Stein's method, Malliavin calculus and infinite-dimensional Gaussian analysis

Giovanni PECCATI*

January 2009 – First draft†

Abstract

This expository paper is a companion of the four one-hour tutorial lectures given in the occasion of the special month *Progress in Stein's Method*, held at the University of Singapore in January 2009. We will explain how one can combine Stein's method with Malliavin calculus, in order to obtain explicit bounds in the normal and Gamma approximation of functionals of infinite-dimensional Gaussian fields. The core of our discussion is based on a series of papers jointly written with I. Nourdin, as well as with I. Nourdin and A. Réveillac.
Key Words: Central limit theorem; Gamma approximation; Gaussian approximation; Gaussian processes; Malliavin calculus; Stein's method; Wiener chaos.
Mathematics Subject Classification: 60F05 · 60G15 · 60H05 · 60H07

Contents

*Équipe Modal'X, Université Paris Ouest - Naterre La Défense, 200 Avenue de la République, 92000 Nanterre and LSTA, Université Paris VI, France. E-mail: `giovanni.peccati@gmail.com`.

†These notes will be posted and updated on my website www.geocities.com/giovannipeccati. Please, send me corrections/comments!

1 Overview and motivation

These lecture notes are an introduction to some new techniques (developed in the recent series of papers [57]–[61], [64] and [70]), bringing together *Stein's method* for normal and non-normal approximation (see e.g. [14], [80], [88] and [89]) and *Malliavin calculus* (see e.g. [35], [42], [43] and [65]). We shall see that the two theories fit together admirably well, and that their interaction leads to some remarkable new results involving central and non-central limit theorems for functionals of infinite-dimensional Gaussian fields.

Roughly speaking, the Gaussian Malliavin calculus is an infinite-dimensional differential calculus, involving operators defined on the class of functionals of a given Gaussian stochastic process. Its original motivation (see again [42], [43] and [65]) has been the obtention of a probabilistic proof of the so-called *Hörmander's theorem* for hypoelliptic operators. One of its most striking and well-established applications (which is tightly related to Hörmander's theorem) is the study of the regularity of the *densities* of random vectors, especially in connection with solutions of stochastic differential equations. Other crucial domains of application are: mathematical finance (see e.g. [44]), the non-anticipative stochastic calculus (see e.g. [65, Chapter 3]), the study of fractional processes (see e.g. [18] and [65, Chapter 5]) and, of course, limit theorems for sequences of functionals of Gaussian fields.

At the core of the Malliavin calculus lies the algebra of the so-called *Malliavin operators*, such as the *derivative operator*, the *divergence operator* and the *Ornstein-Uhlenbeck semigroup*. We will see that all these objects can be successfully characterized in terms of the *chaotic representation property*, stating that every square-integrable functional of a given Gaussian field is indeed an infinite orthogonal series of *multiple stochastic Wiener-Itô integrals* of increasing orders. As discussed in Section 7, the Malliavin operators are linked by several identities, all revolving around a fundamental result known as the *(infinite-dimensional) integration by parts formula*. It is interesting to note that this formula contains as a special case the "Stein's identity"

$$\mathbb{E}\left[f'\left(N\right)-Nf\left(N\right)\right]=0, \tag{1.1}$$

where $N\sim\mathcal{N}\left(0,1\right)$ and f is a smooth function verifying $\mathbb{E}\left|f'\left(N\right)\right|<\infty$. Also, we will see in Section 7.1 that equation (1.1) enters very naturally in the proof of one of the basic results of Malliavin calculus, that is, the *closability* of derivative operators (see Propostion 7.1 below). Other connections between Stein's method and Malliavin-type operators can be found in the papers by Hsu [32] and Decresuefond and Savy [17].

We will start our journey by describing a specific example involving quadratic functionals of a Brownian motion, and we will discuss the difficulties and drawbacks that are related with techniques that are not based on Stein's method, like for instance the *method of moments and cumulants*.

We stress that the applications of the theory presented in this paper go far beyond the examples that are discussed below: in particular, a great impetus is given by applications to limit theorems for functionals of fractional Gaussian processes. See for instance [57], [58], [60], or the lecture notes [53], for a discussion of this issue. We also point out the monograph [59] (in preparation).

In what follows, all random elements are implicitly defined on a suitable probability space $(\Omega,\mathcal{F},\mathbb{P})$.

Acknowledgements. I thank A. Barbour, L. Chen and K. P. Choi for their kind invitation and for their generous hospitality and support. I am grateful to I. Nourdin for useful remarks.

2 Preliminary example: exploding quadratic Brownian functionals *without* Stein's method

As anticipated in the Introduction, the theory developed in these lectures allows to apply Stein's method and Malliavin calculus to the study of the Gaussian and non-Gaussian approximation of non-linear functionals of *infinite-dimensional Gaussian fields*. By an infinite-dimensional Gaussian field we simply mean an infinite collection of jointly Gaussian real-valued random variables, such that the associated linear Gaussian space contains an *infinite i.i.d. (Gaussian) sequence*. We will implicitly prove in Section 4.1 that any infinite-dimensional Gaussian field can be represented in terms of an adequate Gaussian measure or, more generally, of an isonormal Gaussian process.

As a simple illustration of the problems we are interested in, we now present a typical situation where one can take advantage of Stein's method, that is: the asymptotic study of the quadratic functionals of a standard Brownian motion. We shall first state a general problem, and then describe two popular methods of solution (along with their drawbacks) *that are not based on Stein's method*. We will see in Section 9 that our Stein/Malliavin techniques can overcome the disadvantages of both approaches.

Observe that, in what follows, we shall sometimes use the notion of *cumulant*. Recall that, given a random variable Y with finite moments of all orders and with characteristic function $\psi_Y(t) = \mathbb{E}[\exp(itY)]$ $(t \in \mathbb{R})$, one defines the sequence of cumulants (sometimes known as *semi-invariants*) of Y, noted $\{\chi_n(Y) : n \geq 1\}$, as

$$\chi_n(Y) = (-i)^n \frac{d^n}{dt^n} \log \psi_Y(t) \,|_{t=0}, \quad n \geq 1. \tag{2.1}$$

For instance, $\chi_1(Y) = \mathbb{E}(Y)$, $\chi_2(Y) = \mathbb{E}[Y - \mathbb{E}(Y)]^2 = \mathbf{Var}(Y)$,

$$\chi_3(Y) = \mathbb{E}\left(Y^3\right) - 3\mathbb{E}\left(Y^2\right)\mathbb{E}(Y) + 2\mathbb{E}(Y)^3,$$

and so on. In general, one deduces from (2.1) that for every $n \geq 1$ the first n moments of Y can be expressed as polynomials in the first n cumulants (and viceversa). Note that (2.1) also implies that the cumulants of order $n \geq 3$ of a Gaussian random variable are equal to zero (recall also that the Gaussian distribution is determined by its moments, and therefore by its cumulants). We refer the reader to [72, Section 3] (but see also [83]) for a self-contained introduction to the basic combinatorial properties of cumulants.

2.1 Statement of the problem

Let $W = \{W_t : t \geq 0\}$ be a standard Brownian motion started from zero. This means that W is a centered Gaussian process such that $W_0 = 0$, W has continuous paths, and $\mathbb{E}[W_t W_s] = t \wedge s$ for every $t, s \geq 0$. See e.g. Revuz and Yor [82] for an exhaustive account of results and techniques related to Brownian motion.

In what follows, we shall focus on a specific property of the paths of W (first pointed out, in a slightly different form, in [37]), namely that

$$\int_0^1 \frac{W_t^2}{t^2} dt = \infty, \quad \text{a.s.-}\mathbb{P}. \tag{2.2}$$

As discussed in [37], and later in [36], relation (2.2) has deep connections with the theory of the (Gaussian) *initial enlargements of filtrations* in continuous-time stochastic calculus. See also [74] and [75] for applications to the study of Brownian local times.

Remark. Define the process $\hat{W}$ as $\hat{W}_0 = 0$ and $\hat{W}_u = uW_{1/u}$ for $u > 0$. A trivial covariance computation shows that $\hat{W}$ is also a standard Brownian motion. By using the change of variable $u = 1/t$, it now follows that property (2.2) is equivalent to the following statement:

$$\int_1^\infty \frac{W_u^2}{u^2} du = \infty, \quad \text{a.s.-}\mathbb{P}.$$

One natural question arising from (2.2) is therefore how to characterize the "rate of explosion", as $\varepsilon \to 0$, of the quantities

$$B_\varepsilon = \int_\varepsilon^1 \frac{W_t^2}{t^2} dt, \quad \varepsilon \in (0,1)\,. \tag{2.3}$$

One typical answer can be obtained by proving that some suitable renormalization of B_ε converges in distribution to a standard Gaussian random variable. By a direct computation, one can prove that $\mathbb{E}[B_\varepsilon] = \log 1/\varepsilon$ and $\mathbf{Var}(B_\varepsilon) \approx 4\log 1/\varepsilon$ [1]. By setting

$$\widetilde{B}_\varepsilon = \frac{B_\varepsilon - \log 1/\varepsilon}{\sqrt{4\log 1/\varepsilon}}, \quad \varepsilon \in (0,1)\,, \tag{2.4}$$

one can therefore meaningfully state the following problem.

Problem I. *Prove that, as $\varepsilon \to 0$,*

$$\widetilde{B}_\varepsilon \xrightarrow{\text{Law}} N \sim \mathcal{N}(0,1)\,, \tag{2.5}$$

where, here and for the rest of the paper, $\mathcal{N}(\alpha,\beta)$ denotes a one-dimensional Gaussian distribution with mean α and variance $\beta > 0$.

We shall solve Problem I by using both the classic *method of cumulants* and a stochastic calculus technique, known as *random time-change*. We will see below that both approaches suffer of evident drawbacks, and also that these difficulties can be successfully eliminated by means of our Stein/Malliavin approach.

2.2 The method of cumulants

The *method of (moments and) cumulants* is a very popular approach to the proof of limit results involving non-linear functionals of Gaussian fields. Its success relies mainly on the following two facts: (1) square-integrable functionals of Gaussian fields can always be represented in terms of (possibly infinite) series of multiple Wiener-Itô integrals (see e.g. Section 4 below and [41]), and (2) moments and cumulants of multiple integrals can be computed (at least formally) by means of well-established combinatorial devices, known as *diagram formulae* (see e.g. [72]). Classic

[1] In what follows, we shall write $\gamma(\varepsilon) \approx \varphi(\varepsilon)$, whenever $\frac{\gamma(\varepsilon)}{\varphi(\varepsilon)} \to 1$, as $\varepsilon \to 0$.

references for the method of cumulants in a Gaussian framework are [5], [7], [27] and, more recently, [26] and [46]. See the surveys by Peccati and Taqqu [72] and Surgailis [92] for more references and more detailed discussions.

In order to apply the method of cumulants to the proof of (2.4), one should start with the classic Itô formula $W_t^2 = 2\int_0^t W_s dW_s + t$, $t \in [0,1]$, and then write

$$\widetilde{B}_\varepsilon = \frac{B_\varepsilon - \log 1/\varepsilon}{\sqrt{4\log 1/\varepsilon}} = \frac{2\int_\varepsilon^1 \left[\int_0^t W_s dW_s\right] t^{-2} dt}{\sqrt{4\log 1/\varepsilon}}, \quad \varepsilon \in (0,1). \tag{2.6}$$

It is a standard result of stochastic calculus that one can interchange deterministic and stochastic integration on the RHS of (2.6) as follows:

$$\begin{aligned}\int_\varepsilon^1 \left[\int_0^t W_s dW_s\right] \frac{dt}{t^2} &= \int_\varepsilon^1 \left[\int_0^1 \mathbf{1}_{\{s<t\}} W_s dW_s\right] \frac{dt}{t^2} = \int_0^1 \left[\int_\varepsilon^1 \mathbf{1}_{\{s<t\}} \frac{dt}{t^2}\right] W_s dW_s \\ &= \int_0^1 \left[(s \vee \varepsilon)^{-1} - 1\right] W_s dW_s.\end{aligned}$$

As a consequence,

$$\widetilde{B}_\varepsilon = \frac{2\int_0^1 \left[(s\vee\varepsilon)^{-1} - 1\right] W_s dW_s}{\sqrt{4\log 1/\varepsilon}} = 2\int_0^1 \int_0^s f_\varepsilon(s,u)\, dW_u dW_s, \tag{2.7}$$

where f_ε is the symmetric and Lebesgue square-integrable function on $[0,1]^2$ given by

$$f_\varepsilon(s,u) = 2\left[(s\vee u\vee\varepsilon)^{-1} - 1\right] \times (4\log 1/\varepsilon)^{-1/2}. \tag{2.8}$$

By anticipating the terminology introduced in Section 4, formula (2.7) simply implies that each random variable $\widetilde{B}_\varepsilon$ is a member of the *second Wiener chaos* associated with W. We can now combine this fact with the results discussed e.g. in [35, Chapter VI] (see also Section 5 below), to deduce that, since the application $\varepsilon \mapsto \mathbf{Var}\left(\widetilde{B}_\varepsilon\right)$ is bounded, then, for every $n \geq 2$,

$$\sup_{\varepsilon>0} \mathbb{E}\left|\widetilde{B}_\varepsilon\right|^n < \infty. \tag{2.9}$$

Since $\mathbb{E}\left(\widetilde{B}_\varepsilon\right) = 0$ and $\mathbf{Var}\left(\widetilde{B}_\varepsilon\right) \to 1$, relation (2.9) implies immediately that (2.5) is proved once it is shown that, as $\varepsilon \to 0$,

$$\chi_n\left(\widetilde{B}_\varepsilon\right) \to 0, \quad \text{for every } n \geq 3. \tag{2.10}$$

To prove (2.10) we make use of a result by Fox and Taqqu [25] (see also [72] for an alternate combinatorial proof), stating that, for every fixed $n \geq 3$, the nth cumulant of $\widetilde{B}_\varepsilon = 2\int_0^1\int_0^s f_\varepsilon(s,u)\, dW_u dW_s$ is given by the following "chained integral"

$$\chi_n\left(\widetilde{B}_\varepsilon\right) = 2^{n-1}(n-1)! \int_{[0,1]^n} f_\varepsilon(t_1,t_2) f_\varepsilon(t_2,t_3) \cdots f_\varepsilon(t_{n-1},t_n) f_\varepsilon(t_n,t_1)\, dx_1 \cdots dx_n, \tag{2.11}$$

obtained by juxtaposing n copies of the kernel f_ε. By plugging (2.8) into (2.11), and after some lengthy (but standard) computations, one obtains that, as $\varepsilon \to 0$,

$$\chi_n\left(\widetilde{B}_\varepsilon\right) \approx c_n \times \left(\log \frac{1}{\varepsilon}\right)^{1-\frac{n}{2}}, \quad \text{for every } n \geq 3, \tag{2.12}$$

where $c_n > 0$ is a finite constant independent of ε. This yields (2.10) and therefore (2.5). The implication (2.12) $\Rightarrow$ (2.10) $\Rightarrow$ (2.5) is a typical application of the method of cumulants to the proof of Central Limit Theorems (CLTs) for functionals of Gaussian fields. In particular, one should note that (2.11) can be equivalently expressed in terms of diagram formulae.

In the following list we pinpoint some of the main disadvantages of this approach. As already discussed, all these difficulties disappear when using the Stein/Malliavin techniques developed in Section 9 below.

D1 Formulae (2.11) and (2.12) characterize the speed of convergence to zero of the cumulants of $\widetilde{B}_\varepsilon$. However, there is no way to deduce from (2.12) an estimate for quantities of the type $d\left(\widetilde{B}_\varepsilon, N\right)$, where d indicates some distance between the law of $\widetilde{B}_\varepsilon$ and the law of N (d can be for instance the *total variation distance*, or the *Wasserstein distance* – see Section 8.1 below)[2].

D2 Relations (2.10) and (2.11) require that, in order to prove the CLT (2.5), one verifies an *infinity* of asymptotic relations, each one involving the estimate of a multiple deterministic integral of increasing order. This task can be computationally quite demanding. Here, (2.12) is obtained by exploiting the elementary form of the kernels f_ε in (2.8).

D3 If one wants to apply the method of cumulants to elements of higher chaoses (for instance, by considering functionals involving Hermite polynomials of degree greater than 3), then one is forced to use diagram formulae that are much more involved than the Fox-Taqqu formula (2.11). Some examples of this situation appear e.g. in [5], [27] and [46]. See [72, Section 3 and Section 7] for an introduction to general diagram formulae for non-linear functionals of random measures.

2.3 Random time-changes

This technique has been used in [74] and [75]; see also [71] and [96] for some analogous results in the context of stable convergence.

Our starting point is once again formula (2.7), implying that, for each $\varepsilon \in (0,1)$, the random variable $\widetilde{B}_\varepsilon$ coincides with the value at the point $t = 1$ of the continuous Brownian martingale

$$t \mapsto M_t^\varepsilon = 2\int_0^t \int_0^s f_\varepsilon(s,u)\, dW_u dW_s, \quad t \in [0,1]. \tag{2.13}$$

It is well-known that the martingale M_t^ε has a quadratic variation equal to

$$\langle M^\varepsilon, M^\varepsilon\rangle_t = 4\int_0^t \left(\int_0^s f_\varepsilon(s,u)\, dW_u\right)^2 ds, \quad t \in [0,1].$$

[2] This assertion is not accurate, although it is kept for dramatic effect. Indeed, we will show in Section 10.2 that the combination of Stein's method and Malliavin calculus exactly allows to deduce Berry-Esséen bounds from estimates on cumulants.

By virtue of a classic stochastic calculus result, known as the *Dambis-Dubins-Schwarz Theorem* (DDS Theorem – see [82, Ch. V]), for every $\varepsilon \in (0,1)$ there exists (on a possibly enlarged probability space) a standard Brownian motion β^ε, initialized at zero and such that

$$M_t^\varepsilon = \beta^\varepsilon_{\langle M^\varepsilon, M^\varepsilon \rangle_t}, \quad t \in [0,1]. \tag{2.14}$$

It is important to stress that the definition of β^ε strongly depends on ε, and that β^ε is in general *not adapted* to the natural filtration of W. Moreover, one has that there exists a (continuous) filtration $\mathcal{G}_s^\varepsilon$, $s \geq 0$, such that β_s^ε is a $\mathcal{G}_s^\varepsilon$-Brownian motion and (for every fixed t) the positive random variable $\langle M^\varepsilon, M^\varepsilon \rangle_t$ is a $\mathcal{G}_s^\varepsilon$-*stopping time*. Formula (2.14) yields in particular that

$$\widetilde{B}_\varepsilon = M_1^\varepsilon = \beta^\varepsilon_{\langle M^\varepsilon, M^\varepsilon \rangle_1}.$$

Now consider a Lipschitz function h such that $\|h'\|_\infty \leq 1$, and observe that, for every $\varepsilon > 0$, $\beta_1^\varepsilon \overset{\text{Law}}{=} N \sim \mathcal{N}(0,1)$. A careful application of the Burkholder-Davis-Gundy (BDG) inequality (in the version stated in [82, Corollary 4.2, Ch. IV]) yields the following estimates:

$$\begin{aligned}
\left|\mathbb{E}[h(\widetilde{B}_\varepsilon)] - \mathbb{E}[h(N)]\right| &= \left|\mathbb{E}[h(\beta^\varepsilon_{\langle M^\varepsilon, M^\varepsilon \rangle_1})] - \mathbb{E}[h(\beta_1^\varepsilon)]\right| \\
&\leq \mathbb{E}\left[\left|\beta^\varepsilon_{\langle M^\varepsilon, M^\varepsilon \rangle_1} - \beta_1^\varepsilon\right|\right] \\
&\leq \mathbb{E}\left[\left|\beta^\varepsilon_{\langle M^\varepsilon, M^\varepsilon \rangle_1} - \beta_1^\varepsilon\right|^4\right]^{\frac{1}{4}} \\
&\leq C\mathbb{E}\left[\left|\langle M^\varepsilon, M^\varepsilon \rangle_1 - 1\right|^2\right]^{\frac{1}{4}} \\
&= C\mathbb{E}\left[\left|4\int_0^1 \left(\int_0^s f_\varepsilon(s,u)\, dW_u\right)^2 ds - 1\right|^2\right]^{\frac{1}{4}},
\end{aligned} \tag{2.15}$$

where C is some universal constant independent of ε. The CLT (2.5) is now obtained from (2.15) by means of a direct computation, yielding that, as $\varepsilon \to 0$,

$$\mathbb{E}\left[\left|4\int_0^1 \left(\int_0^s f_\varepsilon(s,u)\, dW_u\right)^2 ds - 1\right|^2\right] \approx \frac{\alpha}{\log 1/\varepsilon} \to 0, \tag{2.16}$$

where $\alpha > 0$ is some constant independent of ε.

Note that this approach is more satisfactory than the method of cumulants. Indeed, the chain of relations starting at (2.15) allows to assess explicitly the Wasserstein distance between the law of $\widetilde{B}_\varepsilon$ and the law of N [3] (albeit the implied rate of $(\log 1/\varepsilon)^{-1/4}$ is suboptimal – see Section 11.1). Moreover, the proof of (2.5) is now reduced to a single asymptotic relation, namely (2.16). However, at least two crucial points make this approach quite difficult to apply in general situations.

[3] Recall that the Wasserstein distance between the law of two variables X_1, X_2 is given by $d_W(X_1, X_2) = \sup |\mathbb{E}[h(X_1)] - \mathbb{E}[h(X_2)]|$, where the supremum is taken over all Lipschitz functions such that $\|h'\|_\infty \leq 1$. See Section 8.1 below.

D4 The application of the DDS Theorem and of the BDG inequality requires an explicit underlying (Brownian) martingale structure. Although it is always possible to represent a given Gaussian field in terms of a Brownian motion, this operation is often quite unnatural and can render the asymptotic analysis very hard. For instance, what happens if one considers quadratic functionals of a multiparameter Gaussian process, or of a Gaussian process which is not a semimartingale (for instance, a fractional Brownian motion with Hurst parameter $H \neq 1/2$)? See [71] for some further applications of random time-changes in a general Gaussian setting.

D5 It is not clear whether this approach can be used in order to deal with expressions of the type (2.15), when h is not Lipschitz (for instance, when h equals the indicator of a Borel set), so that it seems difficult to use these techniques in order to assess other distances, like the *total variation distance* or the *Kolmogorov distance*.

Starting from the next section, we will describe the main objects and tools of stochastic analysis that are involved in our techniques.

3 Gaussian measures

Let $(Z, \mathcal{Z})$ be a Polish space, with $\mathcal{Z}$ the associated Borel σ-field, and let μ be a positive σ-finite measure over $(Z, \mathcal{Z})$ with no atoms (that is, $\mu(\{z\}) = 0$, for every $z \in Z$). We denote by $\mathcal{Z}_\mu$ the class of those $A \in \mathcal{Z}$ such that $\mu(A) < \infty$. Note that, by σ-additivity, the σ-field generated by $\mathcal{Z}_\mu$ coincides with $\mathcal{Z}$.

Definition 3.1 *A* ***Gaussian measure*** *on $(Z, \mathcal{Z})$ with* ***control*** *μ is a centered Gaussian family of the type*

$$G = \{G(A) : A \in \mathcal{Z}_\mu\}, \tag{3.1}$$

verifying the relation

$$\mathbb{E}[G(A)G(B)] = \mu(A \cap B), \quad \forall A, B \in \mathcal{Z}_\mu. \tag{3.2}$$

The Gaussian measure G is also called a ***white noise based on*** *μ.*

Remarks. (a) A Gaussian measure such as (3.1)–(3.2) always exists (just regard G as a centered Gaussian process indexed by $\mathcal{Z}_\mu$, and then apply the usual Kolmogorov criterion).

(b) Relation (3.2) implies that, for every pair of disjoint sets $A, B \in \mathcal{Z}_\mu$, the random variables $G(A)$ and $G(B)$ are independent. When this property is verified, one usually says that G is a *completely random measure* (or, equivalently, an *independently scattered random measure*). The concept of a completely random measure can be traced back to Kingman's seminal paper [38]. See e.g. [39], [72], [92] and [93] for a discussion around general (for instance, Poisson) completely random measures.

