2] \to 1$, the convergence $\mathbb{E}[X_n^4] \to 3$ implies that X_n converges in distribution towards $N \sim \mathscr{N}(0, 1)$.

Proof. First note that $\Gamma[X] - p = (\mathrm{L} + p\mathbf{Id})\frac{1}{2}H_2(X)$, where $H_2(x) = x^2 - 1$ is the second Hermite polynomial. Also, for any $i \ge 1$, denote $J_i : L^2(\mathbb{P}) \to \mathbf{Ker}\left(\mathrm{L} + i\mathbf{Id}\right)$ the orthogonal projections onto the subspace $\mathbf{Ker}\left(\mathrm{L} + i\mathbf{Id}\right)$. Then

$$
\begin{aligned}
\mathbb{E}\left[(\Gamma[X]-p)^2\right] &= \mathbb{E}\left[(\mathrm{L}+p\mathbf{Id})\frac{1}{2}H_2(X)^2\right] \\
&= \frac{1}{4}\mathbb{E}\left[\mathrm{L}H_2(X)(\mathrm{L}+p\mathbf{Id})H_2(X)\right] + \frac{p}{4}\mathbb{E}\left[H_2(X)(\mathrm{L}+p\mathbf{Id})H_2(X)\right] \\
&= \frac{1}{4}\sum_{i=1}^{2p}(-i)(p-i)\mathbb{E}\left[J_i(H_2(X))^2\right] + \frac{p}{4}\mathbb{E}\left[H_2(X)(\mathrm{L}+p\mathbf{Id})H_2(X)\right] \\
&\leq \frac{p}{2}\mathbb{E}\left[H_2(X)(\mathrm{L}+p\mathbf{Id})\frac{1}{2}H_2(X)\right] \\
&= \frac{p}{2}\left\{\mathbb{E}\left[H_2(X)\Gamma[X]\right] - p\mathbb{E}\left[H_2(X)\right]\right\} \\
&= \frac{p}{2}\left\{\frac{1}{3}\mathbb{E}\left[\Gamma[H_3(X),X]\right] - p\mathbb{E}[X^2] + p\right\} \\
&= \frac{p}{2}\left\{\frac{1}{3}\mathbb{E}\left[-H_3(X)\mathrm{L}X\right] - p\mathbb{E}[X^2] + p\right\} \\
&= \frac{p}{2}\left\{\frac{p}{3}\mathbb{E}\left[XH_3(X)\right] - p\mathbb{E}[X^2] + p\right\} \qquad (H_3(x) = x^3 - 3x) \\
&= \frac{p^2}{6}\left\{\mathbb{E}[X^4] - 6\mathbb{E}[X^2] + 3\right\}.
\end{aligned}
$$

For the last part note that

$$\mathbb{E}[X^4] - 6\mathbb{E}[X^2] + 3 = \mathbb{E}[X^4] - 3\mathbb{E}[X^2]^2 + 3(\mathbb{E}[X^2]-1)^2.$$

□

Towards generalization, now we explore the idea used in the previous proof. Let $\mathbb{R}_k[T]$ stand for the ring of all polynomials of T of degree at most k over $\mathbb{R}$. Let $X \in \mathcal{H}_p$ be an eigenfunction of the generator L with eigenvalue $-p$, i.e. $-\mathrm{L}X = pX$. We consider the following map:

$$\mathcal{M}_k : \begin{cases} \mathbb{R}_k[T] \times \mathbb{R}_k[T] & \longrightarrow & \mathbb{R} \\ (P,Q) & \longmapsto & \mathbb{E}\big[Q(X)(\mathrm{L}+kp\mathbf{Id})P(X)\big]. \end{cases}$$

Remark 15. Notice that the mapping $\mathcal{M}_k$ strongly depends on the eigenfunction X. We also remark that thanks to Hyperconractivity property, $\mathcal{M}_k$ is well defined.

Theorem 12. *The mapping $\mathcal{M}_k$ is bilinear, symmetric and non-negative. Moreover its matrix representation over the canonical basis $\{1,T,T^2,\cdots,T^k\}$ is given by $p\mathbf{M}_k$ where*

$$\mathbf{M}_k = \Big((k - \frac{ij}{i+j-1})\mathbb{E}[X^{i+j}]\Big)_{0\leq i,j\leq k} \tag{52}$$

with the convention that $\frac{ij}{i+j-1} = 0$ for $(i,j) = (0,1)$ or $(1,0)$.

Proof. Expectation is a linear operator, so the bilinearity property follows. Symmetry proceeds from the symmetry of the diffusive generator L. To prove positivity of the matrix $\mathbf{M}_k$, using the fundamental assumption (51) we obtain that for any polynomial P of degree $\leq k$,

$$P(X) \in \bigoplus_{i\leq kp} \mathbf{Ker}(\mathrm{L}+i\mathbf{Id}).$$

Therefore,

$$
\begin{aligned}
\mathbb{E}\big[\big((\mathrm{L}+kp\mathbf{Id})P(X)\big)^2\big] &= \mathbb{E}[\mathrm{L}P(X)(\mathrm{L}+kp\mathbf{Id})P(X)] \\
&\quad + kp\mathbb{E}[P(X)(\mathrm{L}+kp\mathbf{Id})P(X)] \\
&= \sum_{i=0}^{kp}(-i)(kp-i)\mathbb{E}[J_i^2(P(X))] \\
&\quad + kp\mathbb{E}[P(X)(\mathrm{L}+kp\mathbf{Id})P(X)] \\
&\leq kp\ \mathcal{M}_k(P,P).
\end{aligned} \tag{53}
$$

Hence $\mathcal{M}_k$ is a positive form. To complete the proof, notice that the (i,j)-component of the matrix $\mathbf{M}_k$ is given by $\mathbb{E}\big[X^j(\mathrm{L}+kp\mathbf{Id})X^i\big]$. So, using the diffusive property of the generator L, we obtain

$$\begin{aligned} X^j(\mathrm{L}+kp\mathbf{Id})X^i &= i(i-1)X^{i+j-2}\Gamma(X) + p(k-i)X^{i+j} \\ &= \frac{i(i-1)}{i+j-1}\Gamma(X^{i+j-1},X) + p(k-i)X^{i+j}. \end{aligned}$$

Therefore

$$\begin{aligned} \mathcal{M}_k(X^i,X^j) &= \frac{i(i-1)}{i+j-1}\mathbb{E}[\Gamma(X^{i+j-1},X)] + p(k-i)\mathbb{E}[X^{i+j}] \\ &= p\frac{i(i-1)}{i+j-1}\mathbb{E}\left[X^{i+j}\right] + p(k-i)\mathbb{E}\left[X^{i+j}\right] \\ &= p\big(\frac{i(i-1)+(k-i)(i+j-1)}{i+j-1}\big)\mathbb{E}\left[X^{i+j}\right] \\ &= p\big(k-\frac{ij}{i+j-1}\big)\mathbb{E}\left[X^{i+j}\right]. \end{aligned}$$

Therefore, (i) all the eigenvalues of matrix $\mathbf{M}_k$ are non-negative and (ii) all the l^{th} leading principal minor of the matrix $\mathbf{M}_k$ are non-negative for $l \le k$. □

The moments matrix $\mathbf{M}_k$ can help one to give non-trivial moment inequalities, sometimes sharper than the existing estimates so far, involving the moments of the eigenfunctions of a generator L. Here is an application where we sharpen the standard fourth moment inequality $\mathbb{E}[X^4] \ge 3\mathbb{E}[X^2]^2$.