(c) Let $B_1, ..., B_n, ...$ be a sequence of *disjoint* elements of $\mathcal{Z}_\mu$, and let G be a Gaussian measure on $(Z, \mathcal{Z})$ with control μ. Then, for every finite $N \geq 2$, one has that $\cup_{n=1}^N B_n \in \mathcal{Z}_\mu$, and, by using (3.2)

$$\mathbb{E}\left[\left(G\left(\cup_{n=1}^N B_i\right) - \sum_{n=1}^N G(B_n)\right)^2\right] = \mu\left(\cup_{n=1}^N B_n\right) - \sum_{n=1}^N \mu(B_n) = 0, \tag{3.3}$$

because μ is a measure, and therefore it is finitely additive. Relation (3.3) implies in particular that

$$G\left(\cup_{n=1}^N B_n\right) = \sum_{n=1}^N G(B_n), \quad \text{a.s.-}\mathbb{P}. \tag{3.4}$$

Now suppose that $\cup_{n=1}^\infty B_n \in \mathcal{Z}_\mu$. Then, by (3.4) and again by virtue of (3.2),

$$\begin{aligned}
\mathbb{E}\left[\left(G\left(\cup_{n=1}^\infty B_n\right) - \sum_{n=1}^N G(B_n)\right)^2\right] &= \mathbb{E}\left[\left(G\left(\cup_{n=1}^\infty B_n\right) - G\left(\cup_{n=1}^N B_i\right)\right)^2\right] \\
&= \mu\left(\cup_{n=N+1}^\infty B_n\right) \underset{N\to\infty}{\to} 0,
\end{aligned}$$

because μ is σ-additive. This entails in turn that

$$G\left(\cup_{n=1}^\infty B_n\right) = \sum_{n=1}^\infty G(B_n), \quad a.s. - \mathbb{P}, \tag{3.5}$$

where the series on the RHS converges in $L^2(\mathbb{P})$. Relation (3.5) simply means that the application

$$\mathcal{Z}_\mu \to L^2(\mathbb{P}) : B \mapsto G(B),$$

is σ-additive, and therefore that *the Gaussian measure G is a σ-additive measure with values in the Hilbert space $L^2(\mathbb{P})$*. This remarkable feature of G is the starting point of the combinatorial theory of multiple stochastic integration (also applying to more general random measures) developed by Engel [24] and Rota and Wallstrom [84]. In particular, the crucial facts used in [84] are the following: (1) for every $n \geq 2$, one can canonically associate with G a $L^2(\mathbb{P})$-valued σ-additive measure on the product space $(Z^n, \mathcal{Z}^n)$, and (2) one can completely develop a theory of stochastic integration with respect to G by exploiting the isomorphism between the *diagonal subsets* of Z^n and the lattice of partitions of the set $\{1, ..., n\}$, and by using the properties of the associated Möbius function. In what follows we will not adopt this (rather technical) point of view. See the survey by Peccati and Taqqu [72] for a detailed and self-contained account of the Engel-Rota-Wallstrom theory.

(d) Note that it is *not* true that, for a Gaussian measure G and for a fixed $\omega \in \Omega$, the application

$$\mathcal{Z}_\mu \to \mathbb{R} : B \mapsto G(B)(\omega)$$

is a σ-additive real-valued (signed) measure.

Notation. For the rest of the paper, we shall write $(Z^n, \mathcal{Z}^n) = (Z^{\otimes n}, \mathcal{Z}^{\otimes n})$, $n \geq 2$, and also $\left(Z^1, \mathcal{Z}^1\right) = \left(Z^{\otimes 1}, \mathcal{Z}^{\otimes 1}\right) = (Z, \mathcal{Z})$. Moreover, we set

$$\mathcal{Z}^n_\mu = \{C \in \mathcal{Z}^n : \mu^n (C) < \infty\}.$$

Examples. (i) Let $Z = \mathbb{R}$, $\mathcal{Z} = \mathcal{B}(\mathbb{R})$, and let λ be the Lebesgue measure. Consider a Gaussian measure G with control λ: then, for every Borel subsets $A, B \in \mathcal{B}(\mathbb{R})$ with finite Lebesgue measure, one has that

$$\mathbb{E}[G(A)\,G(B)] = \lambda(A \cap B) = \int_{A \cap B} \lambda(dx). \tag{3.6}$$

In particular, the random function

$$t \mapsto W_t \triangleq G([0, t]), \quad t \geq 0, \tag{3.7}$$

defines a centered Gaussian process such that $W_0 = 0$ and $\mathbb{E}[W_t W_s] = \lambda([0, t] \cap [0, s]) = s \wedge t$, that is, W is a *standard Brownian motion* started from zero. Note that, in order to meet the usual definition of a standard Brownian motion, one should select an appropriate *continuous* version of the process W appearing in (3.7).

(ii) Fix $d \geq 2$, let $Z = \mathbb{R}^d$, $\mathcal{Z} = \mathcal{B}\left(\mathbb{R}^d\right)$, and let λ^d be the Lebesgue measure on $\mathbb{R}^d$. If G is a Gaussian measure with control λ^d, then, for every $A, B \in \mathcal{B}\left(\mathbb{R}^d\right)$ with finite Lebesgue measure, one has that

$$\mathbb{E}[G(A)\,G(B)] = \int_{A \cap B} \lambda^d(dx_1, ..., dx_d).$$

It follows that the application

$$(t_1, ...t_d) \mapsto \mathbf{W}(t_1, ..., t_d) \triangleq G([0, t_1] \times \cdots \times [0, t_d]), \quad t_i \geq 0, \tag{3.8}$$

defines a centered Gaussian process such that

$$\mathbb{E}[\mathbf{W}(t_1, ..., t_d)\,\mathbf{W}(s_1, ..., s_d)] = \prod_{i=1}^{d} (s_i \wedge t_i),$$

that is, $\mathbf{W}$ is a *standard Brownian sheet* on $\mathbb{R}^d_+$.

4 Wiener-Itô integrals

In this section, we define single and multiple Wiener-Itô integrals with respect to Gaussian measures. The main interest of this construction will be completely unveiled in Section 5.3, where we will prove that Wiener-Itô integrals are indeed the basic building blocks of *any* square-integrable functional of a given Gaussian measure. Our main reference is Chapter 1 in Nualart's monograph [65]. Other strongly suggested readings are the books by Dellacherie *et al.* [19] and Janson [35]. See also the original paper by Itô [34] (but beware of the diagonals! – see Masani [49]), as well as [24], [39], [41], [43], [82], [84], [92], [93].

4.1 Single integrals and the first Wiener chaos

Let $(Z, \mathcal{Z}, \mu)$ be a Polish measure space, with μ σ-finite and non-atomic. We denote by $L^2(Z, \mathcal{Z}, \mu) = L^2(\mu)$ the Hilbert space of real-valued functions on $(Z, \mathcal{Z})$ that are square-integrable with respect to μ. We also write $\mathcal{E}(\mu)$ to indicate the subset of $L^2(\mu)$ composed of *elementary functions*, that is, $f \in \mathcal{E}(\mu)$ if and only if

$$f(z) = \sum_{i=1}^{M} a_i \mathbf{1}_{A_i}(z), \quad z \in Z, \tag{4.1}$$

where $M \geq 1$ is finite, $a_i \in \mathbb{R}$, and the sets A_i are pairwise disjoint elements of $\mathcal{Z}_\mu$. Plainly, $\mathcal{E}(\mu)$ is a linear space and $\mathcal{E}(\mu)$ is *dense* in $L^2(\mu)$.

Now consider a Gaussian measure G on $(Z, \mathcal{Z})$, with control μ. The next result establishes the existence of single Wiener-Itô integrals with respect to G.

Proposition 4.1 *There exists a unique linear isomorphism $f \mapsto G(f)$, from $L^2(\mu)$ into $L^2(\mathbb{P})$, such that*

$$G(f) = \sum_{i=1}^{M} a_i \times G(A_i) \tag{4.2}$$

for every elementary function $f \in \mathcal{E}(\mu)$ of the type (4.1).

Proof. For every $f \in \mathcal{E}(\mu)$, set $G(f)$ to be equal to (4.2). Then, by using (3.2) one has that, for every pair $f, f' \in \mathcal{E}(\mu)$,

$$\mathbb{E}\left[G(f)\, G(f')\right] = \int_Z f(z)\, f'(z)\, \mu(dz). \tag{4.3}$$

Since $\mathcal{E}(\mu)$ is dense in $L^2(\mu)$, the proof is completed by the following (standard) approximation argument. If $f \in L^2(\mu)$ and $\{f_n\}$ is a sequence of elementary kernels converging to f, then (4.3) implies that $\{G(f_n)\}$ is a Cauchy sequence in $L^2(\mathbb{P})$, and one defines $G(f)$ to be the $L^2(\mathbb{P})$ limit of $G(f_n)$. One easily verifies that the definition of $G(f)$ does not depend on the chosen approximating sequence $\{f_n\}$. The application $f \mapsto G(f)$ is therefore well-defined, and (by virtue of (4.3)) it is a linear isomorphism from $L^2(\mu)$ into $L^2(\mathbb{P})$. ■

The random variable $G(f)$ is usually written as

$$\int_Z f(z)\, G(dz), \quad \int_Z f dG, \quad I_1^G(f) \quad \text{or} \quad I_1(f), \tag{4.4}$$

(note that in the last formula the symbol G is omitted) and it is called the *Wiener-Itô stochastic integral* of f with respect to G. By inspection of the previous proof, one sees that Wiener-Itô integrals verify the isometric relation

$$\mathbb{E}\left[G(f)\, G(h)\right] = \int_Z f(z)\, h(z)\, \mu(dz) = \langle f, h \rangle_{L^2(\mu)}, \quad \forall f, h \in L^2(\mu). \tag{4.5}$$

Observe also that $\mathbb{E}[G(f)] = 0$, and therefore (4.5) implies that every random vector of the type $(G(f_1), ..., G(f_d))$, $f_i \in L^2(\mu)$, is a d-dimensional centered Gaussian vector with covariance matrix $\Gamma(i,j) = \langle f_i, f_j \rangle_{L^2(\mu)}$, $1 \leq i,j \leq d$. If $B \in \mathcal{Z}_\mu$, we write interchangeably $G(B)$ or $G(\mathbf{1}_B)$ (the two objects coincide, thanks to (4.2)). Plainly, the Gaussian family

$$C_1(G) = \left\{G(f) : f \in L^2(\mu)\right\} \tag{4.6}$$

coincides with the $L^2(\mathbb{P})$-closed linear space generated by G. One customarily says that (4.6) is the *first Wiener chaos* associated with G. Observe that, if $\{e_i : i \geq 1\}$ is an orthonormal basis of $L^2(\mu)$, then $\{G(e_i) : i \geq 1\}$ is an i.i.d. Gaussian sequence with zero mean and common unitary variance.

4.2 Multiple integrals

For every $n \geq 2$, we write $L^2(Z^n, \mathcal{Z}^n, \mu^n) = L^2(\mu^n)$ to indicate the Hilbert space of real-valued functions that are square-integrable with respect to μ^n. Given a function $f \in L^2(\mu^n)$, we denote by $\widetilde{f}$ its canonical symmetrization, that is

$$\widetilde{f}(z_1, ..., z_n) = \frac{1}{n!}\sum_\pi f\left(z_{\pi(1)}, ...z_{\pi(n)}\right), \tag{4.7}$$

where the sum runs over all permutations π of the set $\{1, ..., n\}$. Note that, by the triangle inequality,

$$\| \widetilde{f} \|_{L^2(\mu^n)} \leq \|f\|_{L^2(\mu^n)}. \tag{4.8}$$

We will consider the following three subsets of $L^2(\mu^n)$.

- $L_s^2(Z^n, \mathcal{Z}^n, \mu^n) = L_s^2(\mu^n)$ is the closed linear subspace of $L^2(\mu^n)$ composed of *symmetric* functions, that is, $f \in L_s^2(\mu^n)$ if and only if: (i) f is square integrable with respect to μ^n, and (ii) for $d\mu^n$-almost every $(z_1, ..., z_n) \in Z^n$,

  $$f(z_1, ..., z_n) = f\left(z_{\pi(1)}, ..., z_{\pi(n)}\right),$$

  for every permutation π of $\{1, ..., n\}$.

- $\mathcal{E}(\mu^n)$ is the subset of $L^2(\mu^n)$ composed of *elementary functions vanishing on diagonals*, that is, $f \in \mathcal{E}(\mu^n)$ if and only if f is a finite linear combination of functions of the type

  $$(z_1, ..., z_n) \mapsto \mathbf{1}_{A_1}(z_1)\, \mathbf{1}_{A_2}(z_2) \cdots \mathbf{1}_{A_n}(z_n) \tag{4.9}$$

  where the sets A_i are pairwise disjoint elements of $\mathcal{Z}_\mu$.

- $\mathcal{E}_s(\mu^n)$ is the subset of $L_s^2(\mu^n)$ composed of *symmetric* elementary functions vanishing on diagonals, that is, $g \in \mathcal{E}_s(\mu^n)$ if and only if $g = \widetilde{f}$ for some $f \in \mathcal{E}(\mu^n)$, where the symmetrization $\widetilde{f}$ is defined according to (4.7).

The following technical result will be used throughout the sequel.

Lemma 4.1 *Fix $n \geq 2$. Then, $\mathcal{E}(\mu^n)$ is dense in $L^2(\mu^n)$, and $\mathcal{E}_s(\mu^n)$ is dense in $L_s^2(\mu^n)$.*

Proof. Since, for every $h \in L_s^2(\mu)$ and every $f \in \mathcal{E}(\mu^n)$, by symmetry,

$$\langle h, f\rangle_{L^2(\mu^n)} = \langle h, \widetilde{f}\rangle_{L^2(\mu^n)},$$

it is enough to prove that $\mathcal{E}(\mu^n)$ is dense in $L^2(\mu^n)$. We shall only provide a detailed proof for $n = 2$ (the general case is analogous – see e.g. [65, Section 1.1.2]). To prove the desired claim, it is therefore sufficient to show that every function of the type $h(z_1, z_2) = \mathbf{1}_A(z_1)\mathbf{1}_B(z_2)$, with $A, B \in \mathcal{Z}_\mu$, is the limit in $L^2(\mu)$ of linear combinations of products of the type $\mathbf{1}_{D_1}(z_1)\mathbf{1}_{D_2}(z_2)$, with $D_1, D_2 \in \mathcal{Z}_\mu$ and $D_1 \cap D_2 = \emptyset$. To do this, define $C_1 = A\backslash B$, $C_2 = B\backslash A$ and $C_3 = A \cap B$, so that

$$h = \mathbf{1}_{C_1}\mathbf{1}_{C_2} + \mathbf{1}_{C_1}\mathbf{1}_{C_3} + \mathbf{1}_{C_3}\mathbf{1}_{C_2} + \mathbf{1}_{C_3}\mathbf{1}_{C_3}.$$

If $\mu(C_3) = 0$, there is nothing to prove. If $\mu(C_3) > 0$, since μ is non-atomic, for every $N \geq 2$ we can find disjoint sets $C_3(i, N) \subset C_3$, $i = 1, ..., N$, such that $\mu(C_3(i, N)) = \mu(C_3)/N$ and $\cup_{i=1}^N C_3(i, N) = C_3$. It follows that

$$\begin{aligned}\mathbf{1}_{C_3}(z_1)\mathbf{1}_{C_3}(z_2) &= \sum_{1\leq i\neq j\leq N}\mathbf{1}_{C_3(i,N)}(z_1)\mathbf{1}_{C_3(j,N)}(z_2) + \sum_{i=1}^N \mathbf{1}_{C_3(i,N)}(z_1)\mathbf{1}_{C_3(i,N)}(z_2) \\ &= h_1(z_1, z_2) + h_2(z_1, z_2).\end{aligned}$$

Plainly, $h_1 \in \mathcal{E}(\mu^n)$, and

$$\|h_2\|^2_{L^2(\mu^n)} = \sum_{i=1}^N \mu(C_3(i, N))^2 = \frac{\mu(C_3)^2}{N}.$$

Since N is arbitrary, we deduce the desired conclusion. ■

Fix $n \geq 2$. It is easily seen that every $f \in \mathcal{E}(\mu^n)$ admits a (not necessarily unique) representation of the form

$$f(z_1, ..., z_n) = \sum_{1\leq i_1,...,i_n\leq M} a_{i_1\cdots i_n}\mathbf{1}_{A_{i_1}}(z_1)\cdots\mathbf{1}_{A_{i_n}}(z_n) \tag{4.10}$$

where $M \geq n$, the real coefficients $a_{i_1\cdots i_n}$ are equal to zero whenever two indices i_k, i_l are equal and $A_1, ..., A_M$ are pairwise disjoint elements of $\mathcal{Z}_\mu$. For every $f \in \mathcal{E}(\mu^n)$ with the form (4.10) we set

$$I_n(f) = \sum_{1\leq i_1,...,i_n\leq M} a_{i_1\cdots i_n} G(A_{i_1})\cdots G(A_{i_n}), \tag{4.11}$$

and we say that $I_n(f)$ is the *multiple stochastic Wiener-Itô integral* (of order n) of f with respect to G. Note that $I_n(f)$ has finite moments of all orders, and that the definition of $I_n(f)$ does not depend on the chosen representation of f. The following result shows in particular that I_n can be extended to a continuous linear operator from $L^2(\mu^n)$ into $L^2(\mathbb{P})$. Note that the third point of the following statement also involves random variables of the form $I_1(g)$, $g \in L^2(\mu)$.

Proposition 4.2 *The random variables $I_n(f)$, $n \geq 1$, $f \in \mathcal{E}(\mu^n)$, enjoy the following properties*

1. *For every n, the application $f \mapsto I_n(f)$ is linear.*

2. *For every n, one has $\mathbb{E}(I_n(f)) = 0$ and $I_n(f) = I_n(\widetilde{f})$.*

3. *For every $n \geq 2$ and $m \geq 1$, for every $f \in \mathcal{E}(\mu^n)$ and $g \in \mathcal{E}(\mu^m)$ (if $m = 1$, one can take $g \in L^2(\mu)$),*

$$\mathbb{E}[I_n(f) I_m(g)] = \begin{cases} 0 & \text{if } n \neq m \\ n!\langle \widetilde{f}, \widetilde{g} \rangle_{L^2(\mu^n)} & \text{if } n = m. \end{cases} \tag{4.12}$$

The proof of Proposition 4.2 follows almost immediately from the defnition (4.11); see e.g. [65, Section 1.1.2] for a complete discussion. By combining (4.12) with (4.8), one infers that I_n can be extended to a linear continuous operator, from $L^2(\mu^n)$ into $L^2(\mathbb{P})$, verifying properties 1, 2 and 3 in the statement of Proposition 4.2. Moreover, the second line on the RHS of (4.12) yields that the application

$$I_n : L_s^2(\mu^n) \to L^2(\mathbb{P}) : f \mapsto I_n(f)$$

(that is, the restriction of I_n to $L_s^2(\mu^n)$) is an isomorphism from $L_s^2(\mu^n)$, endowed with the modified scalar product $n!\langle \cdot, \cdot \rangle_{L^2(\mu^n)}$, into $L^2(\mathbb{P})$. For every $n \geq 2$, the $L^2(\mathbb{P})$-closed vector space

$$C_n(G) = \left\{ I_n(f) : f \in L^2(\mu^n) \right\} \tag{4.13}$$

is called the *nth Wiener chaos* associated with G. One conventionally sets

$$C_0(G) = \mathbb{R}. \tag{4.14}$$

Note that (4.12) implies that $C_n(G) \perp C_m(G)$ for $n \neq m$, where " $\perp$ " indicates orthogonality in $L^2(\mathbb{P})$.

Remark (The case of Brownian motion). We consider the case where $(Z, \mathcal{Z}) = (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ and μ is equal to the Lebesgue measure. As already observed, one has that the process $t \mapsto W_t = G([0,t])$, $t \geq 0$, is a standard Brownian motion started from zero. Also, for every $f \in L^2(\mu)$,

$$I_1(f) = \int_{\mathbb{R}_+} f(t)\, G(dt) = \int_0^\infty f(t)\, dW_t, \tag{4.15}$$

where the RHS of (4.15) indicates a standard Itô integral with respect to W. Moreover, for every $n \geq 2$ and every $f \in L^2(\mu^n)$

$$I_n(f) = n! \int_0^1 \left[\int_0^{t_1} \int_0^{t_2} \cdots \int_0^{t_{n-1}} \widetilde{f}(t_1, ..., t_n)\, dW_{t_n} \cdots dW_{t_2} \right] dW_{t_1}, \tag{4.16}$$

where the RHS of (4.16) stands for a usual Itô-type stochastic integral, with respect to W, of the stochastic process

$$t \mapsto \varphi(t) = n! \int_0^t \int_0^{t_2} \cdots \int_0^{t_{n-1}} \widetilde{f}(t_1, ..., t_n)\, dW_{t_n} \cdots dW_{t_2}, \quad t_1 \geq 0.$$

Note in particular that $\varphi(t)$ is adapted to the filtration $\sigma\{W_u : u \leq t\}$, $t \geq 0$, and also

$$\mathbb{E}\left[\int_0^\infty \varphi^2(t)\,dt\right] < \infty.$$

Both equalities (4.15) and (4.16) can be easily proved for elementary functions that are constant on intervals (for (4.15)) or on rectangles (for (4.16)), and the general results are obtained by standard density arguments.

Remark (One more digression on the Engel-Rota-Wallstrom theory). As already evoked on page 10, in [24] and [84] it is proved that one can canonically associate to G a σ-additive $L^2(\mathbb{P})$-valued product measure on $(Z^n, \mathcal{Z}^n)$, say G^n. One can therefore prove that, for every $n \geq 2$ and every $f \in L^2(\mu^n)$, the random variable $I_n(f)$ has indeed the form

$$\begin{aligned} I_n(f) &= \int_{Z^n} f(z_1,...,z_n)\,\mathbf{1}_{D_0^n}(z_1,...,z_n)\,G^n(dz_1,...,dz_n) \qquad (4.17)\\ &= \int_{D_0^n} f(z_1,...,z_n)\,G^n(dz_1,...,dz_n), \end{aligned}$$

where D_0^n indicates the *purely non-diagonal set*

$$D_0^n = \{(z_1,...,z_n) : z_i \neq z_j \quad \forall i \neq j\}.$$

See [72] for a complete discussion of this point.

5 Multiplication formulae

5.1 Contractions and multiplications

The concept of *contraction* plays a fundamental role in the theory developed in this paper.

Definition 5.1 *Let μ be a σ-finite and non-atomic measure on the Polish space $(Z,\mathcal{Z})$. For every $q,p \geq 1$, $f \in L_s^2(\mu^p)$, $g \in L_s^2(\mu^q)$ and every $r = 0,...,q\wedge p$, the **contraction of order** r of f and g is the function $f \otimes_r g$ of $p+q-2r$ variables defined as follows: for $r = 1,...,p\wedge q$ and $(t_1,\dots,t_{p-r},s_1,\dots,s_{q-r}) \in Z^{p+q-2r}$,*

$$\begin{aligned} & f \otimes_r g(t_1,\dots,t_{p-r},s_1,\dots,s_{q-r}) \\ &= \int_{Z^r} f(z_1,\dots,z_r,t_1,\dots,t_{p-r})g(z_1,\dots,z_r,s_1,\dots,s_{q-r})\mu^r(dz_1...dz_r), \qquad (5.1) \end{aligned}$$

and, for $r = 0$,

$$\begin{aligned} f \otimes_r g(t_1,\dots,t_p,s_1,\dots,s_q) &= f \otimes g(t_1,\dots,t_p,s_1,\dots,s_q) \qquad (5.2)\\ &= f(t_1,\dots,t_{p-r})g(s_1,\dots,s_{q-r}). \end{aligned}$$

Note that, if $p = q$, then $f \otimes_p g = \langle f,g\rangle_{L^2(\mu^p)}$. For instance, if $p = q = 2$, one has

$$f \otimes_1 g(t,s) = \int_Z f(z,t)\,g(z,s)\,\mu(dz), \qquad (5.3)$$

$$f \otimes_2 g = \int_{Z^2} f(z_1,z_2)\,g(z_1,z_2)\,\mu^2(dz_1,dz_2). \qquad (5.4)$$

By an application of the Cauchy-Schwarz inequality, it is straightforwrd to prove that, for every $r = 0, ..., q \wedge p$, the function $f \otimes_r g$ is an element of $L^2\left(\mu^{p+q-2r}\right)$. Note that $f \otimes_r g$ is in general not symmetric (although f and g are): we shall denote by $f \widetilde{\otimes}_r g$ the canonical symmetrization of $f \otimes_r g$, as given in (4.7).

For the rest of this section, G is a Gaussian measure on the Polish space $(Z, \mathcal{Z})$, with non-atomic and σ-finite control μ; I_p indicates a multiple Wiener-Itô integral with respect to G, as defined in Section 4.2. The next result is a multiplication formula for multiple Wiener-Itô integrals. It will be crucial for the rest of this paper.

Theorem 5.1 *For every $p, q \geq 1$ and every $f \in L^2(\mu^p)$, $g \in L^2(\mu^q)$,*

$$I_p(f)\, I_q(g) = \sum_{r=0}^{p\wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}\left(\widetilde{f} \otimes_r \widetilde{g}\right). \tag{5.5}$$

Theorem 5.1, whose proof is omitted, can be established by at least two routes, namely by induction (see [65, Proposition 1.1.3]), or by using the concept of "diagonal measure" in the context of the Engel-Rota-Wallstrom theory (see [72, Section 6.4]).

Remark. Recall the notation (4.6), (4.13) and (4.14). Formula (5.5) implies that, for every $m \geq 1$, a random variable belonging to the space $\oplus_{j=0}^m C_j(G)$ (where " $\oplus$ " stands for an orthogonal sum in $L^2(\mathbb{P})$) has finite moments of any order. More precisely, for every $p > 2$ and every $n \geq 1$, one can prove that there exists a universal constant $c_{p,n} > 0$, such that

$$\mathbb{E}\left[|I_n(f)|^p\right]^{1/p} \leq c_{n,p} \mathbb{E}\left[I_n(f)^2\right]^{1/2}, \tag{5.6}$$

$\forall\, f \in L^2(\mu^n)$ (see e.g. [35, Ch. V]). Finally, on every finite sum of Wiener chaoses $\oplus_{j=0}^m C_j(G)$ and for every $p \geq 1$, the topology induced by the $L^p(\mathbb{P})$ convergence is equivalent to the L^0-topology induced by convergence in probability, that is, convergence in probability is equivalent to convergence in L^p, for every $p \geq 1$. This fact has been first proved by Schreiber in [86] – see also [35, Chapter VI]. One can also prove that *the law of a non-zero random variable living in a finite sum of Wiener chaoses always admits a density.*

5.2 Multiple stochastic integrals as Hermite polynomials

Definition 5.2 *The sequence of* ***Hermite polynomials*** $\{H_q : q \geq 0\}$ *on* $\mathbb{R}$*, is defined via the following relations:* $H_0 \equiv 1$ *and, for* $q \geq 1$,

$$H_q(x) = (-1)^q e^{\frac{x^2}{2}} \frac{d^q}{dx^q} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}. \tag{5.7}$$

For instance, $H_1(x) = 1$, $H_2(x) = x^2 - 1$ *and* $H_3(x) = x^3 - 3x$.

Recall that the sequence $\{(q!)^{-1/2} H_q : q \geq 0\}$ is an orthonormal basis of $L^2(\mathbb{R}, (2\pi)^{-1/2} e^{-x^2/2}dx)$. Several relevant properties of Hermite polynomials can be deduced from the following formula, valid for every $t, x \in \mathbb{R}$,

$$\exp\left(tx - \frac{t^2}{2}\right) = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x). \tag{5.8}$$

For instance, one deduces immediately from the previous expression that

$$\begin{aligned} \frac{d}{dx} H_n(x) &= nH_{n-1}(x), \quad n \geq 1, && (5.9) \\ H_{n+1}(x) &= xH_n(x) - nH_{n-1}(x), \quad n \geq 1. && (5.10) \end{aligned}$$

The next result uses (5.5) and (5.10) in order to establish an explicit relation between multiple stochastic integrals and Hermite polynomials.

Proposition 5.1 *Let $h \in L^2(\mu)$ be such that $\|h\|_{L^2(\mu)} = 1$, and, for $n \geq 2$, define*

$$h^{\otimes n}(z_1, .., z_n) = h(z_1) \times \cdots \times h(z_n), \quad (z_1, ..., z_n) \in Z^n.$$

Then,

$$I_n\left(h^{\otimes n}\right) = H_n(G(h)) = H_n(I_1(h)). \tag{5.11}$$

Proof. Of course, $H_1(I_1(h)) = I_1(h)$. By the multiplication formula (5.5), one has therefore that, for $n \geq 2$,

$$I_n\left(h^{\otimes n}\right) I_1(h) = I_{n+1}\left(h^{\otimes n+1}\right) + nI_{n-1}\left(h^{\otimes n-1}\right),$$

and the conclusion is obtained from (5.10), and by recursion on n. ■

Remark. By using the relation $\mathbb{E}\left[I_n(h^{\otimes n}) I_n(g^{\otimes n})\right] = n! \langle h^{\otimes n}, g^{\otimes n}\rangle_{L^2(\mu^n)} = n! \langle h, g\rangle^n_{L^2(\mu)}$, we infer from (5.11) that, for every jointly Gaussian random variables (U, V) with zero mean and unitary variance,

$$\mathbb{E}\left[H_n(U) H_m(V)\right] = \begin{cases} 0 & \text{if} \quad m \neq n \\ n!\mathbb{E}\left[UV\right]^n & \text{if} \quad m = n. \end{cases}$$

5.3 Chaotic decompositions

By combining (5.8) and (5.11), one obtains the following fundamental decomposition of the square-integrable functionals of G.

Theorem 5.2 (Chaotic decomposition) *For every $F \in L^2(\sigma(G), \mathbb{P})$ (that is, F is a square-integrable functional of G), there exists a unique sequence $\{f_n : n \geq 1\}$, with $f_n \in L^2_s(\mu^n)$, such that*

$$F = \mathbb{E}[F] + \sum_{n=1}^{\infty} I_n(f_n), \tag{5.12}$$

where the series converges in $L^2(\mathbb{P})$.

Proof. Fix $h \in L^2(\mu)$ such that $\|h\|_{L^2(\mu)} = 1$, as well as $t \in \mathbb{R}$. By using (5.8) and (5.11), one obtains that

$$\exp\left(tG(h) - \frac{t^2}{2}\right) = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(G(h)) = 1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} I_n\left(h^{\otimes n}\right). \tag{5.13}$$

Since $\mathbb{E}\left[\exp\left(tG(h)-\frac{t^2}{2}\right)\right]=1$, one deduces that (5.12) holds for every random variable of the form $F=\exp\left(tG(h)-\frac{t^2}{2}\right)$, with $f_n=\frac{t^n}{n!}h^{\otimes n}$. The conclusion is obtained by observing that the linear combinations of random variables of this type are dense in $L^2(\sigma(G),\mathbb{P})$. ■

Remarks. (1) Proposition 4.2, together with (5.12), implies that

$$\mathbb{E}\left[F^2\right]=\mathbb{E}[F]^2+\sum_{n=1}^{\infty}n!\,\|f_n\|^2_{L^2(\mu^n)}. \tag{5.14}$$

(2) By using the notation (4.6), (4.13) and (4.14), one can reformulate the statement of Theorem 5.2 as follows:

$$L^2(\sigma(G),\mathbb{P})=\bigoplus_{n=0}^{\infty}C_n(G),$$

where " $\oplus$ " indicates an infinite orthogonal sum in $L^2(\mathbb{P})$.

(3) By inspection of the proof of Theorem 5.2, we deduce that the linear combinations of random variables of the type $I_n(h^{\otimes n})$, with $n\geq 1$ and $\|h\|_{L^2(\mu)}=1$, are dense in $L^2(\sigma(G),\mathbb{P})$. This implies in particular that the random variables $I_n(h^{\otimes n})$ generate the nth Wiener chaos $C_n(G)$.

(4) The first proof of (5.12) dates back to Wiener [99]. See also McKean [50], Nualart and Schoutens [68] and Stroock [91]. See e.g. [19], [35], [39], [43] and [72] for further references and results on chaotic decompositions.

6 Isonormal Gaussian processes

In this section we briefly show how to generalize the previous results to the case of an *isonormal Gaussian process*. These objects have been introduced by Dudley in [22], and are a natural generalization of the Gaussian measures introduced above. In particular, the concept of an isonormal Gaussian process can be very useful in the study of fractional fields. See e.g. Pipiras and Taqqu [76, 77, 78], or the second edition of Nualart's book [65]. For a general approach to Gaussian analysis by means of Hilbert space techniques, and for further details on the subjects discussed in this section, the reader is referred to Janson [35].

6.1 General definitions and examples

Let $\mathfrak{H}$ be a real separable Hilbert space with inner product $\langle\cdot,\cdot\rangle_{\mathfrak{H}}$. In what follows, we will denote by

$$X=X(\mathfrak{H})=\{X(h):h\in\mathfrak{H}\}$$

an *isonormal Gaussian process* over $\mathfrak{H}$. This means that X is a centered real-valued Gaussian family, indexed by the elements of $\mathfrak{H}$ and such that

$$\mathbb{E}\left[X(h)X(h')\right]=\langle h,h'\rangle_{\mathfrak{H}},\quad\forall h,h'\in\mathfrak{H}. \tag{6.1}$$

In other words, relation (6.1) means that X is a centered Gaussian Hilbert space (with respect to the inner product canonically induced by the covariance) isomorphic to $\mathfrak{H}$.

Example (*Euclidean spaces*). Fix an integer $d \geq 1$, set $\mathfrak{H} = \mathbb{R}^d$ and let $(e_1, ..., e_d)$ be an orthonormal basis of $\mathbb{R}^d$ (with respect to the usual Euclidean inner product). Let $(Z_1, ..., Z_d)$ be a Gaussian vector whose components are i.i.d. $N(0,1)$. For every $h = \sum_{j=1}^d c_j e_j$ (where the c_j are real and uniquely defined), set $X(h) = \sum_{j=1}^d c_j Z_j$ and define $X = \left\{ X(h) : h \in \mathbb{R}^d \right\}$. Then, X is an isonormal Gaussian process over $\mathbb{R}^d$.

Example (*Gaussian measures*). Let $(Z, \mathcal{Z}, \mu)$ be a measure space, where μ is positive, σ-finite and non-atomic. Consider a completely random Gaussian measure $G = \{G(A) : A \in \mathcal{Z}_\mu\}$ (as defined in Section 3), where $\mathcal{Z}_\mu = \{A \in \mathcal{Z} : \mu(A) < \infty\}$. Set $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$ (thus, for every $h, h' \in \mathfrak{H}$, $\langle h, h' \rangle_{\mathfrak{H}} = \int_Z h(z) h'(z) \mu(dz)$) and, for every $h \in \mathfrak{H}$, define $X(h) = I_1(h)$ to be the Wiener-Itô integral of h with respect to G, as defined in (4.4). Recall that $X(h)$ is a centered Gaussian random variable with variance given by $\|h\|_{\mathfrak{H}}^2$. Then, relation (4.5) implies that the collection $X = \left\{ X(h) : h \in L^2(Z, \mathcal{Z}, \mu) \right\}$ is an isonormal Gaussian process over $L^2(Z, \mathcal{Z}, \mu)$.

Example (*Isonormal processes built from covariances*). Let $Y = \{Y_t : t \geq 0\}$ be a real-valued centered Gaussian process indexed by the positive axis, and set $R(s,t) = \mathbb{E}[Y_s Y_t]$ to be the covariance function of Y. Then, one can embed Y into some isonormal Gaussian process as follows: (i) define $\mathcal{E}$ as the collection of all finite linear combinations of indicator functions of the type $\mathbf{1}_{[0,t]}$, $t \geq 0$; (ii) define $\mathfrak{H} = \mathfrak{H}_R$ to be the Hilbert space given by the closure of $\mathcal{E}$ with respect to the inner product

$$\langle f, h \rangle_R := \sum_{i,j} a_i c_j R(s_i, t_j),$$

where $f = \sum_i a_i \mathbf{1}_{[0,s_i]}$ and $h = \sum_j c_j \mathbf{1}_{[0,t_j]}$ are two generic elements of $\mathcal{E}$; (iii) for $h = \sum_j c_j \mathbf{1}_{[0,t_j]} \in \mathcal{E}$, set $X(h) = \sum_j c_j Y_{t_j}$; (iv) for $h \in \mathfrak{H}_R$, set $X(h)$ to be the $L^2(\mathbb{P})$ limit of any sequence of the type $X(h_n)$, where $\{h_n\} \subset \mathcal{E}$ converges to h in $\mathfrak{H}_R$. Note that such a sequence $\{h_n\}$ necessarily exists and may not be unique (however, the definition of $X(h)$ does not depend on the choice of the sequence $\{h_n\}$). Then, by construction, the Gaussian space $\{X(h) : h \in \mathfrak{H}\}$ is an isonormal Gaussian process over $\mathfrak{H}_R$. See Janson [35, Ch. 1] or Nualart [65] for more details on this construction.

Example (*Even functions and symmetric measures*). Other classic examples of isonormal Gaussian processes are given by objects of the type

$$X_\beta = \{X_\beta(\psi) : \psi \in \mathfrak{H}_{E,\beta}\},$$

where β is a real non-atomic symmetric measure on $(-\pi, \pi]$ (that is, $\beta(dx) = \beta(-dx)$), and

$$\mathfrak{H}_{E,\beta} = L_E^2((-\pi, \pi], d\beta) \tag{6.2}$$

stands for the collection of *real* linear combinations of complex-valued *even* functions that are square-integrable with respect to β (recall that a complex-valued function ψ is even if $\overline{\psi(x)} = \psi(-x)$). The class $\mathfrak{H}_{E,\beta}$ is indeed a real separable Hilbert space, endowed with the inner product

$$\langle \psi_1, \psi_2 \rangle_\beta = \int_{-\pi}^{\pi} \psi_1(x) \psi_2(-x) \beta(dx) \in \mathbb{R}. \tag{6.3}$$

This type of construction is used in the spectral theory of time series, and is often realized by means of a complex-valued Gaussian measure (see e.g., [7, 27, 41, 92]) .

6.2 Hermite polynomials and Wiener chaos

We shall now show how to extend the notion of *Wiener chaos* to the case of an isonormal Gaussian process.

Definition 6.1 *From now on, the symbol $\mathcal{A}_\infty$ will denote the class of those sequences $\alpha = \{\alpha_i : i \geq 1\}$ such that: (i) each α_i is a nonnegative integer, (ii) α_i is different from zero only for a finite number of indices i. A sequence of this type is called a* ***multiindex.*** *For $\alpha \in \mathcal{A}_\infty$, we use the notation $|\alpha| = \sum_i \alpha_i$. For $q \geq 1$, we also write*

$$\mathcal{A}_{\infty,q} = \{\alpha \in \mathcal{A}_\infty : |\alpha| = q\}.$$

Remark on notation. Fix $q \geq 2$. Given a real separable Hilbert space $\mathfrak{H}$, we denote by $\mathfrak{H}^{\otimes q}$ and $\mathfrak{H}^{\odot q}$, respectively, the qth *tensor power* of $\mathfrak{H}$ and the qth *symmetric tensor power* of $\mathfrak{H}$ (see e.g. [35]). We conventionally set $\mathfrak{H}^{\otimes 1} = \mathfrak{H}^{\odot 1} = \mathfrak{H}$.

We recall four classic facts concerning tensors powers of Hilbert spaces (see e.g. [35]).

(I) The spaces $\mathfrak{H}^{\otimes q}$ and $\mathfrak{H}^{\odot q}$ are real separable Hilbert spaces, such that $\mathfrak{H}^{\odot q} \subset \mathfrak{H}^{\otimes q}$.

(II) Let $\{e_j : j \geq 1\}$ be an orthonormal basis of $\mathfrak{H}$; then, an orthonormal basis of $\mathfrak{H}^{\otimes q}$ is given by the collection of all tensors of the type

$$e_{j_1} \otimes \cdots \otimes e_{j_q}, \quad j_1, ..., j_d \geq 1.$$

(III) Let $\{e_j : j \geq 1\}$ be an orthonormal basis of $\mathfrak{H}$ and endow $\mathfrak{H}^{\odot q}$ with the inner product $(\cdot,\cdot)_{\mathfrak{H}^{\otimes q}}$; then, an orthogonal (and, in general, *not* orthonormal) basis of $\mathfrak{H}^{\odot q}$ is given by all elements of the type

$$\mathbf{e}(j_1, ..., j_q) = \mathbf{sym}\left\{e_{j_1} \otimes \cdots \otimes e_{j_q}\right\}, \quad 1 \leq j_1 \leq ... \leq j_q < \infty, \tag{6.4}$$

where $\mathbf{sym}\{\cdot\}$ stands for a canonical symmetrization.

(IV) If $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$, where μ is σ-finite and non-atomic, then $\mathfrak{H}^{\otimes q}$ can be identified with $L^2(Z^q, \mathcal{Z}^q, \mu^q)$ and $\mathfrak{H}^{\odot q}$ can be identified with $L^2_s(Z^q, \mathcal{Z}^q, \mu^q)$, where $L^2_s(Z^q, \mathcal{Z}^q, \mu^q)$ is the subspace of $L^2(Z^q, \mathcal{Z}^q, \mu^q)$ composed of symmetric functions.

Now observe that, once an orthonormal basis of $\mathfrak{H}$ is fixed and due to the symmetrization, each element $\mathbf{e}(j_1, ..., j_q)$ in (6.4) can be completely described in terms of a unique multiindex $\alpha \in \mathcal{A}_{\infty,q}$, as follows: (i) set $\alpha_i = 0$ if $i \neq j_r$ for every $r = 1, ..., q$, (ii) set $\alpha_j = k$ for every $j \in \{j_1, ..., j_q\}$ such that j is repeated exactly k times in the vector $(j_1, ..., j_q)$ ($k \geq 1$).

Examples. (i) The multiindex $(1, 0, 0, ...)$ is associated with the element of $\mathfrak{H}$ given by e_1.

(ii) Consider the element $\mathbf{e}(1,7,7)$. In $(1,7,7)$ the number 1 is not repeated and 7 is repeated twice, hence $\mathbf{e}(1,7,7)$ is associated with the multiindex $\alpha \in \mathcal{A}_{\infty,3}$ such that $\alpha_1 = 1$, $\alpha_7 = 2$ and $\alpha_j = 0$ for every $j \neq 1,7$, that is, $\alpha = (1,0,0,0,0,0,2,0,0,...)$.

(iii) The multindex $\alpha = (1,2,2,0,5,0,0,0,...)$ is associated with the element of $\mathfrak{H}^{\odot 10}$ given by $\mathbf{e}(1,2,2,3,3,5,5,5,5,5)$.

In what follows, given $\alpha \in \mathcal{A}_{\infty,q}$ $(q \geq 1)$, we shall write $\mathbf{e}(\alpha)$ in order to indicate the element of $\mathfrak{H}^{\odot q}$ uniquely associated with α.

Definition 6.2 *For every $h \in \mathfrak{H}$, we set $I_1^X(h) = I_1(h) = X(h)$. Now fix an orthonormal basis $\{e_j : j \geq 1\}$ of $\mathfrak{H}$: for every $q \geq 2$ and every $h \in \mathfrak{H}^{\odot q}$ such that*

$$h = \sum_{\alpha \in \mathcal{A}_{\infty,q}} c_\alpha e(\alpha)$$

(with convergence in $\mathfrak{H}^{\odot q}$, endowed with the inner product $\langle \cdot, \cdot \rangle_{\mathfrak{H}^{\otimes q}}$), we set

$$I_q^X(h) = I_q(h) = \sum_{\alpha \in \mathcal{A}_{\infty,q}} c_\alpha \prod_j H_{\alpha_j}(X(e_j)), \tag{6.5}$$

*where the products only involve the non-zero terms of each multiindex α, and H_m indicates the mth Hermite polynomial . For $q \geq 1$, the collection of all random variables of the type $I_q(h)$, $h \in \mathfrak{H}^{\odot q}$, is called the q**th Wiener chaos associated with** X and is denoted by $C_q(X)$. One sets conventionally $C_0(X) = \mathbb{R}$.*

Examples. (i) If $h = e(\alpha)$, where $\alpha = (1,1,0,0,0,...) \in \mathcal{A}_{\infty,2}$, then

$$I_2(h) = H_1(X(e_1)) H_1(X(e_2)) = X(e_1) X(e_2).$$

(ii) If $\alpha = (1,0,1,2,0,...) \in \mathcal{A}_{\infty,4}$, then

$$\begin{aligned} I_4(h) &= H_1(X(e_1)) H_1(X(e_3)) H_2(X(e_4)) \\ &= X(e_1) X(e_3) \left(X(e_4)^2 - 1\right) \\ &= X(e_1) X(e_3) X(e_4)^2 - X(e_1) X(e_3). \end{aligned}$$

(iii) If $\alpha = (3,1,1,0,0,...) \in \mathcal{A}_{\infty,5}$, then

$$\begin{aligned} I_5(h) &= H_3(X(e_1)) H_1(X(e_2)) H_1(X(e_3)) \\ &= \left(X(e_1)^3 - 3X(e_1)\right) X(e_2) X(e_3) \\ &= X(e_1)^3 X(e_2) X(e_3) - 3X(e_1) X(e_2) X(e_3). \end{aligned}$$

The following result collects some well-known facts concerning Wiener chaos and isonormal Gaussian processes. In particular: the first point characterizes the operators I_q^X as isomorphisms; the third point is an equivalent of the chaotic representation property for Gaussian measures, as stated in formula (5.12); the fourth point establishes a formal relation between random variables of the type $I_q^X(h)$ and the multiple Wiener-Itô integrals introduced in Section 4.2 (see [65, Ch. 1] for proofs and further discussions of all these facts).

Proposition 6.1 *1. For every $q \geq 1$, the qth Wiener chaos $C_q(X)$ is a Hilbert subspace of $L^2(\mathbb{P})$, and the application*

$$h \mapsto I_q(h), \quad h \in \mathfrak{H}^{\odot q},$$

defines a Hilbert space isomorphism between $\mathfrak{H}^{\odot q}$, endowed with the scalar product $q! \langle \cdot, \cdot \rangle_{\mathfrak{H}^{\otimes q}}$, and $C_q(X)$.

2. *For every $q, q' \geq 0$ such that $q \neq q'$, the spaces $C_q(X)$ and $C_{q'}(X)$ are orthogonal in $L^2(\mathbb{P})$.*

3. *Let F be a functional of the isonormal Gaussian process X satisfying $\mathbb{E}[F(X)^2] < \infty$: then, there exists a unique sequence $\{f_q : q \geq 1\}$ such that $f_q \in \mathfrak{H}^{\odot q}$, and*

$$F = \mathbb{E}(F) + \sum_{q=1}^{\infty} I_q(f_q) = \sum_{q=0}^{\infty} I_q(f_q), \tag{6.6}$$

where we have used the notation $I_0(f_0) = \mathbb{E}(F)$, and the series converges in $L^2(\mathbb{P})$.

4. *Suppose that $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$, where μ is σ-finite and non-atomic. Then, for $q \geq 2$, the symmetric power $\mathfrak{H}^{\odot q}$ can be identified with $L_s^2(Z^q, \mathcal{Z}^q, \mu^q)$ and, for every $f \in \mathfrak{H}^{\odot q}$, the random variable $I_q(f)$ coincides with the Wiener-Itô integral of f with respect to the Gaussian measure given by $A \mapsto X(\mathbf{1}_A)$, $A \in \mathcal{Z}_\mu$.*

Remark. The combination of Point 1 and Point 2 in the statement of Proposition 6.1 implies that, for every $q, q' \geq 1$,

$$\mathbb{E}\left[I_q(f)\, I_{q'}(f')\right] = \mathbf{1}_{q=q'} q! \langle f, f' \rangle_{\mathfrak{H}^{\otimes q}}.$$

From the previous statement, one also deduces the following Hilbert space isomorphism:

$$L^2(\sigma(X)) \simeq \bigoplus_{q=0}^{\infty} \mathfrak{H}^{\odot q}, \tag{6.7}$$

where $\simeq$ stands for a Hilbert space isomorphism, and each symmetric power $\mathfrak{H}^{\odot q}$ is endowed with the modified scalar product $q! \langle \cdot, \cdot \rangle_{\mathfrak{H}^{\otimes q}}$. The direct sum on the RHS of (6.7) is called the *symmetric Fock space* associated with $\mathfrak{H}$.