Theorem 13. *If X is a non-zero eigenfunction of generator* L*, then*

$$\frac{\mathbb{E}[X^4]}{3} - \mathbb{E}[X^2]^2 \ge \frac{\mathbb{E}[X^3]^2}{2\mathbb{E}[X^2]}. \tag{54}$$

Proof. The moments matrix $\mathbf{M}_2$ associated to X is given by

$$\mathbf{M}_2(X) = \begin{pmatrix} 2 & 0 & 2\mathbb{E}[X^2] \\ 0 & \mathbb{E}[X^2] & \mathbb{E}[X^3] \\ 2\mathbb{E}[X^2] & \mathbb{E}[X^3] & \frac{2}{3}\mathbb{E}[X^4] \end{pmatrix}. \tag{55}$$

Hence, we infer that

$$\det(\mathbf{M}_2) = 4\mathbb{E}[X^2]\big\{\frac{\mathbb{E}[X^4]}{3} - \mathbb{E}[X^2]^2\big\} - 2\mathbb{E}[X^3]^2 \ge 0,$$

which immediately implies (54). □

Remark 16. We stress that for a sequence $X_n \in \mathbf{Ker}(\mathrm{L}+p\mathbf{Id})$ for each $n \ge 1$, the convergence

$$\frac{\mathbb{E}[X_n^4]}{3} - \mathbb{E}[X_n^2]^2 - \frac{\mathbb{E}[X^3]^2}{2\mathbb{E}[X^2]} \to 0, \tag{56}$$

does not necessarily imply that X_n converges in distribution towards $\mathscr{N}(0,1)$. The reason is that the convergence (56) does not guarantees that $\mathbb{E}[X_n^3] \to 0$!.

The following proposition states a non-trivial inequality between the second, fourth and sixth moments of eigenfunctions of L.

Proposition 10. *If X is an eigenfunction of* L*, then*

$$\mathbb{E}[X^4]^2 \le \frac{3}{5}\mathbb{E}[X^6]\mathbb{E}[X^2]. \tag{57}$$

Notice that this inequality is an equality when the distribution of X is Gaussian.

Proof. The moments matrix $\mathbf{M}_3$ associated to X has the form

$$\mathbf{M}_3 = \begin{pmatrix} 3 & \star & 3\mathbb{E}[X^2] & \star \\ \star & 2\mathbb{E}[X^2] & \star & 2\mathbb{E}[X^4] \\ 3\mathbb{E}[X^2] & \star & \frac{5}{3}\mathbb{E}[X^4] & \star \\ \star & 2\mathbb{E}[X^4] & \star & \frac{6}{5}\mathbb{E}[X^6] \end{pmatrix}. \tag{58}$$

Since this matrix is positive, we have in particular

$$\begin{vmatrix} 2\mathbb{E}[X^2] & 2\mathbb{E}[X^4] \\ 2\mathbb{E}[X^4] & \frac{6}{5}\mathbb{E}[X^6] \end{vmatrix} \geq 0,$$

which gives the claimed inequality. □

Using Proposition 10, we can prove the following interesting sixth moment theorem.

Corollary 8. *A sequence $\{X_n\}_{n\geq 1}$ such that $X_n \in \boldsymbol{Ker}(\mathrm{L} + p\boldsymbol{Id})$ for each $n \geq 1$, converges in distribution toward the standard Gaussian law if and only if $\mathbb{E}[X_n^2] \to 1$ and $\mathbb{E}[X_n^6] \to 15$.*

Proof. By Proposition (10), for $X \in \mathbf{Ker}\,(\mathrm{L} + p\mathbf{Id})$, we have

$$\mathbb{E}[X^6] \geq \frac{5}{3}\frac{\mathbb{E}[X^4]^2}{\mathbb{E}[X^2]} \geq \frac{5}{3}\frac{(3\mathbb{E}[X^2]^2)^2}{\mathbb{E}[X^2]} = 15\mathbb{E}[X^2]^3.$$

Therefore, for the sequence $\{X_n\}_{n\geq 1}$ in $\mathbf{Ker}\,(\mathrm{L} + p\mathbf{Id})$, if $\mathbb{E}[X_n^2] \to 1$ and $\mathbb{E}[X_n^6] \to 15$, then from the previous chain of inequalities, we deduce that $\mathbb{E}[X_n^4] \to 3$. Hence, the sequence $\{X_n\}_{n\geq 1}$ converges in distribution toward $\mathcal{N}(0,1)$ according to the fourth moment theorem. □

Exercise 8. *Let $X \in \boldsymbol{Ker}(\mathrm{L} + p\boldsymbol{Id})$ for some $p \geq 2$ such that $\mathbb{E}[X^2] = 1$. Using Proposition 10, show that if p be an odd integer and $\kappa_4(X) \geq 3$, then $\kappa_6(X) \geq 0$. Recall that*

$$\kappa_6(X) = \mathbb{E}[X^6] - 15\mathbb{E}[X^2]\mathbb{E}[X^4] - 10\mathbb{E}[X^3]^2 + 30\mathbb{E}[X^2]^3.$$

[Hint: if p be odd then $\mathbb{E}[X^3] = 0$. See [13, Remark 8.4.5]. Any idea how to relax the condition $\kappa_4(X) \geq 3$?

Exercise 8 propels us to the following conjecture (known as Γ_2**- conjecture**). It is related to the non-Gaussian target distribution $N_1 \times N_2$ where N_1, $N_2 \sim \mathcal{N}(0,1)$ are independent, see the next section for more details.

Conjecture 1. *For any $X \in \boldsymbol{Ker}(\mathrm{L} + p\boldsymbol{Id})$, we have that $\kappa_6(X) \geq 0$. In fact, in a different context some computations suggest that*

$$\boldsymbol{Var}(\Gamma_2[X]) \leq C_p\,\kappa_6(X),$$

where $\Gamma_2[X] := \frac{1}{p}\Gamma[X, -\mathrm{L}^{-1}\Gamma[X]]$. This has to be compared with the fact that

$$\boldsymbol{Var}(\Gamma[X]) \leq C_p\kappa_4(X).$$

4.4 New central limit theorems: The even moment theorem and generalization

The main goal of this section is to proof the following substantial generalization of the fourth moment theorem called the even moment theorem allowing to replace the fourth moment with <u>any</u> even moments in the fourth moment theorem !

Theorem 14. *Let L be a Ornstein-Uhlenbeck generator (and so having properties (a)-(b)-(c)), $p \geq 1$ be an eigenvalue of $-\mathrm{L}$, and $\{X_n\}_{n\geq 1}$ a sequence of elements in $\boldsymbol{Ker}\,(\mathrm{L} + p\boldsymbol{Id})$ for all $n \geq 1$, such that $\lim_{n\to\infty} \mathbb{E}\left[X_n^2\right] = 1$. Then, for any integer $k \geq 2$, as $n \to \infty$, we have*

$$X_n \overset{law}{\to} \mathcal{N}(0,1) \quad \text{if and only if} \quad \mathbb{E}[X_n^{2k}] \to \mathbb{E}[N^{2k}] = (2k-1)!!. \tag{59}$$

To this end, let us go really further and explore more the positivity of the bilinear mapping $\mathcal{M}_k$ for every $k \geq 1$. We denote by $\{H_k\}_{k\geq 0}$ the family of Hermite polynomials defined (as before) by the recursive relation

$$H_0(x) = 1, \quad H_1(x) = x, \quad H_{k+1}(x) = xH_k(x) - kH_{k-1}(x). \tag{60}$$

Let us first recall that $H_k' = kH_{k-1}$. For any $k \geq 2$, we define the polynomial W_k as

$$W_k(x) = (2k-1)\Big(x \int_0^x H_k(t)H_{k-2}(t)\mathrm{d}t - H_k(x)H_{k-2}(x)\Big), \tag{61}$$

and the family $\mathscr{P}$ as

$$\mathscr{P} = \Big\{ P \;\Big|\; P(x) = \sum_{k=2}^{m} \alpha_k W_k(x);\; m \geq 2, \alpha_k \geq 0,\, 2 \leq k \leq m \Big\}. \tag{62}$$

The family $\mathscr{P}$ encodes interesting properties of central convergence which are the content of the two next lemmas.