6.3 Contractions and products

We start by introducing the notion of *contraction* in the context of powers of Hilbert spaces.

Definition 6.3 *Consider a real separable Hilbert space $\mathfrak{H}$, and let $\{e_i : i \geq 1\}$ be an orthonormal basis of $\mathfrak{H}$. For every $n, m \geq 1$, every $r = 0, ..., n \wedge m$ and every $f \in \mathfrak{H}^{\odot n}$ and $g \in \mathfrak{H}^{\odot m}$, we define the contraction of order r, of f and g, as the element of $\mathfrak{H}^{\otimes n+m-2r}$ given by*

$$f \otimes_r g = \sum_{i_1, \ldots, i_r = 1}^{\infty} \langle f, e_{i_1} \otimes \cdots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}} \otimes \langle g, e_{i_1} \otimes \cdots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}}, \tag{6.8}$$

and we denote by $f \widetilde{\otimes}_r g$ its symmetrization.

Remark. One can prove the following result: if $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$, $f \in \mathfrak{H}^{\odot n} = L_s^2(Z^n, \mathcal{Z}^n, \mu^n)$ and $g \in \mathfrak{H}^{\odot m} = L_s^2(Z^m, \mathcal{Z}^m, \mu^m)$, then the definition of the contraction $f \otimes_r g$ given in (6.8) and the one given in (5.1) coincide.

The next result extends the product formula (5.5) to the case of isonormal Gaussian processes. The proof (which is left to the reader) can be obtained from Theorem 5.1, by using the fact that every real separable Hilbert space is isomorphic to a space of the type $L^2(Z, \mathcal{Z}, \mu)$, where μ is σ-finite and non-atomic.

Proposition 6.2 *Let X be an isonormal Gaussian process over some real separable Hilbert space $\mathfrak{H}$. Then, for every $n, m \geq 1$, $f \in \mathfrak{H}^{\odot n}$ and $g \in \mathfrak{H}^{\odot m}$,*

$$I_n(f) I_m(g) = \sum_{r=0}^{m \wedge n} r! \binom{m}{r} \binom{n}{r} I_{n+m-2r}\left(f \widetilde{\otimes}_r g\right), \tag{6.9}$$

where the symbol $(\widetilde{\ })$ indicates a symmetrization, the contraction $f \otimes_r g$ is defined in (6.8), and for $m = n = r$, we write

$$I_0\left(f \widetilde{\otimes}_n g\right) = \langle f, g \rangle_{\mathfrak{H}^{\otimes n}}.$$

7 A handful of operators from Malliavin calculus

We shall now describe some Malliavin-type operators, that turn out to be fundamental tools for the analysis to follow. For the sake of generality, we will first provide the definitions and the main properties in the case of an isonormal Gaussian process, and then specialize our discussion to Gaussian measures. Given an isonormal Gaussian process $X = \{X(h) : h \in \mathfrak{H}\}$, these operators involve the following real Hilbert spaces:

- $L^2(\sigma(X), \mathbb{P}) = L^2(\sigma(X))$ is the Hilbert space of real-valued integrable functionals of X, endowed with the usual scalar product $\langle F_1, F_2 \rangle_{L^2(\sigma(X))} = \mathbb{E}[F_1 \times F_2]$;
- For $k \geq 1$, $L^2\left(\sigma(X), \mathbb{P}; \mathfrak{H}^{\otimes k}\right) = L^2\left(\sigma(X); \mathfrak{H}^{\otimes k}\right)$ is the Hilbert space of $\mathfrak{H}^{\otimes k}$-valued functionals of X, endowed with the scalar product $\langle F_1, F_2 \rangle_{L^2\left(\sigma(X); \mathfrak{H}^{\otimes k}\right)} = \mathbb{E}\left[\langle F_1, F_2 \rangle_{\mathfrak{H}^{\otimes k}}\right]$.

In the particular case where $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$, the space $L^2\left(\sigma(X); \mathfrak{H}^{\otimes k}\right)$ can be identified with the class of stochastic processes $u(z_1, ..., z_k, \omega)$ that are $\mathcal{Z}^k \otimes \sigma(X)$ - measurable, and verify the integrability condition

$$\mathbb{E}\left[\int_{Z^k} u(z_1, ..., z_k)^2 \mu^k(dz_1, ..., dz_k)\right] < \infty. \tag{7.1}$$

Our presentation is voluntarily succinct and incomplete, as we prefer to focus on the computations and results that are specifically relevant for the interaction with Stein's method. It follows that the content of this section cannot replace the excellent discussions around Malliavin calculus that one can find in the probabilistic literature: see e.g. Janson [35], Malliavin [43] and Nualart [65].

In what follows, we shall also use the following notation: for every $n \geq 1$,

- $C_p^\infty(\mathbb{R}^n)$ is the class of infinitely differentiable functions f on $\mathbb{R}^n$ such that f and its derivatives have polynomial growth;
- $C_b^\infty(\mathbb{R}^n)$ is the class of infinitely differentiable functions f on $\mathbb{R}^n$ such that f and its derivatives are bounded;
- $C_0^\infty(\mathbb{R}^n)$ is the class of infinitely differentiable functions f on $\mathbb{R}^n$ such that f has compact support.

7.1 Derivatives

7.1.1 Definition and characterization of the domain

Let $X = \{X(h) : h \in \mathfrak{H}\}$ be an isonormal Gaussian process. The Malliavin derivative operator of order k transforms elements of $L^2(\sigma(X))$ into elements of $L^2\left(\sigma(X); \mathfrak{H}^{\otimes k}\right)$. Formally, one starts by defining the class $\mathcal{S}(X) \subset L^2(\sigma(X))$ of *smooth* functionals of X, as the collection of random variables of the type

$$F = f(X(h_1), ..., X(h_m)), \tag{7.2}$$

where $f \in C_p^\infty(\mathbb{R}^n)$ and $h_i \in \mathfrak{H}$.

Definition 7.1 *Let $F \in \mathcal{S}(X)$ be as in (7.2).*

1. *The derivative DF of F is the $\mathfrak{H}$-valued random element given by*

$$DF = \sum_{i=1}^{m} \frac{\partial}{\partial x_i} f(X(h_1), ..., X(h_m))\, h_i; \tag{7.3}$$

we shall sometimes use the notation $DF = D^1F$.

2. *For $k \geq 2$, the kth derivative of F, denoted by D^kF, is the element of $L^2\left(\sigma(X); \mathfrak{H}^{\otimes k}\right)$ given by*

$$D^kF = \sum_{i_1,...,i_k=1}^{k} \frac{\partial^k}{\partial x_{i_1} \cdots \partial x_{i_k}} f(X(h_1), ..., X(h_m))\, h_{i_1} \otimes \cdots \otimes h_{i_k}. \tag{7.4}$$

Example. Let $h \in \mathfrak{H}$ be such that $\|h\|_{\mathfrak{H}} = 1$. Then, for every $q \geq 1$, $I_q(h) = H_q(X(h))$, where H_q is the qth Hermite polynomial, and one has therefore that

$$D^k I_q(h) = \begin{cases} q(q-1)\cdots(q-k+1)\, H_{q-k}(X(h))\, h^{\otimes k}, & \text{if } k \leq q \\ 0, & \text{if } k > q. \end{cases}$$

In particular, $DX(h) = h$.

Remarks. (a) The polynomial growth condition implies that for every $F \in \mathcal{S}(X)$, every $k \geq 1$ and every $h \in \mathfrak{H}^{\otimes k}$, the real-valued random variables F and $\left\langle D^kF, h\right\rangle_{\mathfrak{H}^{\otimes k}}$ have finite moments of all orders.

(b) If, $F, J \in \mathcal{S}(X)$, then $FJ \in \mathcal{S}(X)$ and

$$D(FJ) = J \times DF + F \times DJ. \tag{7.5}$$

Proposition 7.1 *The operator $D^m : \mathcal{S}(X) \to L^2\left(\sigma(X); \mathfrak{H}^{\otimes k}\right)$ is closable.*

Proof. It is interesting to provide a complete proof of this result in the case $k = 1$, since this involves the use (1.1) (the case $k \geq 2$ is analogous). All we have to prove is that, if $F_N \in \mathcal{S}(X)$ is a sequence converging to zero in $L^2(\mathbb{P})$ and DF_N converges to η in $L^2(\sigma(X); \mathfrak{H})$, then necessarily $\eta = 0$.

We start by observing that for every $F \in \mathcal{S}(X)$ and $h \in \mathfrak{H}$, one has the following integration by parts formula:

$$\mathbb{E}\left[\langle DF, h\rangle_{\mathfrak{H}}\right] = \mathbb{E}\left[FX(h)\right]. \tag{7.6}$$

To prove (7.6), first observe that we can always assume (without loss of generality) that $\|h\|_{\mathfrak{H}} = 1$, and that F has the form $F = f(X(h_1), X(h_2), ..., X(h_m))$, where $h_1 = h$ and $h_1, h_2..., h_m$ is an orthonormal system in $\mathfrak{H}$. Now write $\Theta(x) = \mathbb{E}[F \mid X(h) = x]$, denote by Θ' the first derivative of Θ with respect to x, and use (1.1) to obtain that

$$\begin{aligned}
\mathbb{E}[FX(h)] &= \mathbb{E}[\Theta(X(h))X(h)] = \mathbb{E}\left[\Theta'(X(h))\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\frac{\partial}{\partial x_1} f(X(h), ..., X(h_m)) \mid X(h)\right]\right] \\
&= \mathbb{E}\left[\frac{\partial}{\partial x_1} f(X(h), ..., X(h_m))\right] = \mathbb{E}\left[\sum_{i=1}^{m} \frac{\partial}{\partial x_i} f(X(h_1), ..., X(h_m)) \langle h_i, h\rangle_{\mathfrak{H}}\right] \\
&= \mathbb{E}\left[\langle DF, h\rangle_{\mathfrak{H}}\right],
\end{aligned}$$

where we have used the fact that $h_1 = h$ and $h_1, ..., h_m$ is an orthonormal system. By considering two smooth functionals $F, J \in \mathcal{S}(X)$, and by using (7.5), we infer from (7.6) that

$$\mathbb{E}\left[J\langle DF, h\rangle_{\mathfrak{H}}\right] = -\mathbb{E}\left[F\langle DJ, h\rangle_{\mathfrak{H}}\right] + \mathbb{E}\left[FJX(h)\right]. \tag{7.7}$$

We now go back to the variables F_N, $N \geq 1$, and η, as defined at the beginning of the proof. Fix $h \in \mathfrak{H}$ and $F \in \mathcal{S}(X)$. By using the fact that F and $\langle DF, h\rangle_{\mathfrak{H}}$ have finite moments of all orders and by exploiting (7.7) in the case $J = F_N$, we have that

$$\begin{aligned}
\left|\mathbb{E}\left[F\langle \eta, h\rangle_{\mathfrak{H}}\right]\right| &= \lim_{N\to\infty} \left|\mathbb{E}\left[F\langle DF_N, h\rangle_{\mathfrak{H}}\right]\right| \\
&\leq \lim_{N\to\infty} \left|\mathbb{E}\left[F_N\langle DF, h\rangle_{\mathfrak{H}}\right]\right| + \lim_{N\to\infty} \left|\mathbb{E}\left[F_N F X(h)\right]\right| \\
&= 0,
\end{aligned}$$

which gives $\langle \eta, h\rangle_{\mathfrak{H}} = 0$, a.s.-$\mathbb{P}$. Since h is arbitrary and $\mathfrak{H}$ is separable, we deduce the desired conclusion. ■

Definition 7.2 *For every $k \geq 1$, the* ***domain*** *of the operator D^k in $L^2(\sigma(X))$, customarily denoted by $\mathbb{D}^{k,2}$, is the closure of the class $\mathcal{S}(X)$ with respect to the seminorm*

$$\|F\|_{k,2} = \left[\mathbb{E}\left(F^2\right) + \sum_{j=1}^{k} \left\|D^j F\right\|^2_{\mathfrak{H}^{\otimes j}}\right]^{\frac{1}{2}}. \tag{7.8}$$

We also set

$$\mathbb{D}^{\infty,2} = \bigcap_{k=1}^{\infty} \mathbb{D}^{k,2} \tag{7.9}$$

Remark. One can actually define more general domains $\mathbb{D}^{k,p}$, $p \geq 1$, as the closure of D^k in $L^p(\sigma(X))$. See [65, pp. 26–27].

The following result provides an important characterization of $\mathbb{D}^{k,2}$, namely that $F \in \mathbb{D}^{k,2}$ if and only if the norms of its chaotic projections decrease sufficienlty fast. The proof can be found in [65, Proposition 1.2.2], and uses the representation (6.5).

Proposition 7.2 *Fix $k \geq 1$. A random variable $F \in L^2(\sigma(X))$ with a chaotic representation (6.6) is an element of $\mathbb{D}^{k,2}$ if and only if the kernels $\{f_q\}$ verify*

$$\sum_{q=1}^{\infty} q^k q! \, \|f_q\|^2_{\mathfrak{H}^{\otimes q}} < \infty, \tag{7.10}$$

and in this case

$$\mathbb{E}\left\|D^k F\right\|^2_{\mathfrak{H}^{\otimes k}} = \sum_{q=k}^{\infty} (q)_k \times q! \, \|f_q\|^2_{\mathfrak{H}^{\otimes q}} \, ,$$

where $(q)_k = q(q-1)\cdots(q-k+1)$ is the Pochammer symbol.

Note that the previous result implies that random variables belonging to a finite sum of Wiener chaoses are in $\mathbb{D}^{k,2}$, for every $k \geq 1$ (they are actually in $\mathbb{D}^{k,p}$, for every k, $p \geq 1$)

7.1.2 The case of Gaussian measures

We now focus on the case where $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$, so that each symmetric power $\mathfrak{H}^{\odot q}$ can be identified with the space of symmetric functions $L^2_s(\mu^q)$, and the integrals $I_q(f)$, $f \in L^2_s(\mu^q)$, are just (multiple) Wiener-Itô integrals of f with respect to the Gaussian measure $G(A) = X(\mathbf{1}_A)$, where $\mu(A) < \infty$.

As already observed in this case the derivative $D^k F$ of $F \in \mathbb{D}^{k,2}$ takes the form of a stochastic process

$$(z_1, ..., z_k) \mapsto D^k_{z_1,...,z_k} F,$$

verifying moreover

$$\mathbb{E}\left[\int_{Z^k} \left(D^k_{z_1,...,z_k} F\right)^2 \mu(dz_1, ..., dz_k)\right] < \infty.$$

The following statement provides a neat algorithm allowing to deduce the explicit form of the first derivative of a general random variable in $\mathbb{D}^{1,2}$. This result can be proved by first focusing on smooth random variables of the type (4.11), and then by using (7.3) as well as a density argument. Note that a similar statement (that one can deduce by recursion – **Exercise!**) holds also for derivatives of order greater than 1.

Proposition 7.3 *Suppose that $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$, and assume that $F \in \mathbb{D}^{1,2}$ admits the chaotic expansion (5.12). Then, a version of the derivative $DF = \{D_z F : z \in Z\}$ is given by*

$$D_z F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, z)), \quad z \in Z,$$

where, for each n and z, the integral $I_{n-1}(f_n(\cdot, z))$ is obtained by integrating the function on Z^{n-1} given by $(z_1, ..., z_{n-1}) \mapsto f_n(z_1, ..., z_{n-1}, z)$.

7.1.3 Remarkable formulae

We now state, without proofs, four important formulae involving Malliavin derivatives. The first three are called "chain rules" and allow to differentiate random variables that are smooth transformations of differentiable functionals (the proof is based on approximation arguments – see [65, Proposition 1.2.3 and Proposition 1.2.4]). The fourth result has been proved by Stroock in [91], and it is often a very useful tool in order to deduce the chaotic decomposition of a given functional (see also McKean [50]). Finally, we point out some computations related to maxima of Gaussian processess: for instance, this result is one of the staples of the recent remarkable paper by Nourdin and Viens [64].

Chain rule #1. Let $\varphi : \mathbb{R}^m \to \mathbb{R}$ be a continuosly differentiable function with bounded partial derivatives. Assume that $F = (F_1, ..., F_m)$ is a vector of elements of $\mathbb{D}^{1,2}$. Then, $\varphi(F) \in \mathbb{D}^{1,2}$, and

$$D\varphi(F) = \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \varphi(F) DF_i. \tag{7.11}$$

Note that (7.11) is consistent with (7.3).

Chain rule #2. Let $\varphi : \mathbb{R}^m \to \mathbb{R}$ be Lipschitz. Assume that $F = (F_1, ..., F_m)$ is a vector of elements of $\mathbb{D}^{1,2}$ such that the law of F *is absolutely continuous on* $\mathbb{R}^m$. Then, $\varphi(F) \in \mathbb{D}^{1,2}$, and formula (7.11) holds.

Chain rule #3. Let F be a finite sum of multiple stochastic integrals. Then, the multiplication formula (6.9) implies that $F^n \in \mathbb{D}^{1,2}$ for every $n \geq 1$, and moreover

$$D(F^n) = nF^{n-1} DF. \tag{7.12}$$

Stroock formula. Suppose that $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$, and assume that $F \in \mathbb{D}^{\infty,2}$ (see (7.9)) admits the chaotic expansion (5.12). Then,

$$f_n(z_1, ..., z_n) = \frac{1}{n!} \mathbb{E}\left[D^n_{z_1,...,z_n} F\right], \quad n \geq 1. \tag{7.13}$$

For instance, consider $F = \exp\left(tX(h) - \frac{t^2}{2}\right)$, where $\|h\|_{L^2(\mu)} = 1$. Then, $\mathbb{E}[F] = 1$,

$$\begin{aligned} D^n_{z_1,...,z_n} F &= t^n h^{\otimes n}(z_1, ..., z_n) F, \quad \text{and} \\ \mathbb{E}\left[D^n_{z_1,...,z_n} F\right] &= t^n h^{\otimes n}(z_1, ..., z_n). \end{aligned}$$

By using (7.13) we therefore obtain an alternate proof of formula (5.13).

Maxima. Suppose that $X(h_i)$, $i = 1, ..., m$ and $h_i \in \mathfrak{H}$, is a finite subset of some isonormal Gaussian process, such that the span of h_i has dimension m. Consider

$$F = \max_{i=1,\ldots,m} X(h_i).$$

Then, $F \in \mathbb{D}^{1,2}$ (indeed, $(z_1, ..., z_m) \mapsto \max z_i$ is Lipschitz), the random variable

$$I_0 = \arg\max_{i=1,\ldots,m} X(h_i)$$

is well defined, and one has that

$$DF = h_{I_0}.$$

This kind of results can also be extended to continuous-time Gaussian processes. For isntance, if $W = \{W_t : t \in [0,1]\}$ is a standard Brownian motion initialized at zero, then $M = \sup_{t\in[0,1]} W_t \in \mathbb{D}^{1,2}$, and

$$D_t M = \mathbf{1}_{[0,T]}(t),$$

where T is the unique random point where W attains its maximum.

7.2 Divergences

7.2.1 Definition and characterization of the domain

Let $X = \{X(h) : h \in \mathfrak{H}\}$ be an isonormal Gaussian process over some real separable Hilbert space $\mathfrak{H}$. We will now study the *divergence operator* δ, which is defined as the adjoint of the derivative D. Recall that D is a closed and unbounded operator from $\mathbb{D}^{1,2} \subset L^2(\sigma(X))$ into $L^2(\sigma(X); \mathfrak{H})$, so that the domain of the operator δ will be some suitable subset of $L^2(\sigma(X); \mathfrak{H})$.

Definition 7.3 *The* ***domain*** *of the divergence operator δ, denoted by* $\mathrm{dom}(\delta)$*, is the collection of all random elements $u \in L^2(\sigma(X); \mathfrak{H})$ such that, for every $F \in \mathbb{D}^{1,2}$,*

$$\left|\mathbb{E}\left[\langle u, DF\rangle_{\mathfrak{H}}\right]\right| \leq c\mathbb{E}\left[F^2\right]^{1/2}, \tag{7.14}$$

where c is a constant depending on u (and not on F). For every $u \in \mathrm{dom}(\delta)$, the random variable $\delta(u)$ is therefore defined as the unique element of $L^2(\sigma(X))$ verifying

$$\mathbb{E}\left[\langle u, DF\rangle_{\mathfrak{H}}\right] = \mathbb{E}\left[F\delta(u)\right], \tag{7.15}$$

for every $F \in \mathbb{D}^{1,2}$ (note that the existence of $\delta(u)$ is ensured by (7.14) and by the Riesz Representation Theorem). Relation (7.15) is called an ***integration by parts formula.***

Remark. By selecting $F = 1$ in (7.15), one deduces that $\mathbb{E}[\delta(u)] = 0$, for every $u \in \mathrm{dom}(\delta)$.

Example. Fix $h \in \mathfrak{H}$. Since, by Cauchy-Schwarz, $\left|\mathbb{E}\left[\langle h, DF\rangle_{\mathfrak{H}}\right]\right| \leq \|h\|_{\mathfrak{H}}\,\mathbb{E}\left[F^2\right]^{1/2}$, we deduce that $h \in \mathrm{dom}(\delta)$ and also, thanks to (7.6), that $\delta(h) = X(h)$.

7.2.2 The case of Gaussian measures

We now consider the case $\mathfrak{H} = (Z, \mathcal{Z}, \mu)$, where $(Z, \mathcal{Z})$ is a Polish space endowed with a σ-finite and non-atomic measure μ. Recall that the aplication $A \mapsto X(\mathbf{1}_A) = G(A)$ defines a Gaussian measure with control μ. In this case, the random variable $\delta(u)$ is called the *Skorohod integral* of u with respect to G. As already observed, in this framework the space $L^2(\sigma(X); \mathfrak{H})$ can be identified with the class of stochastic processes $u(z, \omega)$ that are $\mathcal{Z} \otimes \sigma(X)$ - measurable, and verify the integrability condition

$$\mathbb{E}\left[\int_{Z^k} u(z)^2 \mu(dz)\right] < \infty. \tag{7.16}$$

By combining (7.16) with (5.12) (and some standard measurability arguments) we infer that every $u \in L^2(\sigma(X); \mathfrak{H})$ admits a representation of the type

$$u(z) = h_0(z) + \sum_{n=1}^{\infty} I_n(h_n(\cdot, z)), \tag{7.17}$$

where $h_0 \in L^2(\mu)$ and, for every $n \geq 1$, h_n is a function on Z^{n+1} which is symmetric in the first n variables, and moreover

$$\mathbb{E}\left[\int_{Z^k} u(z)^2 \mu(dz)\right] = \sum_{n=0}^{\infty} n! \left\|h_n\right\|^2_{L^2(\mu^{n+1})} < \infty. \tag{7.18}$$

The next result provides a characterization of the operator δ as well as of its domain, in terms of chaotic decompositions. The proof can be found in [65, Section 1.3.2].

Proposition 7.4 *Let $\mathfrak{H} = (Z, \mathcal{Z}, \mu)$ as above, and let $u \in L^2(\sigma(X); \mathfrak{H})$ verify (7.16)–(7.18). Then, $u \in \mathrm{dom}(\delta)$ if and only if*

$$\sum_{n=0}^{\infty} (n+1)! \left\|\widetilde{h}_n\right\|^2_{L^2(\mu^{n+1})} < \infty, \tag{7.19}$$

where $\widetilde{h}_n$ indicates the canonical symmetrization of h_n. In this case, one has moreover that

$$\delta(u) = \sum_{n=0}^{\infty} I_{n+1}\left(\widetilde{h}_n\right),$$

where, thanks to (7.19), the series converges in $L^2(\mathbb{P})$.

Examples. (1) Suppose $\mu(Z) < \infty$, and let $u(z) = X(\mathbf{1}_Z)\mathbf{1}_Z(z) = G(Z)\mathbf{1}_Z(z)$. Then, $\delta(u) = I_2(\mathbf{1}_Z \otimes \mathbf{1}_Z) = G(Z)^2 - \mu(Z)$.