Lemma 6. *Let* L *be a Ornstein-Uhlenbeck generator (and so having properties (a)-(b)-(c)). Let P be a polynomial belonging to $\mathscr{P}$. Then*

(1) If $N \sim \mathscr{N}(0,1)$, $\mathbb{E}[P(N)] = 0$.

(2) If X is an eigenvalue of L*, $\mathbb{E}[P(X)] \geq 0$.*

Proof. It is enough to prove that $E[W_k(X)] \geq 0$ and $E[W_k(N)] = 0$. Using the diffusive property (48), the fact that $-\mathrm{L}X = pX$ and the recursive property of Hermite polynomials, we obtain that

$$\begin{aligned}
(\mathrm{L} + kp\,\mathbf{Id})H_k(X) &= H_k''(X)\Gamma(X) + H_k'(X)\mathrm{L}(X) + kpH_k(X) \\
&= H_k''(X)\Gamma(X) - pXH_k'(X) + kpH_k(X) \\
&= H_k''(X)\big(\Gamma(X) - p\big) \\
&= k(k-1)H_{k-2}(X)\big(\Gamma(X) - p\big).
\end{aligned} \tag{63}$$

Therefore,

$$\begin{aligned}
\mathcal{M}_k(H_k) &= \mathbb{E}\big[H_k(X)\,(\mathrm{L} + kp\,\mathbf{Id})\,H_k(X)\big] \\
&= k(k-1)\mathbb{E}\big[H_k(X)H_{k-2}(X)\big(\Gamma(X) - p\big)\big].
\end{aligned} \tag{64}$$

Next, by the integration by parts formula (50), we have

$$\begin{aligned}
\mathbb{E}\big[H_k(X)H_{k-2}(X)\big(\Gamma(X) - p\big)\big] &= \mathbb{E}\left[\Gamma\left(\int_0^X H_k(t)H_{k-2}(t)\mathrm{d}t, X\right)\right] \\
&\quad - p\mathbb{E}\left[H_k(X)H_{k-2}(X)\right] \\
&= p\;\mathbb{E}\big[X\int_0^X H_k(t)H_{k-2}(t)\mathrm{d}t - H_k(X)H_{k-2}(X)\big] \\
&= \frac{p}{2k-1}\;\mathbb{E}\big[W_k(X)\big].
\end{aligned} \tag{65}$$

Hence,

$$\mathcal{M}_k(H_k) = \frac{pk(k-1)}{2k-1}\;\mathbb{E}\big[W_k(X)\big],$$

and the inequality $\mathbb{E}[W_k(X)] \geq 0$ follows from the positivity of the bilinear form $\mathcal{M}_k$. Finally, choosing $X = N$ be a standard Gaussian random variable living in the first Wiener chaos (i.e. $p = 1$) with variance 1, then $\Gamma(N) = p = 1$ and computation (65) shows that $\mathbb{E}[W_k(N)] = 0$ for every $k \geq 2$. Hence $\mathbb{E}[P(N)] = 0$ for every $P \in \mathscr{P}$. ☐

Lemma 7. *Let* L *be a Ornstein-Uhlenbeck generator (and so having properties (a)-(b)-(c)). Let* $p \geq 1$ *and* $\{X_n\}_{n\geq 1}$ *a sequence of elements in* $\boldsymbol{Ker}(\mathrm{L}+p\boldsymbol{Id})$ *for all* $n \geq 1$*. Let* $P = \sum_{k=2}^{m} \alpha_k W_k \in \mathscr{P}$ *such that* $\alpha_2 \neq 0$*. Then, as* $n \to \infty$*, we have*

$$X_n \overset{law}{\to} \mathscr{N}(0,1) \quad \text{if and only if} \quad \mathbb{E}[P(X_n)] \to \mathbb{E}[P(N)] = 0.$$

Proof. In virtue of Lemma 6,

$$\begin{aligned} \mathbb{E}[P(X_n)] &= \sum_{k=2}^{m} \alpha_k \mathbb{E}[W_k(X_n)] \\ &\geq \alpha_2 \mathbb{E}[W_2(X_n)] \\ &= \alpha_2 \Big(\mathbb{E}[X_n^4] - 6\mathbb{E}[X_n^2] + 3 \Big). \end{aligned}$$

This leads to

$$0 \leq \mathbb{E}[X_n^4] - 6\mathbb{E}[X_n^2] + 3 \leq \frac{1}{\alpha_2} \mathbb{E}[P(X_n)].$$

By assumption, $\mathbb{E}[P(X_n)] \to 0$, so $\mathbb{E}[X_n^4] - 6\mathbb{E}[X_n^2] + 3 \to 0$. On the other hand

$$\mathbb{E}[X_n^4] - 6\mathbb{E}[X_n^2] + 3 = \mathbb{E}[X_n^4] - 3\mathbb{E}[X_n^2]^2 + 3(\mathbb{E}[X_n^2] - 1)^2.$$

Thus, we obtain that $\mathbb{E}[X_n^2] \to 1$ and $\mathbb{E}[X_n^4] \to 3$, and we can use fourth moment Theorem to conclude. □

Proof. Proof of Theorem 14: Taking into account Lemmas 6 and 7, we are left to find a suitable polynomial $T_k \in \mathscr{P}$ of the form

$$T_k(x) = x^{2k} - \alpha_k x^2 + \beta_k, \qquad \alpha_k, \beta_k \in \mathbb{R}. \tag{66}$$

Stress that for such a polynomial, according to Lemma 6, the function $\phi_k : x \mapsto \mathbb{E}[T_k(xN)]$ must be positive and vanish at $x = 1$. Hence, we must have $\phi_k(1) = \phi_k'(1) = 0$. This leads us to the following system of equations

$$\begin{cases} (2k-1)!! - \alpha_k + \beta_k = 0, \\ 2k\,(2k-1)!! - 2\alpha_k = 0. \end{cases}$$

Therefore, the coefficients α_k and β_k are necessarily given by

$$\alpha_k = k\,(2k-1)!! \quad \text{and} \quad \beta_k = (k-1)\,(2k-1)!!.$$

It remains to check that the corresponding polynomial

$$T_k(x) = x^{2k} - k\,(2k-1)!!\, x^2 + (k-1)\,(2k-1)!! \in \mathscr{P}.$$

To this end, one needs to show that T_k can be expanded over the polynomials $\{W_k\}_{k\geq 2}$ with positive coefficients. This is the message of the next proposition. It turns out that the coefficient $\alpha_{2,k}$ in front of the polynomial W_2 is strictly positive, and so one can conclude the proof by using Lemma 7. In fact, the proof of the next Proposition is rather involved and the interested reader can consult [2, Appendix] for details.

Proposition 11. *Let* $k \geq 2$*, and* $T_k(x) = x^{2k} - k\,(2k-1)!!\, x^2 + (k-1)\,(2k-1)!!$*. Then*

$$T_k(x) = \sum_{i=2}^{k} \alpha_{i,k}\, W_i(x), \tag{67}$$

where

$$\alpha_{i,k} = \frac{(2k-1)!!}{2^{i-1}(2i-1)(i-2)!} \begin{pmatrix} k \\ i \end{pmatrix} \int_0^1 (1-u)^{-1/2} u^{i-2} \Big(1 - \frac{u}{2}\Big)^{k-i} du.$$

In particular, $T_k \in \mathscr{P}$ *and* $\alpha_{2,k} > 0$ *for all* $k \geq 1$*.*

□

A natural question: is it possible to replace the second moment with some other even moments in the even moment theorem 14? To be more precise, let $N \sim \mathscr{N}(0,1)$, and assume that for some pair (k,l) of positive integers $(k \neq l)$, we have $\mathbb{E}[X_n^{2k}] \to \mathbb{E}[N^{2k}]$ and $\mathbb{E}[X_n^{2l}] \to \mathbb{E}[N^{2l}]$. We want to know if this implies that X_n converges in distribution toward $\mathscr{N}(0,1)$. Our approach (as in the even moment theorem 14) would consist in deducing the existence of a non-trivial polynomial $T_{k,l} \in \mathscr{P}$ such that $\mathbb{E}\left[T_{k,l}(X_n)\right] \to 0$. Natural candidates are polynomials of the form

$$T_{k,l}(x) = x^{2l} + \alpha x^{2k} + \beta,$$

where $\alpha, \beta \in \mathbb{R}$. Using the same arguments as in the proof of Theorem 14, one can show that the condition $P \in \mathscr{P}$ entails necessarily that $\alpha = \frac{l(2l-1)!!}{k(2k-1)!!}$ and $\beta = \left(\frac{l}{k} - 1\right)(2k-1)!!$. Then, the question becomes: does the polynomial $T_{k,l}$ belong to family $\mathscr{P}$? We exhibit the decomposition of $T_{k,l}$ for each pair of integers in the set $\Theta = \{(2,3); (2,4); (2,5); (3,4); (3,5)\}$:

$$\begin{aligned}
T_{2,3}(x) &= x^6 - \tfrac{15}{2}x^4 + \tfrac{15}{2} &&= W_3(x) + \tfrac{\mathbf{5}}{\mathbf{2}} W_2(x)\\
T_{2,4}(x) &= x^8 - 70x^4 + 105 &&= W_4(x) + \tfrac{\mathbf{84}}{\mathbf{5}} W_3(x) + \mathbf{28} W_2(x)\\
T_{2,5}(x) &= x^{10} - \tfrac{1575}{2}x^4 + \tfrac{2835}{2} &&= W_5(x) + \tfrac{\mathbf{180}}{\mathbf{7}} W_4(x) + \mathbf{234} W_3(x) + \tfrac{\mathbf{585}}{\mathbf{2}} W_2(x)\\
T_{3,4}(x) &= x^8 - \tfrac{28}{3}x^6 + 35 &&= W_4(x) + \tfrac{\mathbf{112}}{\mathbf{5}} W_3(x) + \tfrac{\mathbf{14}}{\mathbf{3}} W_2(x)\\
T_{3,5}(x) &= x^{10} - 105x^6 + 630 &&= W_5(x) + \tfrac{\mathbf{180}}{\mathbf{7}} W_4(x) + \mathbf{129} W_3(x) + \mathbf{30} W_2(x).
\end{aligned}$$

The coefficients of each decomposition are positive, thus, for each pair $(k,l) \in \Theta$, the convergence of the $2k^{\text{th}}$ and $2l^{\text{th}}$ moments entails the central convergence. Unfortunately, it comes up that the polynomial $T_{4,5}$ does not belong to family $\mathscr{P}$:

$$T_{4,5}(x) = x^{10} - \frac{45}{4}x^8 + \frac{945}{4} = W_5(x) + \frac{405}{28}W_4(x) + W_3(x) \,\mathbf{-}\, \frac{\mathbf{45}}{\mathbf{2}} W_2(x).$$

Consequently, the convergence of the 8th and 10th moments for characterizing central convergence remains an open problem in the field!

4.5 Gaussian product conjecture

Here we aim to give another application of the Markov triplet approach in one of the outstanding conjecture in probability theory known as **Gaussian product conjecture**. See [11] and some other works by late Wenbo. Li (`http://wenbo.li.muchloved.com/`).

Conjecture 2. *Let $(X_1, \cdots, X_d)$ be a center Gaussian vector. The Gaussian product conjecture states that for all $r \geq 1$:*

$$\mathbb{E}\left(X_1^{2r} \times \cdots \times X_d^{2r}\right) \geq \mathbb{E}(X_1^{2r}) \times \cdots \times \mathbb{E}(X_d^{2r}).$$

Remark 17. (i) Case $r = 1$ solved by Frenkel [9] using exclusively tools taken from linear algebra such as Hafnians, Pfaffians.

(ii) Case $r = 2$ remains unsolved but supported by computer simulations.

(iii) The case of complex Gaussian solved by Arias de Reyna (1998).

Now, we are going to give a new proof of Gaussian product conjecture for the case $r = 2$, and at the same time generalize it to random vectors having multiple integrals in entries possibly of different orders.

Theorem 15. *[12] Let $X = (X_1, \cdots, X_d) = (I_{p_1}(f_1), \cdots, I_{p_d}(f_d))$, and $p_i \geq 1$ for all $1 \leq i \leq d$. Then*

$$\mathbb{E}\left(X_1^2 \times \cdots \times X_d^2\right) \geq \mathbb{E}(X_1^2) \times \cdots \times \mathbb{E}(X_d^2).$$

Proof. Since $X_1, \cdots, F_d$ are eigenfunctions of L with eigenvalues $p_1, \cdots, p_d$, therefore

$$X_1 \times \cdots \times X_d \in \bigoplus_{i \leq p_1 + \cdots + p_d} \mathbf{Ker}(\mathrm{L} + i\mathbf{Id}).$$

Hence,

$$\mathbb{E}\Big(X_1 \times \cdots \times X_d (\mathrm{L} + (p_1 + \cdots + p_d)\mathbf{Id})[X_1 \times \cdots \times X_d]\Big) \geq 0.$$

Using the relation between operators L and Γ one has

$$(\mathrm{L} + (p_1 + \cdots + p_d)\mathbf{Id})[X_1 \times \cdots \times X_d] = \sum_{i \neq j} \left(\prod_{k \neq i,j} X_k \right) \Gamma[X_i, X_j].$$

Therefore,

$$\mathbb{E}\left(\prod_i X_i \sum_{i \neq j} \left(\prod_{k \neq i,j} X_k \right) \Gamma[X_i, X_j] \right) \geq 0.$$

Equivalently, by Malliavin integration by parts, we have

$$\sum_{i=1}^d \mathbb{E}\Big(\mathrm{L}[X_i^2] \prod_{j \neq i} X_j^2\Big) \leq 0.$$

Now, we proceed with induction on the orders of involved multiple integrals. Using again relation between L and Γ: $\mathrm{L}[X_i^2] = 2\Gamma[X_i, X_i] - 2p_i X_i^2$, we infer that

$$\begin{aligned}(p_1 + \cdots + p_d)\mathbb{E}\left(X_1^2 \times \cdots \times X_d^2\right) = \sum_{i=1}^d \mathbb{E}\Big(\Gamma[X_i, X_i] \prod_{j \neq i} X_j^2\Big) \\ - \frac{1}{2} \sum_{i=1}^d \mathbb{E}\big(\mathrm{L}[X_i^2] \prod_{j \neq i} X_j^2\big).\end{aligned}$$

Therefore,

$$(p_1 + \cdots + p_d)\mathbb{E}\left(X_1^2 \times \cdots \times X_d^2\right) \geq \sum_{i=1}^d \mathbb{E}\Big(\Gamma[X_i, X_i] \prod_{j \neq i} X_j^2\Big).$$