(2) Suppose $\mu(Z) < \infty$, let $Z_0 \subset Z$ and define $u(z) = X(\mathbf{1}_Z)\mathbf{1}_{Z_0}(z) = G(Z)\mathbf{1}_{Z_0}(z)$. Then, $\delta(u) = 2^{-1} I_2(\mathbf{1}_Z \otimes \mathbf{1}_{Z_0} + \mathbf{1}_{Z_0} \otimes \mathbf{1}_Z)$.

Remark. Suppose that $(Z, \mathcal{Z}, \mu) = ([0,1], \mathcal{B}([0,1]), dt)$, where dt stands for the Lebesgue measure, and write W_t, $t \in [0,1]$, to indicate the standard Brownain motion $t \mapsto G([0,t])$. We denote by $\mathcal{F}_t$ the filtration generated by W and by the $\mathbb{P}$-null sets of $\sigma(G)$, and we say that a stochastic process $u(t,\omega)$ is *adapted*, if $u(t) \in \mathcal{F}_t$ for every $t \in [0,1]$. If u is adapted and $\mathbb{E}\left(\int_0^1 u(t)^2\, dt\right) < \infty$, then the Itô stochastic integral $\int_0^1 u(t)\, dW_t$ is a well-defined element of $L^2(\sigma(X))$. Moreover, in this case one has that

$$\delta(u) = \int_0^1 u(t)\, dW_t$$

(see [65, Proposition 1.3.11]).

7.2.3 A formula on products

We conclude with a general (useful) formula involving products of Malliavin differentiable random variables and elements of $\mathrm{dom}(\delta)$. The framework is that of a general isonormal process $X = \{X(h) : h \in \mathfrak{H}\}$.

Proposition 7.5 *Let $F \in \mathbb{D}^{1,2}$ and $u \in \mathrm{dom}(\delta)$ be such that: (i) $Fu \in L^2(\sigma(X);\mathfrak{H})$, (ii) $F\delta(u) \in L^2(\sigma(X))$, and (iii) $\langle DF, u\rangle_{\mathfrak{H}} \in L^2(\sigma(X))$. Then, $Fu \in \mathrm{dom}(\delta)$, and also*

$$\delta(Fu) = F\delta(u) - \langle DF, u\rangle_{\mathfrak{H}}. \tag{7.20}$$

Proof. Consider a random variable G equal to the RHS (7.2), with $f \in C_0^\infty(\mathbb{R}^m)$. Then,

$$\begin{aligned} \mathbb{E}\left[\langle DG, Fu\rangle_{\mathfrak{H}}\right] &= \mathbb{E}\left[\langle FDG, u\rangle_{\mathfrak{H}}\right] = \mathbb{E}\left[\langle D(FG) - GDF, u\rangle_{\mathfrak{H}}\right] \\ &= \mathbb{E}\left[\left(F\delta(u) - \langle DF, u\rangle_{\mathfrak{H}}\right) G\right]. \end{aligned}$$

Since random variables such as G generate $\sigma(X)$, the conclusion is obtained. ■

7.3 The Ornstein-Uhlenbeck Semigroup and Mehler's formula

7.3.1 Definition, Mehler's formula and vector-valued Markov processes

Let $X = \{X(h) : h \in \mathfrak{H}\}$ be an isonormal Gaussian process over some real separable Hilbert space $\mathfrak{H}$.

Definition 7.4 *The* ***Ornstein-Uhlenbeck semigroup*** *$\{T_t : t \geq 0\}$ is the set of contraction operators defined as*

$$T_t(F) = \mathbb{E}(F) + \sum_{q=1}^{\infty} e^{-qt} I_q(f_q) = \sum_{q=0}^{\infty} e^{-qt} I_q(f_q), \tag{7.21}$$

for every $t \geq 0$ and every $F \in L^2(\sigma(X))$ as in (6.6).

The Ornstein-Uhlenbeck semigroup plays a fundamental role in our theory. Its relevance for Stein's method is not new: see for instance the so-called "Barbour-Götze generator approach", introduced in [2] and [29] (see [80] for a survey). As another example, see [57], [62] and the discussion contained in Section 9, where it is shown that the use of the semigroup $\{T_t\}$ leads to infinite-dimensional generalizations of the *second order Stein/Poincaré inequalities* proved by Chatterjee in [9]. Another striking connection between Stein's method and the Ornstein-Uhlenbeck semigroup will be exploited in Section 9.2, where we will use the properties of the generator of $\{T_t\}$ in order to provide a proof of a *multi-dimensional Stein's Lemma* which is completely based on Malliavin calculus.

We shall now present two alternative representations of the operators T_t. The first one is known as *Mehler's formula*, and provides a mixture-type characterization of the operator T_t. The second implies that the semigroup $\{T_t\}$ is indeed associated with a Markov process with values in $\mathbb{R}^{\mathfrak{H}}$.

First representation: Mehler's formula. We consider an independent copy of X, noted X', and we suppose that the two isonormal processes X and X' are defined on the same probability space. Note that X and X' are indeed random elements with values in $\mathbb{R}^{\mathfrak{H}}$ (the space of real-valued functions on $\mathfrak{H}$), and that every random variable $F \in L^2(\sigma(X))$ can be indeed identified with a $\sigma(X)$-measurable mapping $F : \mathbb{R}^{\mathfrak{H}} \to \mathbb{R}$, which is uniquely defined up to elements of $\sigma(X)$ with $\mathbb{P}$-measure zero. We now fix $t \geq 0$ and consider the process

$$Z_t(h) = e^{-t} X(h) + \sqrt{1 - e^{-2t}} X'(h), \quad h \in \mathfrak{H}.$$

It is clear that Z_t is another isonormal process over $\mathfrak{H}$, and therefore $Z_t \overset{\text{Law}}{=} X$. Given $F \in L^2(\sigma(X))$, we can therefore meaningfully consider the random variable $F(Z_t) = F\left(e^{-t}X + \sqrt{1-e^{-2t}}X'\right)$, obtained by applying to Z_t a version of the mapping (from $\mathbb{R}^{\mathfrak{H}}$ into $\mathbb{R}$) associated with F. Now write, for $t \geq 0$,

$$\Theta_t F(x) = \mathbb{E}[F(Z_t) \mid X = x], \quad x \in \mathbb{R}^{\mathfrak{H}}, \tag{7.22}$$

and let the class of operators $\{T_t\}$ be defined as in (7.21). The following representation of $\{T_t\}$ is known as *Mehler's formula*: for every $t \geq 0$ and every $F \in L^2(\sigma(X))$,

$$\Theta_t F(X) = T_t(F), \tag{7.23}$$

or, equivalently,

$$T_t(F) = \mathbb{E}\left[F\left(e^{-t}a + \sqrt{1-e^{-2t}}X'\right)\right]\Big|_{a=X}. \tag{7.24}$$

To prove (7.23), one should first observe that, for every $t \geq 0$, one has that

$$\mathbb{E}\left(\Theta_t F(X)^2\right) \leq \mathbb{E}\left(F^2\right) \quad \text{and} \quad \mathbb{E}\left(T_t(F)^2\right) \leq \mathbb{E}\left(F^2\right),$$

that is, both Θ_t and T_t are contraction operators from $L^2(\sigma(X))$ into itself. By a density argument, it is now sufficient to verify that $\Theta_t F$ and $T_t F$ agree for every random variable of the type $F = \exp\left(uX(h) - \frac{u^2}{2}\right)$, where $u \in \mathbb{R}$ and $\|h\|_{\mathfrak{H}} = 1$. Indeed, from (5.13) and (7.21) we infer that

$$T_t(F) = 1 + \sum_{q=1}^{\infty} e^{-qt} \frac{u^q}{q!} I_q\left(h^{\otimes q}\right);$$

on the other hand, by using (5.8), (5.11) and (6.5),

$$\begin{aligned}\Theta_t F(X) &= \exp\left(e^{-t}uX(h) - \frac{u^2}{2}e^{-2t}\right) = \sum_{q=0}^{\infty} \frac{e^{-qt}u^q}{q!} H_q(X(h)) \\ &= 1 + \sum_{n=1}^{\infty} e^{-qt}\frac{u^q}{q!} I_q\left(h^{\otimes q}\right),\end{aligned}$$

yielding the desired conclusion.

Second representation: vector-valued Markov process. We now give a representation of $\{T_t\}$ as the semigroup associated with a Markov process with values in $\mathbb{R}^{\mathfrak{H}}$. To do this, we consider an auxiliary isonormal Gaussian process B over $\widehat{\mathfrak{H}} = \mathfrak{H} \otimes L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), 2dx)$ (note the factor 2), where dx stands for the Lebesgue measure. Also, for $t \geq 0$ we denote by e_t the element of $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), 2dx)$ given by the mapping $x \mapsto e^{-(t-x)}\mathbf{1}_{x<t}$. We define a process $X_t(h)$, on $\mathbb{R}_+ \times \mathfrak{H}$, as

$$X_t(h) = B(h \otimes e_t), \quad t \geq 0, \quad h \in \mathfrak{H}.$$

We easily verify that: (i) for every fixed $h \in \mathfrak{H}$, the process $t \mapsto X_t(h)$ is a centered Gaussian process with covariance function $\mathbb{E}(X_t(h) X_s(h)) = \exp(-|t-s|)\|h\|^2_{\mathfrak{H}}$, that is, $t \mapsto X_t(h)$ is a *real-valued Ornstein-Uhlenbeck process* with parameters 1 and $\|h\|^2_{\mathfrak{H}}$, and (ii) for every fixed $t \geq 0$, the Gaussian family $X_t = \{X_t(h) : h \in \mathfrak{H}\}$ defines an isonormal Gaussian process over $\mathfrak{H}$ (that is, $X_t \overset{\text{Law}}{=} X$). Now fix $F \in L^2(\sigma(X))$, and consider the previously described associated mapping $F : \mathbb{R}^{\mathfrak{H}} \to \mathbb{R}$. One can verify the following (see [65, p. 57]) alternative representation of $\{T_t\}$: for every $t, s \geq 0$,

$$\mathbb{E}[F(X_{t+s}) \mid X_u(h) : u \leq s,\ h \in \mathfrak{H}] = T_t(F)(X_s).$$

Note that in the previous formula we have identified $T_t(F)$ with a suitable mapping from $\mathbb{R}^{\mathfrak{H}}$ into $\mathbb{R}$.

The reader is also referred to the paper by Meyer [51] for further discussions around the Ornstein-Uhlenbeck semigroup.

7.3.2 The generator of the Ornstein-Uhlenbeck semigroup and its inverse

We shall now define the operator L, known as the *generator of the Ornstein-Uhlenbeck semigroup*, in the framework of an isonormal Gaussian process $X = \{X(h) : h \in \mathfrak{H}\}$.

Definition 7.5 *Let $F \in L^2(\sigma(X))$ admit the representation (6.6). We define the operator L as*

$$LF = -\sum_{q=0}^{\infty} qI_q(f_q), \tag{7.25}$$

provided the previous series converges in $L^2(\mathbb{P})$. *This implies that the domain of* L, *denoted by* $\mathrm{dom}(L)$, *is given by*

$$\mathrm{dom}(L) = \left\{ F \in L^2(\sigma(X)), F = \sum_{q=0}^{\infty} I_q(f_q) : \sum_{q=1}^{\infty} q^2 q! \|f_q\|^2_{\mathfrak{H}^{\otimes q}} < \infty \right\}. \tag{7.26}$$

The following result is proved [65, Proposition 1.4.2].

Proposition 7.6 *Let* $\{T_t\}$ *be given by (7.21). For every* $F \in L^2(\sigma(X))$, *the following two statements are equivalent.*

1. $F \in \mathrm{dom}(L)$.
2. *As* $t \to 0$, $t^{-1}(T_t(F) - F)$ *converges in* $L^2(\mathbb{P})$, *and the limit equals the series appearing on the RHS of (7.25).*

It follows that L *is the* ***infinitesimal generator of the Ornstein-Uhlenbeck semigroup*** $\{T_t\}$.

Now note that the image of L coincides with the set

$$L_0^2(\sigma(X)) = \{F \in L^2(\sigma(X)) : \mathbb{E}(F) = 0\},$$

and also that $LF = L(F - \mathbb{E}(F))$. This last property implies that the mapping $L : L^2(\sigma(X)) \to L_0^2(\sigma(X))$ *is not injective*. It is nonetheless possible to define the application $L^{-1} : L_0^2(\sigma(X)) \to L_0^2(\sigma(X))$, which is the inverse mapping of the restriction of L to the set $L_0^2(\sigma(X))$.

Definition 7.6 *Let* $F \in L_0^2(\sigma(X))$ *admit the representation (6.6), with* $\mathbb{E}(F) = I_0(f_0) = 0$. *We define the operator* L^{-1} *as*

$$L^{-1}F = -\sum_{q=1}^{\infty} \frac{1}{q} I_q(f_q). \tag{7.27}$$

Note that the series on the RHS of (7.27) is convergent in $L^2(\mathbb{P})$ for every $F \in L_0^2(\sigma(X))$. The following property is easily verified: for every $F \in L_0^2(\sigma(X))$, one has that $L^{-1}F \in \mathrm{dom}(L)$, and $LL^{-1}F = F$ (L^{-1} is sometimes called the *pseudo-inverse* of L – see e.g. [53]).

7.4 The connection between δ, D and L: first consequences

Let $X = \{X(h) : h \in \mathfrak{H}\}$ be an isonormal Gaussian process. The following result provides a neat connection between the three operators D, δ and L, and is the actual "bridge" between Stein's method and Malliavin calculus. Needless to say, it is one of the staples of the analysis to follow.

Theorem 7.1 *For every $F \in L^2(\sigma(X))$, one has that $F \in \mathrm{dom}(L)$ if and only if $F \in \mathbb{D}^{1,2}$ and $DF \in \mathrm{dom}(\delta)$. In this case, one has moreover that*

$$\delta DF = -LF. \tag{7.28}$$

Proof. It is enough to prove this result for $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$, where μ is σ-finite and non-atomic. In this case, the first part of the statement is easily proved by using the characterizations of $\mathbb{D}^{1,2}$ and $\mathrm{dom}(\delta)$ given respectively in (7.10) (for $k = 1$) and (7.19), that one shall combine with (7.26) and the representation

$$D_z F = \sum_{q=1} q I_{q-1}(f_q(\cdot, z)), \quad z \in Z$$

(where $F = \sum I_q(f_q)$ is the chaotic decomposition of F). To prove (7.28), we observe that, by a density argument, it is sufficient to consider a random variable having the form $F = I_q(f_q)$, where $q \geq 1$. In this case,

$$D_z F = q I_{q-1}(f_q(\cdot, z)),$$

and, since f_q is symmetric, $\delta DF = q I_q(f_q) = -LF$. This yields the desired conclusion. ■

We now present three crucial consequences of Theorem 7.1. The first one characterizes L as a second-order differential operator.

Proposition 7.7 *Let $F \in \mathcal{S}$ have the form $F = f(X(h_1), ..., X(h_d))$, with $f \in C_p^\infty(\mathbb{R}^d)$. Then, $F \in \mathrm{dom}(L)$, and moreover*

$$\begin{aligned} LF &= \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j} f(X(h_1), ..., X(h_d)) \langle h_i, h_j \rangle_{\mathfrak{H}} - \\ &\qquad \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(X(h_1), ..., X(h_d)) X(h_i). \end{aligned} \tag{7.29}$$

Proof. We know that $F \in \mathbb{D}^{1,2}$ and also

$$DF = \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(X(h_1), ..., X(h_d)) h_i.$$

By using Proposition 7.5, one sees immediately that $DF \in \mathrm{dom}(\delta)$, and moreover

$$\begin{aligned} \delta DF &= \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(X(h_1), ..., X(h_d)) X(h_i) \\ &\quad - \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j} f(X(h_1), ..., X(h_d)) \langle h_i, h_j \rangle_{\mathfrak{H}}. \end{aligned}$$

The conclusion follows from (7.28). ■

The next result will be fully exploited in Section 9: it is the starting point of the paper [57].

Theorem 7.2 (See [57]) *Let $F \in \mathbb{D}^{1,2}$ be such that $\mathbb{E}(F) = 0$, and consider a function $g : \mathbb{R} \to \mathbb{R}$. Assume that*

- *either g is continuously differentiable with a bounded first derivative*
- *or g is Lipschitz and the law of F is absolutely continuous.*

Then,

$$\mathbb{E}[Fg(F)] = \mathbb{E}\left[g'(F)\left\langle DF, -DL^{-1}F\right\rangle_{\mathfrak{H}}\right] = \mathbb{E}\left[g'(F)\,\mathbb{E}\left[\left\langle DF, -DL^{-1}F\right\rangle_{\mathfrak{H}} \mid F\right]\right] \tag{7.30}$$

If, $F = I_q(f)$, where $q \geq 1$ and $f \in \mathfrak{H}^{\odot q}$, then $L^{-1}F = -q^{-1}F$, and (7.30) becomes

$$\mathbb{E}[Fg(F)] = \frac{1}{q}\mathbb{E}\left[g'(F)\left\|DF\right\|_{\mathfrak{H}}^{2}\right] = \frac{1}{q}\mathbb{E}\left[g'(F)\,\mathbb{E}\left[\left\|DF\right\|_{\mathfrak{H}}^{2} \mid F\right]\right] \tag{7.31}$$

Proof. Observe that, thanks to (7.11), $g(F) \in \mathbb{D}^{1,2}$ and $Dg(F) = g'(F)\,DF$. Now write $F = LL^{-1}F = -\delta DL^{-1}F = \delta\left(-DL^{-1}F\right)$, so that by (7.15),

$$\begin{aligned}
\mathbb{E}[Fg(F)] &= \mathbb{E}\left[\delta\left(-DL^{-1}F\right)g(F)\right] = \mathbb{E}\left[\left\langle Dg(F), -DL^{-1}F\right\rangle_{\mathfrak{H}}\right] \\
&= \mathbb{E}\left[g'(F)\left\langle DF, -DL^{-1}F\right\rangle_{\mathfrak{H}}\right] = \mathbb{E}\left[g'(F)\,\mathbb{E}\left[\left\langle DF, -DL^{-1}F\right\rangle_{\mathfrak{H}} \mid F\right]\right]
\end{aligned}$$

The last part of the statement is straightforward. ■

Remarks. (1) The quantity $\left\langle DF, -DL^{-1}F\right\rangle_{\mathfrak{H}}$ will be crucial for the rest of the paper. Observe that this object can be directly represented in terms of the Ornstein-Uhlenbeck semigroup as follows

$$\left\langle DF, -DL^{-1}F\right\rangle_{\mathfrak{H}} = \int_0^\infty \left\langle DF, T_t DF\right\rangle_{\mathfrak{H}} e^{-t}dt, \tag{7.32}$$

or, with a more probabilistic twist,

$$\left\langle DF, -DL^{-1}F\right\rangle_{\mathfrak{H}} = \mathbb{E}\left[\left\langle DF, T_Y DF\right\rangle_{\mathfrak{H}} \mid X\right], \tag{7.33}$$

where Y is an exponential random variable with unitary parameter, independent of X.

(2) By using the second equality in (7.30), it is not difficult to prove that, for every $F \in \mathbb{D}^{1,2}$,

$$\mathbb{E}\left[\left\langle DF, -DL^{-1}F\right\rangle_{\mathfrak{H}} \mid F\right] \geq 0, \quad \text{a.s.-}\mathbb{P} \tag{7.34}$$

(see [57]).

(3) According to Goldstein and Reinert [28], for F as in the statement of Theorem 7.2, there exists a random variable F^* having the F-*zero biased* distribution, that is, F^* is such that, for every smooth function g,

$$E[g'(F^*)] = E[Fg(F)].$$

By using (7.30), one sees that

$$E[g'(F^*)] = E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\, g'(F)].$$

This implies that the conditional expectation $E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}|F]$ is a version of the Radon-Nikodym derivative of the law of F^* with respect to the law of F, whenever the two laws are equivalent.

To conclude, we present a characterization of the moments of a random variable in a fixed Wiener chaos.

Proposition 7.8 (See [56]) *Fix an integer $q \geq 2$ and set $F = I_q(f)$, with $f \in \mathfrak{H}^{\odot q}$. Then, for every integer $n \geq 0$, we have*

$$\mathbb{E}\left[F^n \|DF\|^2_{\mathfrak{H}}\right] = \frac{q}{n+1}\mathbb{E}\left[F^{n+2}\right]. \tag{7.35}$$

Proof. By using (7.12),

$$\begin{aligned}
\mathbb{E}\left[F^n \|DF\|^2_{\mathfrak{H}}\right] &= \mathbb{E}\left[F^n \langle DF, DF\rangle_{\mathfrak{H}}\right] = \frac{1}{n+1}\mathbb{E}\left[\langle DF, D(F^{n+1})\rangle_{\mathfrak{H}}\right] \\
&= \frac{1}{n+1}\mathbb{E}\left[\delta DF \times F^{n+1}\right] \\
&= \frac{q}{n+1}\mathbb{E}\left[F^{n+2}\right],
\end{aligned}$$

where we have used the fact that $\delta DF = -LF = qF$. ■

We will see that (7.35) can be a very effective alternative to the *combinatorial diagram formulae* that are customarily used in order to compute moments of chaotic random variables (see e.g. [72] or [92]).

8 Enter Stein's method

We are heading steadily towards the crux of these lectures, where we will show how to combine Malliavin's calculus with Stein's method, in order to assess the accuracy of the normal and non-normal approximation of functionals of isonormal Gaussian processes.

Before performing this task, it is necessary to recall some basic results involving Stein's method and distances between probability measures.

8.1 Distances between probability distributions

Let Y and Z be two random variables with values in $\mathbb{R}^d$. In what follows, we shall focus on *distances*, between the law of Y and the law of Z, having the following form

$$d_{\mathcal{G}}(Y,Z) = \sup\left\{\left|\mathbb{E}\left[g(Y)\right] - \mathbb{E}\left[g(Z)\right]\right| : g \in \mathcal{G}\right\}, \tag{8.1}$$

where $\mathcal{G}$ is some suitable class of functions. Our choices for $\mathcal{G}$ will always refer to one of the following examples.

- By taking $\mathcal{G} = \{g : \|g\|_L \leq 1\}$, where $\|\cdot\|_L$ is the usual Lipschitz seminorm given by

  $$\|g\|_L = \sup_{x \neq y} \frac{|g(x) - g(y)|}{\|x - y\|_{\mathbb{R}^d}},$$

  (with $\|\cdot\|_{\mathbb{R}^d}$ the usual Euclidian norm on $\mathbb{R}^d$) one obtains the *Wasserstein* (or *Kantorovich-Wasserstein*) distance.

- By taking $\mathcal{G}$ equal to the collection of all indicators $\mathbf{1}_B$ of Borel sets, one obtains the *total variation distance*.
- By taking $\mathcal{G}$ equal to the class of all indicators functions $\mathbf{1}_{(-\infty,z_1]} \cdots \mathbf{1}_{(-\infty,z_d]}$, $(z_1, ..., z_d) \in \mathbb{R}^d$, one has the *Kolmogorov* distance.

In what follows, we shall sometimes denote by $d_W(.,.)$, $d_{TV}(.,.)$ and $d_{Kol}(.,.)$, respectively, the Wasserstein, total variation and Kolmogorov distances. Observe that $d_{TV}(.,.) \geq d_{Kol}(.,.)$. Moreover, the topologies induced by d_W, d_{TV} and d_{Kol} are stronger than the topology of convergence in distribution (see e.g. [22, Ch. 11] for an account of results involving distances on spaces of probability measures).

8.2 Stein's method in dimension one

We shall now give a short account of Stein's method, which is basically a set of techniques allowing to evaluate distances of the type (8.1) by means of differential operators. As already recalled in the Introduction, this theory has been initiated by Stein in the path-breaking paper [88], and then further developed in the monograph [89]. For a comprehensive introduction, see the two surveys [14] and [80]. In this section, we will apply Stein's method to two types of one dimensional approximations, namely Gaussian and (centered) Gamma. As before, we shall denote by $\mathcal{N}(0,1)$ a standard Gaussian random variable. The centered Gamma random variables we are interested in have the form

$$F(\nu) \stackrel{\text{Law}}{=} 2G(\nu/2) - \nu, \quad \nu > 0, \tag{8.2}$$

where $G(\nu/2)$ has a Gamma law with parameter $\nu/2$. This means that $G(\nu/2)$ is a (a.s. strictly positive) random variable with density

$$\phi(x) = \frac{x^{\frac{\nu}{2}-1} \mathrm{e}^{-x}}{\Gamma(\nu/2)} \mathbf{1}_{(0,\infty)}(x),$$

where Γ is the usual Gamma function. Observe in particular that, if $\nu \geq 1$ is an integer, then $F(\nu)$ has a centered χ^2 distribution with ν degrees of freedom.