Now, using the fact that $\Gamma[X_i, X_i] = \|D_t X_i\|^2_{L^2(\mu)} = \|p_i I_{p_i - 1}(f_i(t, \bullet))\|^2$, stochastic Fubini Theorem, and induction, we obtain that

$$\begin{aligned}\sum_{i=1}^d \mathbb{E}\Big(\Gamma[X_i, X_i] \prod_{j \neq i} X_j^2\Big) &= \sum_{i=1}^d \mathbb{E}\Big(\int_A p_i^2\, I^2_{p_i - 1}(f_i(t, \bullet))\mu(\mathrm{d}t) \times \prod_{j \neq i} X_j^2\Big), \\ &= \sum_{i=1}^d \int_A p_i^2\, \mathbb{E}\Big(I^2_{p_i - 1}(f_i(t, \bullet)) \times \prod_{j \neq i} X_j^2\Big)\mu(\mathrm{d}t) \\ &\geq \sum_{i=1}^d \int_A p_i^2\, \mathbb{E}(I^2_{p_i - 1}(f_i(t, \bullet)))\mu(\mathrm{d}t) \times \prod_{j \neq i} \mathbb{E}(X_j^2), \qquad \text{(by induction)} \\ &= \sum_{i=1}^d \mathbb{E}\big(p_i^2 \int_A I^2_{p_i - 1}(f_i(t, \bullet))\mu(\mathrm{d}t)\big) \times \prod_{j \neq i} \mathbb{E}(X_j^2) \\ &= \sum_{i=1}^d \mathbb{E}[\Gamma(X_i, X_i)] \times \prod_{j \neq i} \mathbb{E}(X_j^2) \\ &= \sum_{i=1}^d p_i \mathbb{E}(X_i^2) \times \prod_{j \neq i} \mathbb{E}(X_j^2) = (p_1 + \cdots + p_d)\mathbb{E}(X_1^2) \times \cdots \times \mathbb{E}(X_d^2).\end{aligned}$$

□

5 Non-Gaussian target distributions

5.1 Cumulant and Malliavin operators

It will turn out that cumulants play more transparent role for convergence toward non-Gaussian target distributions. Main references for this section are [3, 8, 16]. Hence, our first aim is to provide an explicit representation of cumulants in terms of Malliavin operators. To this end, it is convenient to introduce the following Malliavin objects that naturally appear using repetition of Malliavin integration by part formula.

Definition 6. Let $F \in \mathbb{D}^\infty$, i.e. infinitely times Malliavin differentiable. The sequence of random variables $\{\Gamma_i(F)\}_{i\geq 0} \subset \mathbb{D}^\infty$ is recursively defined as follows. Set $\Gamma_0(F) = F$ and, for every $i \geq 1$,

$$\Gamma_i(F) = \langle DF, -DL^{-1}\Gamma_{i-1}(F)\rangle.$$

For instance, one has that $\Gamma_1(F) = \langle DF, -DL^{-1}F\rangle$. The following statement provides an explicit expression for $\Gamma_s(F)$, $s \geq 1$, when F has the form of a multiple integral.

Proposition 12 (See e.g. Chapter 8 in [13])**.** *Let $p \geq 2$, and assume that $F = I_p(f)$. Then, for any $i \geq 1$, we have*

$$\Gamma_i(F) = \sum_{r_1=1}^{p} \cdots \sum_{r_i=1}^{[ip-2r_1-\ldots-2r_{i-1}]\wedge p} c_p(r_1,\ldots,r_i)\mathbf{1}_{\{r_1<p\}}\ldots\mathbf{1}_{\{r_1+\ldots+r_{i-1}<\frac{ip}{2}\}} \times I_{(i+1)p-2r_1-\ldots-2r_i}\big((\ldots(f\widetilde{\otimes}_{r_1}f)\widetilde{\otimes}_{r_2}f)\ldots f)\widetilde{\otimes}_{r_i}f\big),$$

where the constants $c_p(r_1,\ldots,r_{i-2})$ are recursively defined as follows:

$$c_p(r) = p(r-1)!\binom{p-1}{r-1}^2,$$

and, for $a \geq 2$,

$$c_p(r_1,\ldots,r_a) = p(r_a-1)!\binom{ap-2r_1-\ldots-2r_{a-1}-1}{r_a-1}\binom{p-1}{r_a-1}c_p(r_1,\ldots,r_{a-1}).$$

Example 4. *Let $F = I_p(f)$ for some $p \geq 2$. Then*

$$\Gamma_1(F) = p\sum_{r=1}^{p}(r-1)!\binom{p-1}{r-1}^2 I_{2p-2r}(f\widetilde{\otimes}_r f)$$

$$\Gamma_2(F) = \sum_{r=1}^{p}\sum_{s=1}^{(2p-2r)\wedge p} p^2(r-1)!(s-1)!\binom{p-1}{r-1}^2\binom{p-1}{s-1}\binom{2p-2r-1}{s-1}I_{3p-2r-2s}\big((f\widetilde{\otimes}_r f)\widetilde{\otimes}_s f\big)$$

The following statement explicitly connects the expectation of the random variables $\Gamma_i(F)$ to the cumulants of F.

Proposition 13 (See again Chapter 8 in [13])**.** *Let $F \in \mathbb{D}^\infty$. Then F has finite moments of every order, and the following relation holds for every $i \geq 0$:*

$$\kappa_{i+1}(F) = i!\mathbb{E}[\Gamma_i(F)]. \tag{68}$$

Lemma 8. *Let $X \in \mathbb{D}^\infty$. Then, the relation*

$$\mathbb{E}(\phi^{(k)}(X)\Gamma_r(X)) = \mathbb{E}(X\phi^{(k-r)}(X)) - \sum_{s=1}^{r}\mathbb{E}(\phi^{(k-s)}(X))\mathbb{E}(\Gamma_{r-s}(X)) \tag{69}$$

holds for every k-times continuously differentiable mapping $\phi : \mathbb{R} \to \mathbb{R}$.

Proof. Use Malliavin integration by part formula (many times) together with Proposition 13. □

5.2 Some relevant properties of the second Wiener chaos

In this subsection, we gather together some properties of the elements of the second Wiener chaos of the Gaussian random measure G with control measure μ; recall that these are random variables having the general form $F = I_2(f)$, with $f \in L^2(\mu^2)$. Notice that, if $f = h \otimes h$, where $h \in L^2(\mu)$ is such that $\|h\| = 1$, then using the product formula, one has $I_2(f) = I_1(h)^2 - 1 \stackrel{\text{law}}{=} N^2 - 1$, where $N \sim \mathscr{N}(0,1)$. To any symmetric kernel $f \in L^2(\mu^2)$, we associate the following *Hilbert-Schmidt* operator

$$A_f : L^2(\mu) \mapsto L^2(\mu); \quad g \mapsto f \otimes_1 g.$$

It is also convenient to introduce the sequence of auxiliary kernels

$$\left\{ f \otimes_1^{(p)} f : p \geq 1 \right\} \subset L^2_{\text{sym}}(\mu^2) \tag{70}$$

defined as follows: $f \otimes_1^{(1)} f = f$, and, for $p \geq 2$,

$$f \otimes_1^{(p)} f = \left(f \otimes_1^{(p-1)} f \right) \otimes_1 f. \tag{71}$$

In particular, $f \otimes_1^{(2)} f = f \otimes_1 f$. Finally, we write $\{\alpha_{f,j}\}_{j\geq 1}$ and $\{e_{f,j}\}_{j\geq 1}$, respectively, to indicate the (not necessarily distinct) eigenvalues of A_f and the corresponding eigenvectors.

Proposition 14 (See e.g. Section 2.7.4 in [13]). *Fix $F = I_2(f)$ with symmetric kernel $f \in L^2(\mu^2)$.*

1. *The following equality holds: $F = \sum_{j\geq 1} \alpha_{f,j}\big(N_j^2 - 1\big)$, where $\{N_j\}_{j\geq 1}$ is a sequence of i.i.d. $\mathscr{N}(0,1)$ random variables, and the series converges in L^2 and almost surely.*

2. *For any $i \geq 2$,*

$$\kappa_i(F) = 2^{i-1}(i-1)! \sum_{j\geq 1} \alpha_{f,j}^i = 2^{i-1}(i-1)! \times \langle f \otimes_1^{(i-1)} f, f \rangle_{L^2(\mu^2)}.$$

3. *The law of the random variable F is completely determined by its moments or equivalently by its cumulants.*

We now fix a symmetric kernel f_∞ such that the its corresponding Hilbert-Schmidt operator A_{f_∞} (see Section 5.2) has a finite number of non-zero eigenvalues, that we denote by $\{\alpha_i\}_{i=1}^k$. To simplify the discussion, we assume that the eigenvalues are all distinct. We want to study convergence in distribution towards the random variable

$$F_\infty := I_2(f_\infty) = \sum_{i=1}^k \alpha_i \left(N_i^2 - 1 \right), \tag{72}$$

where $\{N_i\}_{i=1}^k$ is the family of i.i.d. $\mathscr{N}(0,1)$ random variables appearing at Point 1 of Proposition 14. Following Nourdin and Poly [16], we define the two crucial polynomials P and Q as follows:

$$Q(x) = \big(P(x)\big)^2 = \left(x \prod_{i=1}^k (x - \alpha_i) \right)^2. \tag{73}$$

Note that, by definition, the roots of Q and P correspond with the set $\{0, \alpha_1, ..., \alpha_k\}$. The following lemma reveals the important role of the polynomials P and Q.

Lemma 9. *Let $F = I_2(f)$ be a generic element of the second Wiener chaos, and write $\{\alpha_{f,j}\}_{j\geq 1}$ for the set of the eigenvalues of the associated Hilbert-Schmidt operator A_f we have*

$$\sum_{r=2}^{deg(Q)} \frac{Q^{(r)}(0)}{r!} \frac{\kappa_r(F)}{2^{r-1}(r-1)!}$$
$$= \sum_{j\geq 1} Q(\alpha_{f,j}) \tag{74}$$
$$= \left\| \sum_{r=1}^{deg(P)} \frac{P^{(r)}(0)}{r!} f \otimes_1^{(r)} f \right\|_{L^2(\mu^2)}^2 \tag{75}$$
$$= \frac{1}{2}\mathbb{E}\left(\sum_{r=1}^{deg(P)} \frac{P^{(r)}(0)}{r!\ 2^{r-1}} \Big(\Gamma_{r-1}(F) - \mathbb{E}(\Gamma_{r-1}(F)) \Big) \right)^2, \tag{76}$$

where the operators $\Gamma_r(\cdot)$ have been introduced in Definition 6. In particular, for the target random variable F_∞ introduced at (72) *one has that*

$$0 = \sum_{r=2}^{deg(Q)} \frac{Q^{(r)}(0)}{r!} \frac{\kappa_r(F_\infty)}{2^{r-1}(r-1)!}$$
$$= \frac{1}{2}\mathbb{E}\left(\sum_{r=1}^{deg(P)} \frac{P^{(r)}(0)}{r!\ 2^{r-1}} \Big(\Gamma_{r-1}(F_\infty) - \mathbb{E}(\Gamma_{r-1}(F_\infty)) \Big) \right)^2. \tag{77}$$

Proof. In view of the second equality at Point 2 of Proposition 14, one has that $\frac{\kappa_r(F)}{2^{r-1}(r-1)!} = \sum_{j\geq 1} \alpha_{f,j}^r$, from which we deduce immediately (74). To prove (75), observe that Point 1 of Proposition 14, together with the product formula, implies that the kernel f admits a representation of the type $f = \sum_{j\geq 1} \alpha_{f,j} \eta_j \otimes \eta_j$, where $\{\eta_j\}$ is some orthonormal system in $L^2(\mu)$ (recall that $L^2(\mu)$ is a separable Hilbert space). It follows that, for $r \geq 1$, one has the representation $f \otimes_1^{(r)} f = \sum_{j\geq 1} \alpha_{f,j}^r \eta_j \otimes \eta_j$, and therefore

$$\sum_{r=1}^{\deg(P)} \frac{P^{(r)}(0)}{r!} f \otimes_1^{(r)} f = \sum_{j\geq 1} \eta_j \otimes \eta_j \sum_{r=1}^{\deg(P)} \frac{P^{(r)}(0)}{r!} \alpha_{f,j}^r.$$

Taking norms on both sides of the previous relation and exploiting the orthonormality of the η_j yields (75). Finally, in order to show (76), it is clearly enough to prove that, for any $r \geq 1$,

$$I_2(f \otimes_1^{(r)} f) = \frac{1}{2^{r-1}} \big\{ \Gamma_{r-1}(F) - \mathbb{E}(\Gamma_{r-1}(F)) \big\}. \tag{78}$$

We proceed by induction on r. It is clear for $r = 1$, because $\Gamma_0(F) = F$ and $\mathbb{E}(F) = 0$. Take $r \geq 2$ and assume that (78) holds true. Notice that, by definition of $\Gamma_r(F)$ and the induction assumption, one has

$$\begin{aligned}
&\Gamma_r(F) \\
&= \langle DF, -DL^{-1}\Gamma_{r-1}(F) \rangle_{L^2(\mu)} = \Big\langle 2I_1(f(t,.)), 2^{r-1} I_1(f \otimes_1^{(r)} f(t,.)) \Big\rangle_{L^2(\mu)} \\
&= 2^r \int_A \Big\{ \langle f(t,.), f \otimes_1^{(r)} f(t,.) \rangle_{L^2(\mu)} + I_2\big(f(t,.) \otimes (f \otimes_1^{(r)} f)(t,.)\big) \Big\} \mathrm{d}\mu(t) \\
&= 2^r \langle f, f \otimes_1^{(r)} f \rangle_{L^2(\mu^2)} + 2^r I_2(f \otimes_1^{(r+1)} f),
\end{aligned}$$

where we have used a standard stochastic Fubini Theorem. This proves that (78) is verified for every $r \geq 1$. The last assertion in the statement follows from (74), as well as the fact that the eigenvalues α_i are all roots of Q. □

Proposition 15. *Let the polynomial P be defined as in (73) and consider again the random variable $F_\infty = I_2(f_\infty)$ defined in (72). Let F be a centered random variable living in a finite sum of Wiener chaoses, i.e. $F \in \bigoplus_{i=1}^{M} \mathcal{H}_i$. Moreover, assume that*

(i) $\kappa_r(F) = \kappa_r(F_\infty)$, *for all $2 \le r \le k+1 = deg(P)$, and*

(ii)

$$\mathbb{E}\left(\sum_{r=1}^{k+1} \frac{P^{(r)}(0)}{r!\; 2^{r-1}} \Big(\Gamma_{r-1}(F) - \mathbb{E}(\Gamma_{r-1}(F)) \Big) \right)^2 = 0.$$

Then, $F \overset{\text{law}}{=} F_\infty$, and $F \in \mathcal{H}_2$.