Standard Gaussian distribution. Let $N \sim \mathcal{N}(0,1)$. Consider a real-valued function $g : \mathbb{R} \to \mathbb{R}$ such that the expectation $E(g(N))$ is well-defined. The *Stein equation* associated with g and N is given by

$$g(x) - E(g(N)) = f'(x) - xf(x), \quad x \in \mathbb{R}. \tag{8.3}$$

A solution to (8.3) is a function f which is Lebesgue a.e.-differentiable, and such that there exists a version of f' verifying (8.3) for every $x \in \mathbb{R}$. The following result is basically due to Stein [88, 89]. The proof of point (i) (whose content is usually referred as *Stein's lemma*) involves a standard use of the Fubini theorem (see e.g. [14, Lemma 2.1]). Point (ii) is proved e.g. in [14, Lemma 2.2]; point (iii) can be obtained by combining e.g. the arguments in [89, p. 25] and [9, Lemma 5.1]; point (iv) is proved e.g. in [8, Lemma 4.3].

Lemma 8.1 (i) *Let W be a random variable. Then, $W \stackrel{\text{Law}}{=} N \sim \mathcal{N}(0,1)$ if, and only if,*

$$E[f'(W) - Wf(W)] = 0, \tag{8.4}$$

for every continuous and piecewise continuously differentiable function f verifying the relation $E|f'(Z)| < \infty$.

(ii) *If $g(x) = \mathbf{1}_{(-\infty,z]}(x)$, $z \in \mathbb{R}$, then (8.3) admits a solution f which is bounded by $\sqrt{2\pi}/4$, piecewise continuously differentiable and such that $\|f'\|_\infty \leq 1$.*

(iii) *If g is bounded by 1/2, then (8.3) admits a solution f which is bounded by $\sqrt{\pi/2}$, Lebesgue a.e. differentiable and such that $\|f'\|_\infty \leq 2$.*

(iv) *If g is absolutely continuous with bounded derivative, then (8.3) has a solution f which is twice differentiable and such that $\|f'\|_\infty \leq \|g'\|_\infty$ and $\|f''\|_\infty \leq 2\|g'\|_\infty$.*

We also recall the relation:

$$2d_{TV}(X,Y) = \sup\{|E(u(X)) - E(u(Y))| : \|u\|_\infty \leq 1\}. \tag{8.5}$$

Note that point (ii) and (iii) (via (8.5)) imply the following bounds on the Kolmogorov and total variation distance between Z and an arbitrary random variable Y:

$$d_{Kol}(Y,Z) \leq \sup_{f \in \mathcal{F}_{Kol}} |E(f'(Y) - Yf(Y))| \tag{8.6}$$

$$d_{TV}(Y,Z) \leq \sup_{f \in \mathcal{F}_{TV}} |E(f'(Y) - Yf(Y))| \tag{8.7}$$

where $\mathcal{F}_{Kol}$ and $\mathcal{F}_{TV}$ are, respectively, the class of piecewise continuously differentiable functions that are bounded by $\sqrt{2\pi}/4$ and such that their derivative is bounded by 1, and the class of piecewise continuously differentiable functions that are bounded by $\sqrt{\pi/2}$ and such that their derivative is bounded by 2.

Analogously, by using (v) along with the relation $\|h\|_L = \|h'\|_\infty$, one obtains

$$d_W(Y,Z) \leq \sup_{f \in \mathcal{F}_W} |E(f'(Y) - Yf(Y))|, \tag{8.8}$$

where $\mathcal{F}_W$ is the class of twice differentiable functions, whose first derivative is bounded by 1 and whose second derivative is bounded by 2.

Centered Gamma distribution. Let $F(\nu)$ be as in (8.2). Consider a real-valued function $g : \mathbb{R} \to \mathbb{R}$ such that the expectation $E[g(F(\nu))]$ exists. The *Stein equation* associated with g and $F(\nu)$ is:

$$g(x) - E[g(F(\nu))] = 2(x+\nu)f'(x) - xf(x), \quad x \in (-\nu, +\infty). \tag{8.9}$$

The following statement collects some slight variations around results proved by Stein [89], Diaconis and Zabell [20], Luk [40], Schoutens [85] and Pickett [94]. See also [80]. It is the "Gamma counterpart" of Lemma 8.1.

Lemma 8.2 (**i**) *Let W be a real-valued random variable whose law admits a density with respect to the Lebesgue measure. Then, $W \overset{\text{Law}}{=} F(\nu)$ if and only if*

$$E[2(W+\nu)_+ f'(W) - Wf(W)] = 0, \tag{8.10}$$

where $a_+ := \max(a,0)$, for every smooth function f such that the mapping $x \mapsto 2(x+\nu)_+ f'(x) - xf(x)$ is bounded.

(**ii**) *If $|g(x)| \le c\exp(ax)$ for every $x > -\nu$ and for some $c > 0$ and $a < 1/2$, and if g is twice differentiable, then (8.9) has a solution f which is bounded on $(-\nu, +\infty)$, differentiable and such that $\|f\|_\infty \le 2\|g'\|_\infty$ and $\|f'\|_\infty \le \|g''\|_\infty$.*

(**iii**) *Suppose that $\nu \ge 1$ is an integer. If $|g(x)| \le c\exp(ax)$ for every $x > -\nu$ and for some $c > 0$ and $a < 1/2$, and if g is twice differentiable with bounded derivatives, then (8.9) has a solution f which is bounded on $(-\nu, +\infty)$, differentiable and such that $\|f\|_\infty \le \sqrt{2\pi/\nu}\|g\|_\infty$ and $\|f'\|_\infty \le \sqrt{2\pi/\nu}\|g'\|_\infty$.*

Now define

$$\begin{aligned} \mathcal{G}_1 &= \{g \in C_b^2 : \|g\|_\infty \le 1, \|g'\|_\infty \le 1, \|g''\|_\infty \le 1\}, & (8.11)\\ \mathcal{G}_2 &= \{g \in C_b^2 : \|g\|_\infty \le 1, \|g'\|_\infty \le 1\}, & (8.12)\\ \mathcal{G}_{1,\nu} &= \mathcal{G}_1 \cap C_b^2(\nu) & (8.13)\\ \mathcal{G}_{2,\nu} &= \mathcal{G}_2 \cap C_b^2(\nu) & (8.14) \end{aligned}$$

where C_b^2 denotes the class of twice differentiable functions with support in $\mathbb{R}$ and with bounded derivatives, and $C_b^2(\nu)$ denotes the subset of C_b^2 composed of functions with support in $(-\nu, +\infty)$. Note that point (ii) in the previous statement implies that, by adopting the notation (8.1) and for every $\nu > 0$ and every real random variable Y (not necessarily with support in $(-\nu, +\infty)$),

$$d_{\mathcal{G}_{1,\nu}}(Y, F(\nu)) \le \sup_{f \in \mathcal{F}_{1,\nu}} |E[2(Y+\nu)f'(Y) - Yf(Y)]| \tag{8.15}$$

where $\mathcal{G}_{1,\nu}$ is the class of differentiable functions with support in $(-\nu, +\infty)$, bounded by 2 and whose derivatives are bounded by 1. Analogously, point (iii) implies that, for every integer $\nu \ge 1$,

$$d_{\mathcal{G}_{2,\nu}}(Y, F(\nu)) \le \sup_{f \in \mathcal{F}_{2,\nu}} |E[2(Y+\nu)f'(Y) - Yf(Y)]|, \tag{8.16}$$

where $\mathcal{F}_{2,\nu}$ is the class of differentiable functions with support in $(-\nu, +\infty)$, bounded by $\sqrt{2\pi/\nu}$ and whose derivatives are also bounded by $\sqrt{2\pi/\nu}$. A little inspection shows that the following estimates also hold: for every $\nu > 0$ and every random variable Y,

$$d_{\mathcal{G}_1}(Y, F(\nu)) \le \sup_{f \in \mathcal{F}_1} |E[2(Y+\nu)_+ f'(Y) - Yf(Y)]|$$

where $\mathcal{F}_1$ is the class of functions (defined on $\mathbb{R}$) that are continuous and differentiable on $\mathbb{R}\backslash\{\nu\}$, bounded by $\max\{2, 2/\nu\}$, and whose derivatives are bounded by $\max\{1, 1/\nu + 2/\nu^2\}$. Analogously, for every integer $\nu \ge 1$,

$$d_{\mathcal{G}_2}(Y, F(\nu)) \le \sup_{f \in \mathcal{F}_2} |E[2(Y+\nu)_+ f'(Y) - Yf(Y)]|, \tag{8.17}$$

where $\mathcal{F}_2$ is the class of functions (on $\mathbb{R}$) that are continuous and differentiable on $\mathbb{R}\backslash\{\nu\}$, bounded by $\max\{\sqrt{2\pi/\nu}, 2/\nu\}$, and whose derivatives are bounded by $\max\{\sqrt{2\pi/\nu}, 1/\nu + 2/\nu^2\}$.

8.3 Multi-dimensional Stein's Lemma: a Malliavin calculus approach

We start by introducing some useful norms over classes of real-valued matrices.

Definition 8.1 (**i**) *The **Hilbert-Schmidt inner product** and the **Hilbert-Schmidt norm** on the class of $d \times d$ real matrices, denoted respectively by $\langle \cdot, \cdot \rangle_{H.S.}$ and $\| \cdot \|_{H.S.}$, are defined as follows: for every pair of matrices A and B, $\langle A, B \rangle_{H.S.} \triangleq \mathrm{Tr}(AB^T)$ and $\|A\|_{H.S.} \triangleq \sqrt{\langle A, A \rangle_{H.S.}}$.*

(**ii**) *The **operator norm** of a $d \times d$ matrix A over $\mathbb{R}$ is given by $\|A\|_{op} \triangleq \sup_{\|x\|_{\mathbb{R}^d}=1} \|Ax\|_{\mathbb{R}^d}$.*

Remark. According to the just introduced notation, we can rewrite the differential characterization of the generator L, as given in (7.29), as follows: for every smooth

$$F = f\left(X\left(h_1\right), ..., X\left(h_d\right)\right),$$

one has that

$$LF = \langle C, \mathrm{Hess} f\left(Z\right) \rangle_{H.S.} - \langle Z, \nabla f\left(Z\right) \rangle_{\mathbb{R}^d}, \tag{8.18}$$

where $\mathrm{Hess} f$ is the Hessian matrix of f, $Z = \left(X\left(h_1\right), ..., X\left(h_d\right)\right)$, and $C = \{C\left(i,j\right) : 1 \leq i,j \leq d\}$ is the covariance matrix given by $C\left(i,j\right) = \langle h_i, h_j \rangle_{\mathfrak{H}}$.

Given a $d \times d$ positive definite symmetric matrix C, we use the notation $\mathcal{N}_d(0, C)$ to indicate the law of a d-dimensional Gaussian vector with zero mean and covariance C. The following result is the d-dimensional counterpart of Stein's Lemma 8.1. Here, we provide a proof (which is taken from [60]) that is almost completely based on Malliavin calculus.

Lemma 8.3 *Fix an integer $d \geq 2$ and let $C = \{C(i,j) : i,j = 1, ..., d\}$ be a $d \times d$ positive definite symmetric real matrix.*

(**i**) *Let Y be a random variable with values in $\mathbb{R}^d$. Then $Y \sim \mathcal{N}_d(0, C)$ if and only if, for every twice differentiable function $f : \mathbb{R}^d \to \mathbb{R}$ such that $\mathbb{E}|\langle C, \mathrm{Hess} f(Y) \rangle_{H.S.} - \langle Y, \nabla f(Y) \rangle_{\mathbb{R}^d}| < \infty$, it holds that*

$$\mathbb{E}[\langle Y, \nabla f(Y) \rangle_{\mathbb{R}^d} - \langle C, \mathrm{Hess} f(Y) \rangle_{H.S.}] = 0. \tag{8.19}$$

(**ii**) *Consider a Gaussian random vector $Z \sim \mathcal{N}_d(0, C)$. Let $g : \mathbb{R}^d \to \mathbb{R}$ belong to $C^2(\mathbb{R}^d)$ with first and second bounded derivatives. Then, the function $U_0(g)$ defined by*

$$U_0 g(x) := \int_0^1 \frac{1}{2t} \mathbb{E}[g(\sqrt{t}x + \sqrt{1-t}Z) - g(Z)] dt$$

is a solution to the following differential equation (with unknown function f):

$$g(x) - \mathbb{E}[g(Z)] = \langle x, \nabla f(x) \rangle_{\mathbb{R}^d} - \langle C, \mathrm{Hess} f(x) \rangle_{H.S.}, \quad x \in \mathbb{R}^d. \tag{8.20}$$

Moreover, one has that

$$\sup_{x \in \mathbb{R}^d} \|\mathrm{Hess}\, U_0 g(x)\|_{H.S.} \leq \|C^{-1}\|_{op} \, \|C\|_{op}^{1/2} \, \|g\|_L. \tag{8.21}$$

Remark. If $C = \sigma^2 \mathbf{I}_d$ for some $\sigma > 0$ (that is, if Z is composed of i.i.d. centered Gaussian random variables with common variance equal to σ^2), then

$$\|C^{-1}\|_{op}\, \|C\|_{op}^{1/2} = \|\sigma^{-2}\mathbf{I}_d\|_{op}\, \|\sigma^2 \mathbf{I}_d\|_{op}^{1/2} = \sigma^{-1}.$$

Proof of Lemma 8.3. We start by proving Point (ii). First observe that, without loss of generality, we can suppose that $Z = (Z_1, ..., Z_d) \triangleq (X(h_1), ...X(h_d))$, where X is an isonormal Gaussian process over some Hilbert space $\mathfrak{H}$, the kernels h_i belong to $\mathfrak{H}$ $(i = 1, ..., d)$, and $\langle h_i, h_j\rangle_{\mathfrak{H}} = \mathbb{E}(X(h_i)X(h_j)) = \mathbb{E}(Z_i Z_j) = C(i,j)$. By using the change of variable $2u = -\log t$, one can rewrite $U_0 g(x)$ as follows

$$U_0 g(x) = \int_0^\infty \{\mathbb{E}[g(e^{-u}x + \sqrt{1-e^{-2u}}Z)] - \mathbb{E}[g(Z)]\}du.$$

Now define $\widetilde{g}(Z) := g(Z) - \mathbb{E}[g(Z)]$, and observe that $\widetilde{g}(Z)$ is by assumption a centered element of $L^2(\sigma(X))$. For $q \geq 0$, denote by $J_q(\widetilde{g}(Z))$ the projection of $\widetilde{g}(Z)$ on the qth Wiener chaos, so that $J_0(\widetilde{g}(Z)) = 0$. According to Mehler's formula (7.24),

$$\mathbb{E}[g(e^{-u}x + \sqrt{1-e^{-2u}}Z)]|_{x=Z} - \mathbb{E}[g(Z)] = \mathbb{E}[\widetilde{g}(e^{-u}x + \sqrt{1-e^{-2u}}Z)]|_{x=Z} = T_u \widetilde{g}(Z),$$

where x denotes a generic element of $\mathbb{R}^d$. In view of (7.21), it follows that

$$U_0 g(Z) = \int_0^\infty T_u \widetilde{g}(Z) du = \int_0^\infty \sum_{q\geq 1} e^{-qu} J_q(\widetilde{g}(Z)) du = \sum_{q\geq 1} \frac{1}{q} J_q(\widetilde{g}(Z)) = -L^{-1}\widetilde{g}(Z).$$

Since g belongs to $C^2(\mathbb{R}^d)$ with bounded first and second derivatives, it is easily seen that the same holds for $U_0 g$. By exploiting the differential representation (8.18), one deduces that

$$\langle Z, \nabla U_0 g(Z)\rangle_{\mathbb{R}^d} - \langle C, \mathrm{Hess} U_0 g(Z)\rangle_{H.S.} = -LU_0 g(Z) = LL^{-1}\widetilde{g}(Z) = g(Z) - \mathbb{E}[g(Z)].$$

Since the matrix C is positive definite, we infer that the support of the law of Z coincides with $\mathbb{R}^d$, and therefore (e.g. by a continuity argument) we obtain that

$$\langle x, \nabla U_0 g(x)\rangle_{\mathbb{R}^d} - \langle C, \mathrm{Hess}\, U_0 g(x)\rangle_{H.S.} = g(x) - E[g(Z)],$$

for every $x \in \mathbb{R}^d$. This yields that the function $U_0 g$ solves the Stein's equation (8.20).

To prove the estimate (8.21), we first recall that there exists a unique non-singular *symmetric* matrix A such that $A^2 = C$, and that one has that $A^{-1}Z \sim \mathcal{N}_d(0, \mathbf{I}_d)$. Now write $U_0 g(x) = h(A^{-1}x)$, where

$$h(x) = \int_0^1 \frac{1}{2t} \mathbb{E}[g_A(\sqrt{t}x + \sqrt{1-t}A^{-1}Z) - g_A(A^{-1}Z)]dt,$$

and $g_A(x) = g(Ax)$. Note that, since $A^{-1}Z \sim \mathcal{N}_d(0, \mathbf{I}_d)$, the function h solves the Stein's equation $\langle x, \nabla h(x)\rangle_{\mathbb{R}^d} - \Delta h(x) = g_A(x) - \mathbb{E}[g_A(Y)]$, where $Y \sim \mathcal{N}_d(0, \mathbf{I}_d)$. We can now use standard arguments (see e.g. the proof of Lemma 3 in [10]) in order to deduce that

$$\sup_{x\in\mathbb{R}^d} \|\mathrm{Hess}\, h(x)\|_{H.S.} \leq \|g_A\|_{Lip} \leq \|A\|_{op}\|g\|_L. \tag{8.22}$$

On the other hand, by noting $h_{A^{-1}}(x) = h(A^{-1}x)$, one obtains by standard computations (recall that A is symmetric) that $\mathrm{Hess}\, U_0 g(x) = \mathrm{Hess}\, h_{A^{-1}}(x) = A^{-1}\mathrm{Hess}\, h(A^{-1}x)A^{-1}$, yielding

$$\begin{aligned}
\sup_{x\in\mathbb{R}^d} \|\mathrm{Hess}\, U_0 g(x)\|_{H.S.} &= \sup_{x\in\mathbb{R}^d} \|A^{-1}\mathrm{Hess}\, h(A^{-1}x)A^{-1}\|_{H.S.} \\
&= \sup_{x\in\mathbb{R}^d} \|A^{-1}\mathrm{Hess}\, h(x)A^{-1}\|_{H.S.} \\
&\leq \|A^{-1}\|_{op}^2 \sup_{x\in\mathbb{R}^d} \|\mathrm{Hess}\, h(x)\|_{H.S.} &\quad (8.23)\\
&\leq \|A^{-1}\|_{op}^2\, \|A\|_{op}\, \|g\|_L &\quad (8.24)\\
&\leq \|C^{-1}\|_{op}\, \|C\|_{op}^{1/2}\, \|g\|_L. &\quad (8.25)
\end{aligned}$$

The chain of inequalities appearing in formulae (8.23)–(8.25) are mainly a consequence of the usual properties of the Hilbert-Schmidt and operator norms. Indeed, to prove inequality (8.23) we used the relations

$$\begin{aligned}
\|A^{-1}\mathrm{Hess}\, h(x)A^{-1}\|_{H.S.} &\leq \|A^{-1}\|_{op}\, \|\mathrm{Hess}\, h(x)A^{-1}\|_{H.S.} \\
&\leq \|A^{-1}\|_{op}\, \|\mathrm{Hess}\, h(x)\|_{H.S.}\, \|A^{-1}\|_{op}\ ;
\end{aligned}$$

relation (8.24) is a consequence of (8.22); finally, to show the inequality (8.25), one uses the fact that

$$\|A^{-1}\|_{op} \leq \sqrt{\|A^{-1}A^{-1}\|_{op}} = \sqrt{\|C^{-1}\|_{op}} \quad \text{and} \quad \|A\|_{op} \leq \sqrt{\|AA\|_{op}} = \sqrt{\|C\|_{op}}\ .$$

We are now left with the proof of Point (i) in the statement. The fact that a vector $Y \sim \mathcal{N}_d(0, C)$ necessarily verifies (8.19) can be proved by standard integration by parts. On the other hand, suppose that Y verifies (8.19). Then, according to Point (ii), for every $g \in C^2(\mathbb{R}^d)$ with bounded first and second derivatives,

$$\mathbb{E}(g(Y)) - \mathbb{E}(g(Z)) = \mathbb{R}(\langle Y, \nabla U_0 g(Y)\rangle_{\mathbb{R}^d} - \langle C, \mathrm{Hess}\, U_0 g(Y)\rangle_{H.S.}) = 0,$$

where $Z \sim \mathcal{N}_d(0, C)$. Since the collection of all such functions g generates the Borel σ-field on $\mathbb{R}^d$, this implies that $Y \overset{\text{Law}}{=} Z$, thus yielding the desired conclusion. ■

9 Explicit bounds using Malliavin operators

9.1 One-dimensional normal approximation

Consider a standard Gaussian random variable $N \sim \mathcal{N}(0, 1)$, as well as a functional F of some isonormal Gaussian process $X = \{X(h) : h \in \mathfrak{H}\}$. We are interested in assessing the distance between the law of N and the law of F by using relations (8.6)–(8.8). As shown in the next statement, which has been first proved in [57], this task is particularly easy if one assumes that F is also Malliavin differentiable.

Theorem 9.1 (See [57]) *Let $F \in \mathbb{D}^{1,2}$ be such that $\mathbb{E}[F] = 0$. Then, one has that*

$$\begin{aligned}
d_W(F, N) &\leq \mathbb{E}\left|1 - \left\langle DF, -DL^{-1}F\right\rangle_{\mathfrak{H}}\right| &\quad (9.1)\\
&\leq \mathbb{E}[(1 - \left\langle DF, -DL^{-1}F\right\rangle_{\mathfrak{H}})^2]^{1/2}
\end{aligned}$$

Moreover, if the law of F is absolutely continuous, then

$$d_{Kol}(F,N) \leq \mathbb{E}\left|1-\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}}\right| \leq \mathbb{E}[(1-\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}})^2]^{1/2} \tag{9.2}$$

$$d_{TV}(F,N) \leq 2\mathbb{E}\left|1-\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}}\right| \leq 2\mathbb{E}[(1-\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}})^2]^{1/2} \tag{9.3}$$

Proof. Let d be one of the four distances d_W, d_{Kol} and d_{TV}, and let $\mathcal{F}$ be the associated functional class $\mathcal{F}_W$, $\mathcal{F}_{Kol}$ or $\mathcal{F}_{FM}$, as defined on page 39. By combining (8.6)–(8.8) with Theorem 7.2, one deduces that

$$\begin{aligned}
d(F,N) &\leq \sup_{f\in\mathcal{F}}\left|\mathbb{E}\left[f'(F)-f(F)F\right]\right| \\
&= \sup_{f\in\mathcal{F}}\left|\mathbb{E}\left[f'(F)-f'(F)\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}}\right]\right| \\
&\leq [\sup_{f\in\mathcal{F}}\left\|f'\right\|_\infty]\times\mathbb{E}\left|1-\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}}\right| \\
&\leq [\sup_{f\in\mathcal{F}}\left\|f'\right\|_\infty]\times\mathbb{E}[(1-\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}})^2]^{1/2},
\end{aligned}$$

where in the last inequality we used Cauchy-Schwarz. Note that, in order to apply Theorem 7.2 when $f\in\mathcal{F}_{Kol}$ or $f\in\mathcal{F}_{FM}$, one needs to assume that F has an absolutely continuous distribution (indeed, in this case f is merely Lipschitz). ■

Remark. Observe that

$$\mathbb{E}[\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}}]=\mathbb{E}\left[F^2\right],$$

and therefore

$$\mathbb{E}[(1-\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}})^2]^{1/2}\leq\left|1-\mathbb{E}\left[F^2\right]\right|+\mathbf{Var}\left(\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}}\right)^{1/2}. \tag{9.4}$$

Note also that $\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}}\in L^1(\mathbb{P})$, but $\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}}$ is not necessarily square-integrable. To have that $\left\langle DF,-DL^{-1}F\right\rangle\in L^2(\mathbb{P})$ one needs further regularity assumptions on F : for instance, if F lives in a finite sum of Wiener chaoses, then $\left\langle DF,-DL^{-1}F\right\rangle_{\mathfrak{H}}\in L^p(\mathbb{P})$ for every $p\geq 1$.