Proof. Let ϕ be a smooth function. Using the integration by parts formula (Lemma **??**), Assumption (**ii**) in the statement and Proposition 13, we obtain

$$\begin{aligned}\mathbb{E}\big(F\phi(F)\big) &= \sum_{r=0}^{k-1} \frac{\kappa_{r+1}(F)}{r!} \mathbb{E}(\phi^{(r)}(F)) + \mathbb{E}(\phi^{(k)}(F)\Gamma_k(F)) \\ &= \sum_{r=0}^{k-1} \frac{\kappa_{r+1}(F)}{r!} \mathbb{E}(\phi^{(r)}(F)) + \frac{\kappa_{k+1}(F)}{k!} \mathbb{E}(\phi^{(k)}(F)) \\ &+ \sum_{r=1}^{k} \frac{2^{k-r+1}\kappa_r(F)}{(r-1)!r!} P^{(r)}(0)\mathbb{E}(\phi^{(k)}(F)) \\ &- \sum_{r=1}^{k} \frac{2^{k-r+1}}{r!} P^{(r)}(0)\mathbb{E}(\phi^{(k)}(F)\Gamma_{r-1}(F))\end{aligned} \tag{79}$$

On the other hand, using (69) we obtain that

$$\begin{aligned}\mathbb{E}(\phi^{(k)}(F)\Gamma_{r-1}(F)) &= \mathbb{E}(F\phi^{(k-(r-1))}(F)) \\ &- \sum_{s=1}^{r-1} \mathbb{E}(\phi^{(k-s)}(F))\mathbb{E}(\Gamma_{r-1-s}(F)).\end{aligned} \tag{80}$$

Using the relation $\mathbb{E}(\Gamma_{r-1-s}(F)) = \kappa_{r-s}(F)/(r-s-1)!$, and therefore deduce that, for every smooth test function ϕ

$$\begin{aligned}\mathbb{E}\big(F\phi(F)\big) \;\; = \;\; & \sum_{r=0}^{k-1} \frac{\kappa_{r+1}(F)}{r!} \mathbb{E}(\phi^{(r)}(F)) + \frac{\kappa_{k+1}(F)}{k!} \mathbb{E}(\phi^{(k)}(F)) \\ & + \sum_{r=1}^{k} \frac{2^{k-r+1}\kappa_r(F)}{(r-1)!r!} P^{(r)}(0)\mathbb{E}(\phi^{(k)}(F)) \\ & - \sum_{r=1}^{k} \frac{2^{k-r+1}}{r!} P^{(r)}(0)\mathbb{E}[F\phi^{(k-(r-1))}(F)] \\ & + \sum_{r=1}^{k} \frac{2^{k-r+1}}{r!} P^{(r)}(0) \sum_{s=1}^{r-1} \mathbb{E}[\phi^{(k-s)}(F)] \frac{\kappa_{r-s}(F)}{(r-s-1)!}.\end{aligned}$$

Considering the test function $\phi(x) = x^n$ with $n > k$, we infer that $\mathbb{E}(F^{n+1})$ can be expressed in a recursive way in terms of the quantities

$$\mathbb{E}(F^n), \mathbb{E}(F^{n-1}), \cdots, \mathbb{E}(F^{n-k}),\ \kappa_2(F), \cdots, \kappa_{k+1}(F)$$

and $P^{(1)}(0), \cdots, P^{(k)}(0)$. Using Assumption (**i**) in the statement together with last assertion in Lemma 9, we see that the moments of the random variable F_∞ also satisfy the same recursive relation. These facts immediately imply that

$$\mathbb{E}\left(F^n\right) = \mathbb{E}\left(F_\infty^n\right), \quad n \ge 1,$$

and the claim follows at once from Point 3 in Proposition 14. To prove that, in fact, $F \in \mathcal{H}_2$, we are going to use the two following remarkable results:

(i) Let $F \in \mathcal{H}_p$ for some $p \geq 2$. Then, there exist two constants a and b and also $x_0 \in \mathbb{R}_+$ such that

$$\exp\{-bx^{2/p}\} \leq \mathbb{P}(|F| > x) \leq \exp\{-ax^{2/p}\} \qquad \forall\, x \geq x_0.$$

(ii) Let $\{F_n\}_{n\geq 1}$ be a sequence of elements in the second Wiener chaos so that F_n converges in distribution toward F_∞. Then $F_\infty \stackrel{\text{law}}{=} I_1(f_{1,\infty}) + I_2(f_{2,\infty})$, and moreover $I_1(f_{1,\infty})$ and $I_1(f_{2,\infty})$ are independent.

Coming back to our proof, now we assume that M is the smallest natural number such that $F \in \bigoplus_{i=1}^{M} \mathcal{H}_i$. Hence $F \notin \bigoplus_{i=1}^{M-1} \mathcal{H}_i$. Therefore, by applying the fact (i) to F, F_∞ and the fact that $F \stackrel{\text{law}}{=} F_\infty$, we deduce that $M = 2$. Let assume that $F = I_1(g)+I_2(h)$ for some $g \in L^2(\mu)$ and $h \in L^2(\mu^2)$. Considering the trivial sequence $\{F_n\}_{n\geq 1}$ such that $F_n = F_\infty$, $n \geq 1$, using the fact that $F \stackrel{\text{law}}{=} F_\infty$ and applying the fact (ii) above, we deduce that $I_1(g)$ is independent of $I_2(h)$. Let $\{\lambda_{f_\infty,k}\}_{k\geq 1}$ and $\{\lambda_{h,k}\}_{k\geq 1}$ denote the eigenvalues corresponding to the Hilbert-Schmidt operator A_{f_∞} and A_h associated with the kernels f_∞ and h respectively (see Section 5.2). Exploiting the independence of $I_1(g)$ and $I_2(h)$ and Point 2 in Proposition 14, we infer that

$$\sum_{k\in\mathbb{N}} \lambda_{f_\infty,k}^{3p} = \sum_{k\in\mathbb{N}} \lambda_{h,k}^{3p} \quad \forall\, p \geq 1.$$

As a result, for some permutation π on N we have $\lambda_{\infty,k} = \lambda_{h,\pi(k)}$ for $k \geq 1$, which in turn implies

$$\sum_{k\in\mathbb{N}} \lambda_{f_\infty,k}^{2} = \sum_{k\in\mathbb{N}} \lambda_{h,k}^{2}. \tag{81}$$

On the other hand, from $F = I_1(g) + I_2(h) \stackrel{\text{law}}{=} F_\infty$, and computing the second cumulant of both sides, one can easily deduce that if $\kappa_2(I_1(g)) = \mathbb{E}(I_1(g))^2 = \|g\|^2 \neq 0$, then the equality (81) cannot hold. It follows that $I_1(g) = 0$, and therefore $F \in \mathcal{H}_2$. □

The next theorem is the main finding of this section. Recall that the random variable F_∞ has been defined in formula (72).

Theorem 16. *Let $\{F_n\}_{n\geq 1}$ be a sequence of random variables such that each F_n lives in a finite sum of chaoses, i.e. $F_n \in \bigoplus_{i=1}^{M} \mathcal{H}_i$ for $n \geq 1$ and some $M \geq 2$ (not depending on n). Consider the following three asymptotic relations, as $n \to \infty$:*

(i) *F_n converges in distribution toward F_∞.*

(ii) *The following relations 1.-2. are in order:*

1. *$\kappa_r(F_n) \to \kappa_r(F_\infty)$, for all $2 \leq r \leq k+1 = deg(P)$, and*
2. *$\mathbb{E}\Big(\sum_{r=1}^{k+1} \frac{P^{(r)}(0)}{r!2^{r-1}} \Big(\Gamma_{r-1}(F_n) - \mathbb{E}(\Gamma_{r-1}(F_n))\Big)\Big)^2 \longrightarrow 0.$*

Then, one has the implications **(ii)** $\to$ **(i)**.

We will now illustrate the our main findings (Theorem 16) of this section by considering the case of a target random variable of the type $F_\infty = I_2(f_\infty)$, where the Hilbert-Schmidt operator A_{f_∞} associated the kernel f_∞ has only two non-zero eigenvalues $\alpha_1 = -\alpha_2 = 1/2$, thus implying that

$$F_\infty = \frac{1}{2}\left(N_1^2 - 1\right) - \frac{1}{2}\left(N_2^2 - 1\right) \stackrel{\text{law}}{=} N_1 \times N_2 \tag{82}$$

where N_1 and N_2 are independent $\mathscr{N}(0,1)$. Note that the associated Polynomial P to such target distribution is given by $P(x) = x(x-1/2)(x+1/2)$. Hence, $P'(0) = -1/4$, $P''(0) = 0$ and $P'''(0) = 3!$.