Of course, the relevance of the bounds (9.1)–(9.3) can only be appreciated through examples. In the forthcoming Section 10, we will prove that these bounds lead indeed to several striking generalizations of some *central limit theorems on Wiener chaos* proved in [66], [67] and [73]. We now provide a first example, taken from [57], where it is shown that (9.3) contains as a special case a technical result proved by Chatterjee in [9].

Example. In [9, Lemma 5.3], Chatterjee has proved the following result. Let $Y=g(V)$, where $V=(V_1,...,V_n)$ is a vector of centered i.i.d. standard Gaussian random variables, and $g:\mathbb{R}^n\to\mathbb{R}$ is a smooth function such that: (i) g and its derivatives have subexponential growth at infinity, (ii) $E(g(V))=0$, and (iii) $E(g(V)^2)=1$. Then, for any Lipschitz function f, one has that

$$\mathbb{E}[Yf(Y)]=\mathbb{E}[S(V)f'(Y)], \tag{9.5}$$

where, for every $v = (v_1, ..., v_n) \in \mathbb{R}^n$,

$$S(v) = \int_0^1 \frac{1}{2\sqrt{t}} \mathbb{E}\left[\sum_{i=1}^n \frac{\partial g}{\partial v_i}(v) \frac{\partial g}{\partial v_i}(\sqrt{t}v + \sqrt{1-t}V)\right] dt, \tag{9.6}$$

so that, for instance, for $N \sim \mathcal{N}(0,1)$ and by using (8.7), Lemma 8.1 (iii), (8.5) and Cauchy-Schwarz inequality,

$$d_{TV}(Y, Z) \leq 2\mathbb{E}[(S(V)-1)^2]^{1/2}. \tag{9.7}$$

We shall prove that (9.5) is a very special case of the first equality in (7.30), and therefore that (9.7) is a special case of (9.3). Observe first that, without loss of generality, we can assume that $V_i = X(h_i)$, where X is an isonormal process over some Hilbert space of the type $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$ and $\{h_1, ..., h_n\}$ is an orthonormal system in $\mathfrak{H}$. Since $Y = g(V_1, \dots, V_n)$, according to (7.3) we have that $D_a Y = \sum_{i=1}^n \frac{\partial g}{\partial x_i}(V) h_i(a)$. On the other hand, since Y is centered and square integrable, it admits a chaotic representation of the form $Y = \sum_{q\geq 1} I_q(\psi_q)$. This implies in particular that $D_a Y = \sum_{q=1}^\infty q I_{q-1}(\psi_q(a,\cdot))$. Moreover, one has that $-L^{-1}Y = \sum_{q\geq 1} \frac{1}{q} I_q(\psi_q)$, so that $-D_a L^{-1} Y = \sum_{q\geq 1} I_{q-1}(\psi_q(a,\cdot))$. Now, let T_z, $z \geq 0$, denote the Ornstein-Uhlenbeck semigroup introduced in (7.21), whose action on random variables $F \in L^2(\sigma(X))$ is given by $T_z(F) = \sum_{q\geq 0} e^{-qz} J_q(F)$, where $J_q(F)$ denotes the projection of F on the qth Wiener chaos. We can write

$$\begin{aligned} \int_0^1 \frac{1}{2\sqrt{t}} T_{\ln(1/\sqrt{t})}(D_a Y) dt &= \int_0^\infty e^{-z} T_z(D_a Y) dz = \sum_{q\geq 1} \frac{1}{q} J_{q-1}(D_a Y) \\ &= \sum_{q\geq 1} I_{q-1}(\psi_q(a,\cdot)) = -D_a L^{-1} Y. \end{aligned} \tag{9.8}$$

Now recall that Mehler's formula (7.24) implies that

$$T_z(f(V)) = \mathbb{E}[f(e^{-z}v + \sqrt{1-e^{-2z}}V)]\Big|_{v=V}, \quad z \geq 0.$$

In particular, by applying this last relation to the partial derivatives $\frac{\partial g}{\partial v_i}$, $i = 1, ..., n$, we deduce from (9.8) that

$$\int_0^1 \frac{1}{2\sqrt{t}} T_{\ln(1/\sqrt{t})}(D_a Y) dt = \sum_{i=1}^n h_i(a) \int_0^1 \frac{1}{2\sqrt{t}} \mathbb{E}\left[\frac{\partial g}{\partial v_i}(\sqrt{t}\, v + \sqrt{1-t}\, V)\right] dt \Bigg|_{v=V}.$$

Consequently, (9.5) follows, since

$$\begin{aligned} \langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}} &= \left\langle \sum_{i=1}^n \frac{\partial g}{\partial v_i}(V) h_i, \sum_{i=1}^n \int_0^1 \frac{1}{2\sqrt{t}}\ \mathbb{E}\left[\frac{\partial g}{\partial v_i}(\sqrt{t}\, v + \sqrt{1-t}\, V)\right] dt \Bigg|_{v=V} h_i \right\rangle_{\mathfrak{H}} \\ &= S(V). \end{aligned}$$

Remark. By inspection of the proof of Theorem 9.1, one sees that it is possible to refine the bounds (9.1)–(9.3) and (9.4), by replacing $\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}$ with $\mathbb{E}\left[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}} \mid F\right]$ and $\mathbf{Var}\left(\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\right)$ with $\mathbf{Var}\left(\mathbb{E}\left[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}} \mid F\right]\right)$.

9.2 Multi-dimensional normal approximation

We now present a multidimensional version of Theorem 9.1 which is based on the multidimensional Stein Lemma 8.1. See [60] and [62] for more results in this direction.

Theorem 9.2 (See [60]) *Fix $d \geq 2$ and let $C = \{C(i,j) : i,j = 1,...,d\}$ be a $d \times d$ positive definite matrix. Suppose that $Z \sim \mathcal{N}_d(0, C)$ and that $F = (F_1, \ldots, F_d)$ is a $\mathbb{R}^d$-valued random vector such that $\mathbb{E}[F_i] = 0$ and $F_i \in \mathbb{D}^{1,2}$ for every $i = 1, \ldots, d$. Then,*

$$\begin{aligned} d_W(F,Z) &\leq \|C^{-1}\|_{op}\ \|C\|_{op}^{1/2}\sqrt{\mathbb{E}\|C-\Phi(DF)\|_{H.S}^2} &(9.9)\\ &= \|C^{-1}\|_{op}\ \|C\|_{op}^{1/2}\sqrt{\sum_{i,j=1}^{d}\mathbb{E}[(C(i,j)-\langle DF_i, -DL^{-1}F_j\rangle_{\mathfrak{H}})^2]}, &(9.10)\end{aligned}$$

where we write $\Phi(DF)$ to indicate the matrix $\Phi(DF) = \{\langle DF_i, -DL^{-1}F_j\rangle_{\mathfrak{H}} : 1 \leq i,j \leq d\}$.

Proof. We start by proving that, for every $g \in C^2(\mathbb{R}^d)$ with bounded first and second derivatives,

$$|\mathbb{E}[g(F)] - \mathbb{E}[g(Z)]| \ \leq \|C^{-1}\|_{op}\ \|C\|_{op}^{1/2}\ \|g\|_L\sqrt{\mathbb{E}\|C-\Phi(DF)\|_{H.S}^2}.$$

To prove such a claim, observe that, according to Point (ii) in Lemma 8.3, $\mathbb{E}[g(F)] - \mathbb{E}[g(Z)] = \mathbb{E}[\langle F, \nabla U_0 g(F)\rangle_{\mathbb{R}^d} - \langle C, \mathrm{Hess} U_0 g(F)\rangle_{H.S.}]$. Now let us write $\partial_{ij}^2 = \frac{\partial^2}{\partial x_i \partial x_j}$; we have that

$$\begin{aligned} & \left|\mathbb{E}[\langle C, \mathrm{Hess}U_0 g(F)\rangle_{H.S.} - \langle F, \nabla U_0 g(F)\rangle_{\mathbb{R}^d}]\right| \\ = & \left|\mathbb{E}\left[\sum_{i,j=1}^{d} C(i,j)\partial_{ij}^2 U_0 g(F) - \sum_{i=1}^{d} F_i \partial_i U_0 g(F)\right]\right| \\ = & \left|\sum_{i,j=1}^{d}\mathbb{E}\left[C(i,j)\partial_{ij}^2 U_0 g(F)\right] - \sum_{i=1}^{d}\mathbb{E}\left[\left(LL^{-1}F_i\right)\partial_i U_0 g(F)\right]\right| \ \text{(since } E(F_i)=0) \\ = & \left|\sum_{i,j=1}^{d}\mathbb{E}\left[C(i,j)\partial_{ij}^2 U_0 g(F)\right] + \sum_{i=1}^{d}\mathbb{E}\left[\delta(DL^{-1}F_i)\partial_i U_0 g(F)\right]\right| \ \text{(since } \delta D = -L) \\ = & \left|\sum_{i,j=1}^{d}\mathbb{E}\left[C(i,j)\partial_{ij}^2 U_0 g(F)\right] - \sum_{i=1}^{d}\mathbb{E}\left[\langle D(\partial_i U_0 g(F)), -DL^{-1}F_i\rangle_{\mathfrak{H}}\right]\right| \ \text{(by (7.15))} \\ = & \left|\sum_{i,j=1}^{d}\mathbb{E}\left[C(i,j)\partial_{ij}^2 U_0 g(F)\right] - \sum_{i,j=1}^{d}\mathbb{E}\left[\partial_{ji}^2 U_0 g(F)\langle DF_j, -DL^{-1}F_i\rangle_{\mathfrak{H}}\right]\right| \ \text{(by (7.11))} \\ = & \left|\sum_{i,j=1}^{d}\mathbb{E}\left[\partial_{ij}^2 U_0 g(F)\left(C(i,j) - \langle DF_i, -DL^{-1}F_j\rangle_{\mathfrak{H}}\right)\right]\right| \\ = & \left|\mathbb{E}\langle \mathrm{Hess}\, U_0 g(F), C - \Phi(DF))\rangle_{H.S.}\right| \\ \leq & \sqrt{\mathbb{E}\|\mathrm{Hess}\, U_0 g(F)\|_{H.S}^2}\sqrt{\mathbb{E}\|C-\Phi(DF)\|_{H.S}^2} \ \text{(by the Cauchy-Schwarz inequality)} \\ \leq & \|C^{-1}\|_{op}\ \|C\|_{op}^{1/2}\ \|g\|_L\sqrt{\mathbb{E}\|C-\Phi(DF)\|_{H.S}^2} \ \text{(by (8.21)).} \end{aligned}$$

To prove the Wasserstein estimate (9.9), it is sufficient to observe that, for every globally Lipschitz function g such that $\|g\|_L \leq 1$, there exists a family $\{g_\varepsilon : \varepsilon > 0\}$ such that:

(i) for each $\varepsilon > 0$, the first and second derivatives of g_ε are bounded;

(ii) for each $\varepsilon > 0$, one has that $\|g_\varepsilon\|_{Lip} \leq \|g\|_L$;

(iii) as $\varepsilon \to 0$, $\|g_\varepsilon - g\|_\infty \downarrow 0$.

For instance, we can choose $g_\varepsilon(x) = E\big[g(x + \sqrt{\varepsilon}S)\big]$ with $S \sim \mathcal{N}_d(0, \mathbf{I}_d)$. ■

Theorem 9.2 will be fully exploited in Section 10.2, where we will obtain bounds on the normal approximation of random vectors woth coordinates living in a fixed Wiener chaos.

9.3 Gamma approximation

We now state a result that can be obtained by combining Malliavin calculus with the Gamma approximations discussed in the second part of Section 8.2 (we shall use the same notation introduced therein). The proof (left to the reader) makes use of (7.34), and of arguments analogous to those displayed in the proof of Theorem 9.1.

Theorem 9.3 *Fix $\nu > 0$ and let $F(\nu)$ have a centered Gamma distribution with parameter ν. Let $G \in \mathbb{D}^{1,2}$ be such that $E(G) = 0$ and the law of G is absolutely continuous with respect to the Lebesgue measure. Then:*

$$d_{\mathcal{G}_1}(G, F(\nu)) \leq K_1 E[(2\nu + 2G - \langle DG, -DL^{-1}G\rangle_{\mathfrak{H}})^2]^{1/2}, \tag{9.11}$$

and, if $\nu \geq 1$ is an integer,

$$d_{\mathcal{G}_2}(G, F(\nu)) \leq K_2 E[(2\nu + 2G - \langle DG, -DL^{-1}G\rangle_{\mathfrak{H}})^2]^{1/2}, \tag{9.12}$$

where $\mathcal{G}_1$ and $\mathcal{G}_2$ are defined in (8.11)–(8.12), $K_1 \triangleq \max\{1, 1/\nu + 2/\nu^2\}$ and $K_2 \triangleq \max\{\sqrt{2\pi/\nu}, 1/\nu + 2/\nu^2\}$.

We will come back to Theorem 9.3 in Section 10.3, where we will present some characterizations of *non-central limit theorems* on a fixed Wiener chaos.

10 Limit Theorems on Wiener chaos

Let $X = \{X(h) : h \in \mathfrak{H}\}$ be an isonormal Gaussian process. In this section, we focus on the Gaussian and Gamma approximations of (vectors of) random variables of the type $F = I_q(f)$, where $q \geq 2$ and $f \in \mathfrak{H}^{\odot q}$. We recall that, according to the chaotic representation property stated in Proposition 6.1-3, random variables of this form are the basic building blocks of every square-integrable functional of X.

In order to appreciate the subtlety of the issues faced in this section, we list some well-known properties of the laws of chaotic random variables.

- If $q = 2$, then there exists a sequence $\{\xi_i : i \geq 1\}$ of i.i.d. centered standard Gaussian random variables such that

$$I_2(f) = \sum_{i=1}^{\infty} \lambda_i \left(\xi_i^2 - 1\right), \tag{10.1}$$

  where the series converges in $L^2(\mathbb{P})$, and $\{\lambda_i : i \geq 1\}$ is the sequence of eigenvalues of the Hilbert-Schmidt operator (from $\mathfrak{H}$ into $\mathfrak{H}$) given by $h \mapsto f \otimes_1 h$, where $\otimes_1$ indicates a contraction of order 1. In particular, $I_2(f)$ admits some finite exponential moment, and the law of $I_2(f)$ is determined by its moments.

- If $q \geq 3$, the law of $I_q(f)$ may not be determined by its moments. See Slud [87].

- For every $q \geq 2$, the random variable $I_q(f)$ *cannot be Gaussian*. See [35, Chapter VI].

- For $q \geq 3$, and except for trivial cases, there does not exist a general explicit formula for the characteristic function of $I_q(f)$.

Note that in the next section we will focus on the total variation distance d_{TV}. However, it will be clear later on that (thanks to Theorem 9.1) all the results extend without difficulties to the Fortet-Mourier, Wasserstein or Kolmogorov distances.

10.1 CLTs in dimension one

Let $N \sim \mathcal{N}(0,1)$. Fix $q \geq 2$ and consider an element of the qth Wiener chaos of X with the form $F = I_q(f)$, where the kernel f is in $\mathfrak{H}^{\odot q}$.

Remark. Since $\mathbb{E}[F] = 0$, the fourth cumulant of F is given by

$$\chi_4(F) = \mathbb{E}\left[F^4\right] - 3\mathbb{E}\left[F^2\right]^2. \tag{10.2}$$

Observe also that $\mathbb{E}\left[N^4\right] = 3$.

We have that $\mathbb{E}\left[\frac{1}{q}\|DF\|^2_{\mathfrak{H}}\right] = \mathbb{E}\left[F^2\right] = q!\,\|f\|^2_{\mathfrak{H}^{\otimes q}}$, and also, by (9.4),

$$d_{TV}(F,N) \leq 2\left|1 - q!\,\|f\|^2_{\mathfrak{H}^{\otimes q}}\right| + 2\sqrt{\mathbf{Var}\left(\frac{1}{q}\|DF\|^2_{\mathfrak{H}}\right)}. \tag{10.3}$$

The following result, which is partially based on the moment formula (7.35), shows that (10.3) yields indeed an important simplification of the method of moments and cumulants.

Proposition 10.1 (See [57]) *Let the above notation and assumptions prevail. Then, the following hold.*

1.

$$\mathbf{Var}\left(\frac{1}{q}\|DF\|^2_{\mathfrak{H}}\right) = q^2 \sum_{p=1}^{q-1} (p-1)!^2 \binom{q-1}{p-1}^4 (2q-2p)!\,\left\|f\widetilde{\otimes}_p f\right\|^2_{\mathfrak{H}^{\otimes 2(q-p)}}. \tag{10.4}$$

2.

$$\chi_4(F) = \mathbb{E}\left[F^4\right] - 3 = 3q\sum_{p=1}^{q-1} p!\,(p-1)!\binom{q}{p}^2\binom{q-1}{p-1}^2 (2q-2p)!\left\|f\widetilde{\otimes}_p f\right\|^2_{\mathfrak{H}^{\otimes 2(q-p)}}. \quad (10.5)$$

3.

$$0 \leq \frac{1}{3q}\chi_4(F) \leq \mathbf{Var}\left(\frac{1}{q}\left\|DF\right\|^2_{\mathfrak{H}}\right) \leq \frac{q-1}{3q}\chi_4(F). \quad (10.6)$$

4.

$$d_{TV}(N,F) \leq 2\left[\left|1-\mathbb{E}\left[F^2\right]\right| + \sqrt{\frac{q-1}{3q}\chi_4(F)}\right]. \quad (10.7)$$

Proof. It suffices to prove the statement when $\mathfrak{H} = L^2(Z,\mathcal{Z},\mu)$, with μ σ-finite and without atoms. In this case, one has that $D_zF = qI_{q-1}(f(\cdot,z))$ and, by the multiplication formula,

$$(D_zF)^2 = q^2\sum_{r=0}^{q-1} r!\binom{q-1}{r}^2 I_{2(q-1-r)}\left(f(\cdot,z)\otimes_r f(\cdot,z)\right).$$

It follows that

$$\begin{aligned}
\frac{1}{q}\left\|DF\right\|^2_{L^2(\mu)} &= \frac{1}{q}\int_Z (D_zF)^2\,\mu(dz) \\
&= q\sum_{r=0}^{q-1} r!\binom{q-1}{r}^2 I_{2(q-1-r)}\left(f\otimes_{r+1} f\right) \\
&= q\sum_{p=1}^{q}(p-1)!\binom{q-1}{p-1}^2 I_{2(q-p)}\left(f\otimes_p f\right), \quad (10.8)
\end{aligned}$$

so that (10.4) follows immediately from the isometry and orthogonality properties of multiple Wiener-Itô integrals. To prove (10.5), we start by observing that, thanks to (7.35) in the case $n=2$,

$$\mathbb{E}\left[F^4\right] = 3\left[F^2\times\frac{1}{q}\left\|DF\right\|^2_{L^2(\mu)}\right]. \quad (10.9)$$

Now, by virtue of the multiplication formula,

$$F^2 = \sum_{p=0}^{q} p!\binom{q}{p}^2 I_{2(q-p)}\left(f\otimes_p f\right),$$

and, by plugging (10.8) into (10.9), we obtain

$$\mathbb{E}\left[F^4\right] = 3q\sum_{p=1}^{q} p!\,(p-1)!\binom{q}{p}^2\binom{q-1}{p-1}^2 (2q-2p)!\left\|f\widetilde{\otimes}_p f\right\|^2_{\mathfrak{H}^{\otimes 2(q-p)}},$$

which is equivalent to (10.5) (note that $\| f\widetilde{\otimes}_q f \|^2_{\mathfrak{H}^{\otimes 0}} = \|f\|^4_{\mathfrak{H}^{\odot q}}$ by definition). Formula (10.6) is obtained by comparing the RHS of (10.4) and (10.5). Finally (10.7) follows from (10.3). ■

Formula (10.7) implies that the fourth cumulant controls the distance between the law of F and a standard Gaussian distribution. The following statement exploits this result, in order to give a neat and exhaustive characterization of CLTs on a fixed Wiener chaos.

Theorem 10.1 *Let $q \geq 2$ and let $F_n = I_q(f_n), n \geq 1$, be a sequence in the qth Wiener chaos such that $\mathbb{E}(F_n^2) \to 1$ as $n \to \infty$. Then, the following five conditions are equivalent as $n \to \infty$.*

(i) $F_n \overset{Law}{\longrightarrow} N \sim \mathcal{N}(0,1)$.

(ii) $d_{TV}(F_n, N) \to 0$.

(iii) *For every $p = 1, ..., q-1$,* $\left\| f_n \widetilde{\otimes}_p f_n \right\|^2_{\mathfrak{H}^{\otimes 2(q-p)}} \to 0$.

(iv) $\mathbf{Var}\left(\frac{1}{q}\|DF_n\|^2_{\mathfrak{H}}\right) \to 0$.

(v) $\chi_4(F_n) = \mathbb{E}\left[F_n^4\right] - 3\mathbb{E}\left(F_n^2\right)^2 \to 0$.

Proof. In view of Proposition 10.1, the implications (v) ⇒ (iv) ⇒ (iii) ⇒ (ii) ⇒ (i) are immediate. To prove (i) ⇒ (v), one can combine the contraction inequality (5.6) with the assumption that $\mathbb{E}\left[F_n^2\right] \to 1$, in order to deduce that, for every $p > 2$ $\sup_n \mathbb{E}|F_n|^p < \infty$. This last relation implies the desired conclusion (for instance, by a uniform integrability argument). ■

Remarks. (1) The equivalence between (i), (iii) and (v) in the statement of Theorem 10.1 has been first proved in [67] by means of the Dambis-Dubins-Schwarz theorem. The equivalence between (iv) and (i) comes from [66]. Incidentally, it is very interesting to compare our techniques based on Stein's method with those developed in [66], where the authors make use of the *differential equation* satisfied by the characteristic function of $N \sim \mathcal{N}(0,1)$. Namely, since $\psi_N(t) = \mathbb{E}[\exp(itN)] = \exp\left(-t^2/2\right)$, one has that ψ_N is the unique solution of

$$\psi'(t) + t\psi(t) = 0, \quad \psi(0) = 1.$$

This approach is close to the so-called "Tikhomirov method" – see [95].

(2) We stress that the implication (ii) ⇒ (i) is not trivial, since the topology induced by the total variation distance (on the class of probabilities on $\mathbb{R}$) is strictly stronger than the topology of weak convergence.

(3) The implication (v) ⇒ (i) yields that, in order to prove a central limit theorem on a fixed Wiener chaos, *it is sufficient to check that the first two even moments of the concerned sequence converge, respectively, to* 1 *and* 3. This is the announced "drastic" simplification of the method of moments and cumulants, as described in Section 2.2.

(4) In [67] it is also proved that $\left\| f_n \widetilde{\otimes}_p f_n \right\|^2_{\mathfrak{H}^{\otimes 2(q-p)}} \to 0$, for every $p = 1, ..., q-1$, if and only if the non-symmetrized norm $\|f_n \otimes_p f_n\|^2_{\mathfrak{H}^{\otimes 2(q-p)}}$ converges to 0 for every $p = 1, ..., q-1$.

(5) Theorem 10.1 and its multidimensional extensions (see the next section) have been applied to a variety of frameworks, such as: quadratic functionals of bivariate Gaussian processes (see

[21]), quadratic functionals of fractional processes (see [67]), high-frequency limit theorems on homogeneous spaces (see [47, 48]), self-intersection local times of fractional Brownian motion (see [33, 66]), needleets analysis on the sphere (see [1]), power variations of iterated processes (see [55]), weighted variations of fractional processes (see [54, 63]) and of related random functions (see [3, 16]).