Corollary 9. *Let $\{F_n\}_{n\geq 1}$ be a sequence of random variables such that each F_n lives in a finite sum of chaoses, i.e. $F_n \in \bigoplus_{i=1}^{M} \mathcal{H}_i$ for $n \geq 1$ and some $M \geq 2$ (not depending on n). Assume that the following asymptotic relations are in order:*

1. $\kappa_r(F_n) \to \kappa_r(F_\infty), \quad r = 2, 3 = deg(P)$, *and*
2. $\mathbb{E}\Big(\Gamma_2(F_n) - \mathbb{E}(\Gamma_2(F_n)) - F_n\Big)^2 \longrightarrow 0.$

Then F_n converges in distribution toward the target random variable $N_1 \times N_2$.

Lemma 10. *Let $F \in \mathbb{D}^\infty$, and $\mathbb{E}(F) = 0$. Then for every $s \geq 1$,*

$$\kappa_{s+2}(F) = 1/2 \times (s+1)!\mathbb{E}\Big(F^2\big(\Gamma_{s-1}(F) - \mathbb{E}\Gamma_{s-1}(F)\big)\Big).$$

In particular,

$$\begin{aligned}
\kappa_4(F) &= \frac{3!}{2}\mathbb{E}\Big(F^2\big(\Gamma_1(F) - \mathbb{E}(\Gamma_1(F))\big)\Big),\\
\kappa_6(F) &= \frac{5!}{2}\mathbb{E}\Big(F^2\big(\Gamma_3(F) - \mathbb{E}(\Gamma_3(F))\big)\Big)\\
&= \frac{5!}{3!}\Big\{\mathbb{E}\Big(F^3\big(\Gamma_2(F) - \mathbb{E}(\Gamma_2(F))\big)\Big) - \kappa_2(F)\kappa_4(F)\Big\}.
\end{aligned}$$

Proof. Using twice Malliavin integration by part formula, we obtain

$$\begin{aligned}
\mathbb{E}\left(F^2\Gamma_{s-1}(F)\right) &= \mathbb{E}(F^2)\mathbb{E}(\Gamma_{s-1}(F)) + 2\mathbb{E}(F\Gamma_s(F))\\
&= \mathbb{E}(F^2)\mathbb{E}(\Gamma_{s-1}(F)) + 2\mathbb{E}(\Gamma_{s+1}(F)).
\end{aligned}$$

Hence

$$\mathbb{E}\Big(F^2\big(\Gamma_{s-1}(F) - \mathbb{E}\Gamma_{s-1}(F)\big)\Big) = 2\mathbb{E}(\Gamma_{s+1}F) = \frac{2}{(s+1)!}\kappa_{s+2}(F).$$

The second equality for κ_6 comes from a direct application of Malliavin integration by part formula. □

Remark 18. (Γ_2-conjecture revisited). Assume that we have a sequence $\{F_n\}_{n\geq 1}$ of multiple integrals of a fixed order $p \geq 2$, and we aim to characterize convergence in distribution of F_n toward the target distribution $F_\infty := N_1 \times N_2$ in terms of convergences of finitely many cumulants. According to Corollary 9 it is enough to have the following convergences: (i) $\kappa_r(F_n) \to \kappa_r(F_\infty)$ for $r = 2, 3$, and (ii) $\mathrm{Var}(\Gamma_2(F_n)) - 2\mathbb{E}\big(F_n(\Gamma_2(F_n) - \mathbb{E}\Gamma_2(F_n))\big) + \kappa_2(F_n) \to 0$. Using Lemma 10 the convergence in (ii) can be rewritten as

$$\mathrm{Var}(\Gamma_2(F_n)) - 1/3\,\kappa_4(F_n) + \kappa_2(F_n) \to 0.$$

Now, assume that we have

$$\begin{aligned}
&\sum_{r=2}^{\deg Q=6} \frac{Q^{(r)}(0)}{r!}\frac{\kappa_r(F_n)}{2^{r-1}(r-1)!} = 1/2^5\Big\{\frac{\kappa_6(F_n)}{5!} - \frac{\kappa_4(F_n)}{3} + \kappa_2(F_n)\Big\}\\
&\qquad \longrightarrow \sum_{r=2}^{\deg Q=6} \frac{Q^{(r)}(0)}{r!}\frac{\kappa_r(F_\infty)}{2^{r-1}(r-1)!} = 0
\end{aligned} \tag{83}$$

Combining these observation, one just left to show that for any multiple integral F, we have the estimate $\mathrm{Var}(\Gamma_2(F)) \leq 1/5!\,\kappa_6(F)$.

References

[1] Azmoodeh, E., Campese, S., Poly, G. (2014) Fourth Moment Theorems for Markov Diffusion Generators. *J. Funct. Anal.* 266, 4, 2341-2359.

[2] Azmoodeh, E., Malicet, D., Mijoule, G., Poly, G. (2014) Generalization of the Nualart-Peccati criterion. *To appear in: The Annals of Probability.*

[3] Azmoodeh, E., Peccati, G., Poly, G. (2014). Convergence towards linear combinations of chi-squared random variables: a Malliavin-based approach. *To appear in: Séminaire de Probabilités* (Special volume in memory of Marc Yor)

[4] Azmoodeh, E., Morlanes, J. I. (2015) Drift parameter estimation for fractional Ornstein-Uhlenbeck process of the second kind. *Statistics*, 49, no. 1, 1-18.

[5] Azmoodeh, E., Viitasaari, L. (2014). Parameter estimation based on discrete observations of fractional Ornstein-Uhlenbeck process of the second kind. *Stat. Inference Stoch. Process.*

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

[7] Chen, L. H. Y., Goldstein, L., Shao, Qi-Man (2011) *Normal approximation by Stein's method.* Probability and its Applications. Springer.

[8] Eichelsbacher, P., Thäle, Ch. (2014). Malliavin-Stein method for Variance-Gamma approximation on Wiener space. Preprint

[9] Frenkel, P. (2008) Pfaffians, Hafnians and products of real linear functionals. *Math. Res. Lett.* 15, no. 2, 351-358.

[10] Hu, Y., Nualart, D. (2010) Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Statist. Probab. Lett.* Vol. 80, no. 11-12, 1030-1038.

[11] Li, W., Wei, A. (2012) Gaussian inequality for expected absolute products. *J. Theoret. Probab.* 25, no. 1, 92-99.

[12] Malicet, D., Nourdin, I., Peccati, G., and Poly, G. (2015). Squared chaotic random variables: new moment inequalities with applications. `http://arxiv.org/abs/1503.02154`

[13] Nourdin, I., Peccati, G. (2012) *Normal Approximations Using Malliavin Calculus: from Stein's Method to Universality.* Cambridge Tracts in Mathematics. Cambridge University.

[14] Nourdin, I., Peccati, G. (2009) Noncentral convergence of multiple integrals. *Ann. Probab.* 37, no. 4, 1412-1426.

[15] Nualart D. (2006) The Malliavin Calculus and Related Topics. *Probability and its Application.* Springer.

[16] Nourdin, I., Poly, G. (2012) Convergence in law in the second Wiener/Wigner chaos. *Electron. Commun. Probab.* 17, no. 36, 1-12.

[17] Peccati, G., Taqqu, M. (2011) Wiener Chaos: Moments, Cumulants and Diagrams. A survey with computer implementation. *Bocconi University Press.* Springer.

[18] Tudor, C.A. (2013) *Analysis of Variations for Self-similar Processes: A Stochastic Calculus Approach.* Springer-Verlag.