10.2 Multi-dimensional CLTs

We keep the framework of the previous section. We are now interested in the normal approximation, in the Wasserstein distance, of *random vectors* of multiple Wiener-Itô integrals (of possibly different orders). In particular, our main tool is the following consequence of Theorem 9.2.

Proposition 10.2 (See [60]) *Fix $d \geq 2$ and $1 \leq q_1 \leq \ldots \leq q_d$. Consider a vector $F = (F_1, \ldots, F_d) = (I_{q_1}(f_1), \ldots, I_{q_d}(f_d))$ with $f_i \in \mathfrak{H}^{\odot q_i}$ for any $i = 1 \ldots, d$. Let $Z \sim \mathcal{N}_d(0, C)$ be a d-dimensional Gaussian vector, with a positive definite covariance matrix C. Then,*

$$d_W(F, Z) \leq \|C^{-1}\|_{op} \, \|C\|_{op}^{1/2} \sqrt{\sum_{1 \leq i,j \leq d} \mathbb{E}\left[\left(C(i,j) - \frac{1}{q_j} \langle DF_i, DF_j \rangle_{\mathfrak{H}}\right)^2\right]}. \tag{10.10}$$

Plainly, the proof of Proposition 10.2 is immediately deduced from the fact that, for every $q \geq 1$, $L^{-1} I_q(f) = -q^{-1} I_q(f)$.

When applying Proposition 10.2 in concrete situations, one can use the following result in order to evaluate the RHS of (10.10).

Lemma 10.1 (See [60]) *Let $F = I_p(f)$ and $G = I_q(g)$, with $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$ $(p, q \geq 1)$. Let a be a real constant. If $p = q$, one has the estimate:*

$$\begin{aligned} \mathbb{E}\left[\left(a - \frac{1}{p} \langle DF, DG \rangle_{\mathfrak{H}}\right)^2\right] &\leq (a - p! \langle f, g \rangle_{\mathfrak{H}^{\otimes p}})^2 \\ &+ \frac{p^2}{2} \sum_{r=1}^{p-1} (r-1)!^2 \binom{p-1}{r-1}^4 (2p-2r)! \big(\|f \otimes_{p-r} f\|_{\mathfrak{H}^{\otimes 2r}}^2 + \|g \otimes_{p-r} g\|_{\mathfrak{H}^{\otimes 2r}}^2\big). \end{aligned} \tag{10.11}$$

On the other hand, if $p < q$, one has that

$$\begin{aligned} \mathbb{E}\left[\left(a - \frac{1}{q} \langle DF, DG \rangle_{\mathfrak{H}}\right)^2\right] &\leq a^2 + p!^2 \binom{q-1}{p-1}^2 (q-p)! \|f\|_{\mathfrak{H}^{\otimes p}}^2 \|g \otimes_{q-p} g\|_{\mathfrak{H}^{\otimes 2p}} \\ &+ \frac{p^2}{2} \sum_{r=1}^{p-1} (r-1)!^2 \binom{p-1}{r-1}^2 \binom{q-1}{r-1}^2 (p+q-2r)! \big(\|f \otimes_{p-r} f\|_{\mathfrak{H}^{\otimes 2r}}^2 + \|g \otimes_{q-r} g\|_{\mathfrak{H}^{\otimes 2r}}^2\big). \end{aligned} \tag{10.12}$$

Remark. One crucial consequence of Lemma 10.1 is that, in order to estimate the right-hand side of (10.10), it is sufficient to asses the quantity $\|f_i \otimes_r f_i\|_{\mathfrak{H}^{\otimes 2(q_i - r)}}$ (for any $i \in \{1, \ldots, d\}$ and $r \in \{1, \ldots, q_i - 1\}$) on the one hand, and $q_i! \langle f_i, f_j \rangle_{\mathfrak{H}^{\otimes q_i}} = \mathbb{E}\left[I_{q_i}(f_i) I_{q_j}(f_j)\right]$ (for any $1 \leq i, j \leq d$ such that $q_i = q_j$) on the other hand.

Proof of Lemma 10.1 (see also [66, Lemma 2]). Without loss of generality, we can assume that $\mathfrak{H} = L^2(Z, \mathcal{Z}, \mu)$, where $(A, \mathcal{Z})$ is a Polish space, and μ is a σ-finite and non-atomic measure. Thus, we can write

$$\begin{aligned}
\langle DF, DG\rangle_{\mathfrak{H}} &= p\,q\,\langle I_{p-1}(f), I_{q-1}(g)\rangle_{\mathfrak{H}} = p\,q\int_Z I_{p-1}\big(f(\cdot,z)\big)I_{q-1}\big(g(\cdot,z)\big)\mu(dz)\\
&= p\,q\int_A \sum_{r=0}^{p\wedge q-1} r!\binom{p-1}{r}\binom{q-1}{r} I_{p+q-2-2r}\big(f(\cdot,z)\widetilde{\otimes}_r g(\cdot,z)\big)\mu(dz)\\
&= p\,q\sum_{r=0}^{p\wedge q-1} r!\binom{p-1}{r}\binom{q-1}{r} I_{p+q-2-2r}(f\widetilde{\otimes}_{r+1}g)\\
&= p\,q\sum_{r=1}^{p\wedge q}(r-1)!\binom{p-1}{r-1}\binom{q-1}{r-1} I_{p+q-2r}(f\widetilde{\otimes}_r g).
\end{aligned}$$

It follows that

$$E\left[\left(a - \frac{1}{q}\langle DF, DG\rangle_{\mathfrak{H}}\right)^2\right] \tag{10.13}$$

$$= \begin{cases} a^2 + p^2\sum_{r=1}^{p}(r-1)!^2\binom{p-1}{r-1}^2\binom{q-1}{r-1}^2(p+q-2r)!\|f\widetilde{\otimes}_r g\|^2_{\mathfrak{H}^{\otimes(p+q-2r)}} & \text{if } p<q,\\[2ex] (a - p!\langle f,g\rangle_{\mathfrak{H}^{\otimes p}})^2 + p^2\sum_{r=1}^{p-1}(r-1)!^2\binom{p-1}{r-1}^4(2p-2r)!\|f\widetilde{\otimes}_r g\|^2_{\mathfrak{H}^{\otimes(2p-2r)}} & \text{if } p=q. \end{cases}$$

If $r < p \le q$ then

$$\begin{aligned}
\|f\widetilde{\otimes}_r g\|^2_{\mathfrak{H}^{\otimes(p+q-2r)}} &\le \|f\otimes_r g\|^2_{\mathfrak{H}^{\otimes(p+q-2r)}} = \langle f\otimes_{p-r} f, g\otimes_{q-r} g\rangle_{\mathfrak{H}^{\otimes 2r}}\\
&\le \|f\otimes_{p-r} f\|_{\mathfrak{H}^{\otimes 2r}}\|g\otimes_{q-r} g\|_{\mathfrak{H}^{\otimes 2r}}\\
&\le \frac{1}{2}\left(\|f\otimes_{p-r} f\|^2_{\mathfrak{H}^{\otimes 2r}} + \|g\otimes_{q-r} g\|^2_{\mathfrak{H}^{\otimes 2r}}\right).
\end{aligned}$$

If $r = p < q$, then

$$\|f\widetilde{\otimes}_p\, g\|^2_{\mathfrak{H}^{\otimes(q-p)}} \le \|f\otimes_p\, g\|^2_{\mathfrak{H}^{\otimes(q-p)}} \le \|f\|^2_{\mathfrak{H}^{\otimes p}}\|g\otimes_{q-p} g\|_{\mathfrak{H}^{\otimes 2p}}.$$

If $r = p = q$, then $f\widetilde{\otimes}_p g = \langle f, g\rangle_{\mathfrak{H}^{\otimes p}}$. By plugging these last expressions into (10.13), we deduce immediately the desired conclusion. ■

The combination of the results presented in this section with Theorem 10.1 lead to the following statement, which is a collection of the main findings contained in the papers by Peccati and Tudor [73] and Nualart and Ortiz-Latorre [66].

Theorem 10.2 (See [66, 73]) *Fix $d \ge 2$ and let $C = \{C(i,j) : i,j = 1,...,d\}$ be a $d\times d$ positive definite matrix. Fix integers $1 \le q_1 \le \ldots \le q_d$. For any $n \ge 1$ and $i = 1,\ldots,d$, let $f_i^{(n)}$ belong to $\mathfrak{H}^{\odot q_i}$. Assume that*

$$F^{(n)} = (F_1^{(n)},\ldots,F_d^{(n)}) = (I_{q_1}(f_1^{(n)}),\ldots,I_{q_d}(f_d^{(n)})) \quad n\ge 1,$$

is such that

$$\lim_{n\to\infty}\mathbb{E}[F_i^{(n)}F_j^{(n)}] = C(i,j), \quad 1\le i,j\le d. \tag{10.14}$$

Then, as $n\to\infty$, the following four assertions are equivalent:

(i) *For every* $1 \le i \le d$, $F_i^{(n)}$ *converges in distribution to a centered Gaussian random variable with variance* $C(i,i)$.

(ii) *For every* $1 \le i \le d$, $\mathbb{E}\left[(F_i^{(n)})^4\right] \to 3C(i,i)^2$.

(iii) *For every* $1 \le i \le d$ *and every* $1 \le r \le q_i - 1$, $\|f_i^{(n)} \otimes_r f_i^{(n)}\|_{\mathfrak{H}^{\otimes 2(q_i - r)}} \to 0$.

(iv) *The vector* $F^{(n)}$ *converges in distribution to a* d*-dimensional Gaussian vector* $\mathcal{N}_d(0, C)$.

Moreover, if $C(i,j) = \delta_{ij}$, *where* δ_{ij} *is the Kronecker symbol, then either one of conditions (i)–(iv) above is equivalent to the following:*

(v) *For every* $1 \le i \le d$, $\|DF_i^{(n)}\|_{\mathfrak{H}}^2 \xrightarrow{L^2} q_i$.

Remark. The crucial implication in the statement of Theorem 10.2 is (i) $\Rightarrow$ (iv), yielding that, for random vectors composed of chaotic random variables and verifying the asymptotic covariance condition (10.14), *componentwise convergence in distribution towards a Gaussian vector always implies joint convergence.* This fact is extremely useful for applications: see for instance [3], [33], [48], [54] and [55].

We conclude this section by pointing out the remarkable fact that, for vectors of multiple Wiener-Itô integrals of arbitrary length, *the Wasserstein distance metrizes the weak convergence towards a Gaussian vector with positive definite covariance.* Once again, this result is not trivial, since the topology induced by the Wasserstein distance is stronger than the topology of weak convergence.

Proposition 10.3 (See [60]) *Fix* $d \ge 2$, *let* C *be a positive definite* $d \times d$ *symmetric matrix, and let* $1 \le q_1 \le \ldots \le q_d$. *Consider vectors*

$$F^{(n)} = (F_1^{(n)}, \ldots, F_d^{(n)}) = (I_{q_1}(f_1^{(n)}), \ldots, I_{q_d}(f_d^{(n)})), \quad n \ge 1,$$

with $f_i^{(n)} \in \mathfrak{H}^{\odot q_i}$ *for every* $i = 1 \ldots, d$. *Assume moreover that* $F^{(n)}$ *satisfies condition (10.14). Then, as* $n \to \infty$, *the following three conditions are equivalent:*

(a) $d_W(F^{(n)}, Z) \to 0$.

(b) *For every* $1 \le i \le d$, $q_i^{-1}\|DF_i^{(n)}\|_{\mathfrak{H}}^2 \xrightarrow{L^2} C(i,i)$ *and, for every* $1 \le i \ne j \le d$,

$$\langle DF_i, -DL^{-1}F_j\rangle_{\mathfrak{H}} = q_j^{-1}\langle DF_i, DF_j\rangle_{\mathfrak{H}} \xrightarrow{L^2} C(i,j).$$

(c) $F^{(n)}$ *converges in distribution to* $Z \sim \mathcal{N}_d(0, C)$.

Proof. Since convergence in the Wasserstein distance implies convergence in distribution, the implication (a) $\to$ (c) is trivial. The implication (b) $\to$ (a) is a consequence of relation (10.10). Now assume that (c) is verified, that is, $F^{(n)}$ converges in law to $Z \sim \mathcal{N}_d(0, C)$ as n goes to infinity. By Theorem 10.2 we have that, for any $i \in \{1, \ldots, d\}$ and $r \in \{1, \ldots, q_i - 1\}$,

$$\|f_i^{(n)} \otimes_r f_i^{(n)}\|_{\mathfrak{H}^{\otimes 2(q_i - r)}} \underset{n\to\infty}{\longrightarrow} 0.$$

By combining Corollary 10.2 with Lemma 10.1, one therefore easily deduces that, since (10.14) is in order, condition (b) must necessarily be satisfied. ■

10.3 A non-central limit theorem (with bounds)

We now present (without proofs) two statements concerning the Gamma approximation of multiple integrals of even order $q \geq 2$. The first result, which is taken from [57], provides an explcit representation for the quantities appearing on the RHS of (9.11) and (9.12).

Proposition 10.4 (See [57]) *Let $q \geq 2$ be an even integer, and let $G = I_q(g)$, where $g \in \mathfrak{H}^{\odot q}$. Then,*

$$\begin{aligned}
& \mathbb{E}[(2\nu + 2G - \langle DG, -DL^{-1}G\rangle_{\mathfrak{H}})^2] = \mathbb{E}[(2\nu + 2G - q^{-1}\|DG\|^2_{\mathfrak{H}})^2] \qquad (10.15)\\
\leq\ & (2\nu - q!\,\|g\|^2_{\mathfrak{H}^{\otimes q}})^2 + \\
& \quad + q^2 \sum_{\substack{r\in\{1,\dots,q-1\}\\ r\neq q/2}} (2q-2r)!(r-1)!^2 \binom{q-1}{r-1}^4 \|g\otimes_r g\|^2_{\mathfrak{H}^{\otimes 2(q-r)}} + \\
& \qquad\qquad + 4q!\,\left\| c_q^{-1} \times g\widetilde{\otimes}_{q/2} g - g\right\|^2_{\mathfrak{H}^{\otimes q}},
\end{aligned}$$

where

$$c_q = \frac{1}{(q/2)!\binom{q-1}{q/2-1}^2} = \frac{4}{(q/2)!\binom{q}{q/2}^2}. \qquad (10.16)$$

The next statement, which is a main result of [56], contains a "non-central" analogous of Theorem 10.1. Recall the definition of the centered Gamma random variables $F(\nu)$, $\nu > 0$, given in (8.2).

Theorem 10.3 (See [56]) *Fix $\nu > 0$, as well as an even integer $q \geq 2$. Define c_q as in (10.16). Then, for any sequence $\{f_k\}_{k\geq 1} \subset \mathfrak{H}^{\odot q}$ verifying*

$$\lim_{k\to\infty} q!\|f_k\|^2_{\mathfrak{H}^{\otimes n}} = \lim_{k\to\infty} \mathbb{E}\big[I_q(f_k)^2\big] = \mathbf{Var}\,(F(\nu)) = 2\nu, \qquad (10.17)$$

the following six conditions are equivalent:

- **(i)** $\lim_{k\to\infty} \mathbb{E}[I_q(f_k)^3] = \mathbb{E}[F(\nu)^3] = 8\nu$ *and* $\lim_{k\to\infty} \mathbb{E}[I_q(f_k)^4] = \mathbb{E}[F(\nu)^4] = 48\nu + 12\nu^2$*;*
- **(ii)** $\lim_{k\to\infty} \mathbb{E}[I_q(f_k)^4] - 12\mathbb{E}[I_q(f_k)^3] = 12\nu^2 - 48\nu$*;*
- **(iii)** $\lim_{k\to\infty} \|f_k\widetilde{\otimes}_{q/2} f_k - c_q \times f_k\|_{\mathfrak{H}^{\otimes q}} = 0$ *and* $\lim_{k\to\infty} \|f_k\widetilde{\otimes}_p f_k\|_{\mathfrak{H}^{\otimes 2(q-p)}} = 0$*, for every* $p = 1, ..., q-1$ *such that* $p \neq q/2$*;*
- **(iv)** $\lim_{k\to\infty} \|f_k\widetilde{\otimes}_{q/2} f_k - c_q \times f_k\|_{\mathfrak{H}^{\otimes q}} = 0$ *and* $\lim_{k\to\infty} \|f_k \otimes_p f_k\|_{\mathfrak{H}^{\otimes 2(q-p)}} = 0$*, for every* $p = 1, ..., q-1$ *such that* $p \neq q/2$*;*
- **(v)** *as* $k\to\infty$, $\|D[I_q(f_k)]\|^2_{\mathfrak{H}} - 2qI_q(f_k) \longrightarrow 2q\nu$ *in* L^2*;*
- **(vi)** *as* $k\to\infty$*, the sequence* $\{I_q(f_k)\}_{k\geq 1}$ *converges in distribution to* $F(\nu)$*.*

Remark. In [56], Theorem 10.3 is not proved with Stein's method, but rather by implementing the "differential approach" initiated by Nualart and Ortiz-Latorre in [66]. However, it is not difficult to see that (9.11), (9.12) and (10.15) can be combined in order to deduce an alternate proof of the implications (iv) ⇒ (v) ⇒ (vi).

11 Two examples

The theory developed in the previous sections (along with its refinements and generalizations – see Section 12) has been already applied in a variety of frameworks. In particular:

- In [57], Theorem 9.1 and Proposition 10.1 are applied in order to deduce explicit Berry-Esséen bounds for the so-called *Breuer-Major CLT* (see [5]), involving Hermite-type transformations of fractional Brownian motion. This analysis is further developed in [4], [58] and [60]
- The paper [58] contains applications to Toepliz quadratic forms in continuous time – see e.g. [30] and the references therein.
- In [60] one can also find multidimensional generalizations of Chatterjee's result (9.7).
- Reference [62] contains an application of (9.3) to the proof of infinite-dimensional second-order Poincaré inequalities on Wiener space.
- In [64], relation (7.30) is exploited in order to provide a new explicit expression for the densities of functionals of isonormal Gaussian processes.
- In [97], one can find applications to tail bounds on Gaussian functionals and polymer models.

Remark. Apart from the previous references, the applications to fractional Brownian motion and density estimation are discussed in the lecture notes [53].

In what follows, we shall present two further applications of the previous results. The first one (basically taken from [58]) focuses on exploding quadratic functionals of a Brownian sheet – thus completing the discussion contained in Section 2. The second one involves Hermite transformations of multiparameter Ornstein-Uhlenebck Gaussian processes, and is new (albeit it is inspired by the last section of [70]).

11.1 Exploding Quadratic functionals of a Brownian sheet

11.1.1 Statement of the problem

Let $d \geq 1$, and let

$$\mathbf{W} = \left\{ \mathbf{W}\left(t_1, ..., t_d\right) : \left(t_1, ..., t_d\right) \in [0,1]^d \right\}$$

be a standard Brownian sheet on $[0,1]^d$. Recall that this means that $\mathbf{W}$ is a continuous centered Gaussian process with a covariance function given by

$$\mathbb{E}\left[\mathbf{W}\left(t_1, ..., t_d\right)\mathbf{W}\left(s_1, ..., s_d\right)\right] = \prod_{j=1}^{d}\left(t_j \wedge s_j\right).$$

By using an appropriate version of the so-called *Jeulin Lemma* (see [36, Lemma 1, p. 44]), one can prove that

$$\int_0^1 \cdots \int_0^1 \left(\frac{\mathbf{W}(t_1, ..., t_d)}{t_1 \cdots t_d}\right)^2 dt_1 \cdots dt_d = \infty, \quad \text{a.s.-}\mathbb{P}.$$

For $\varepsilon \in (0,1)$, we can now define the random variable

$$B_\varepsilon^d = \int_\varepsilon^1 \cdots \int_\varepsilon^1 \left(\frac{\mathbf{W}(t_1, ..., t_d)}{t_1 \cdots t_d}\right)^2 dt_1 \cdots dt_2.$$

A standard computation yields $\mathbb{E}\left[B_\varepsilon^d\right] = (\log 1/\varepsilon)^d$, and also

$$\mathbf{Var}\left(B_\varepsilon^d\right) \approx (4 \log 1/\varepsilon)^d, \quad \text{as } \varepsilon \to 0.$$

By setting

$$\widetilde{B}_\varepsilon^d \triangleq \frac{B_\varepsilon^d - (\log 1/\varepsilon)^d}{(4 \log 1/\varepsilon)^{\frac{d}{2}}},$$

one can therefore state the following generalization of **Problem I**, as stated at the end of Section 2.1.

Problem II. *Prove that, as $\varepsilon \to 0$, $\widetilde{B}_\varepsilon^d \overset{\text{Law}}{\to} N \sim \mathcal{N}(0,1)$.*

Remark. See [21] for applications of quadratic functionals of Brownian sheets to tests of independence.

11.1.2 Interlude: optimal rates for second chaos sequences

In order to give an exhaustive answer to Problem II, we state (without proof) a result concerning sequences in the second Wiener chaos of a given isonormal Gaussian process $X = \{X(h) : h \in \mathfrak{H}\}$. It gives a simple criterion (based on cumulants) allowing to determine whether, for a sequence in the second chaos, the rate of convergence implied by (10.7) is optimal. Note that the forthcoming formula (11.1) is just a rewriting of (10.7), which we added for the convenience of the reader. We also use the notation

$$\Phi(z) = \mathbb{P}[N \le z], \quad \text{where} \quad N \sim \mathcal{N}(0,1).$$

Proposition 11.1 (See [58]) *Let $F_n = I_2(f_n)$, $n \ge 1$, be such that $f_n \in \mathfrak{H}^{\odot 2}$, and write $\chi_p^{(n)} = \chi_p(F_n)$, $p \ge 1$. Assume that $\chi_2^{(n)} = E(F_n^2) \longrightarrow 1$ as $n \to \infty$. Then, as $n \to \infty$, $F_n \overset{\text{Law}}{\longrightarrow} N \sim \mathcal{N}(0,1)$ if and only if $\chi_4^{(n)} \longrightarrow 0$. In this case, we have moreover*

$$d_{Kol}(F_n, N) \le \sqrt{\frac{\chi_4^{(n)}}{6} + (\chi_2^{(n)} - 1)^2}. \tag{11.1}$$

If, in addition, we have, as $n \to \infty$,

$$\frac{\chi_2^{(n)} - 1}{\frac{\chi_4^{(n)}}{6} + (\chi_2^{(n)} - 1)^2} \longrightarrow 0, \tag{11.2}$$

$$\frac{\chi_3^{(n)}}{\sqrt{\frac{\chi_4^{(n)}}{6} + (\chi_2^{(n)} - 1)^2}} \longrightarrow \alpha \quad and \quad \frac{\chi_8^{(n)}}{\left(\frac{\chi_4^{(n)}}{6} + (\chi_2^{(n)} - 1)^2\right)^2} \longrightarrow 0, \tag{11.3}$$

then

$$\frac{\mathbb{P}(F_n \leq z) - \Phi(z)}{\sqrt{\frac{\chi_4^{(n)}}{6} + (\chi_2^{(n)} - 1)^2}} \longrightarrow \frac{\alpha}{3!} \frac{1}{\sqrt{2\pi}} \left(1 - z^2\right) e^{-\frac{z^2}{2}}, \quad as\ n \to \infty. \tag{11.4}$$

In particular, if $\alpha \neq 0$, there exists $c \in (0,1)$ and $n_0 \geq 1$ such that, for any $n \geq n_0$,

$$d_{TV}(F_n, N) \geq d_{Kol}(F_n, N) \geq c \sqrt{\frac{\chi_4^{(n)}}{6} + (\chi_2^{(n)} - 1)^2}. \tag{11.5}$$

11.1.3 A general statement

The next result provides an exhaustive solution to Problem II (and therefore to Problem I).

Proposition 11.2 *For every $d \geq 1$, there exist constants $0 < c(d) < C(d) < \infty$ and $0 < \eta(d) < 1$, depending uniquely on d, such that, for every $\varepsilon \in (0,1)$,*

$$d_{TV}[\widetilde{B}_\varepsilon^d, N] \leq C(d)(\log 1/\varepsilon)^{-d/2} \tag{11.6}$$

and, for $\varepsilon < \eta(d)$,

$$d_{TV}[\widetilde{B}_\varepsilon^d, N] \geq c(d)(\log 1/\varepsilon)^{-d/2}. \tag{11.7}$$

This yields that, as $\varepsilon \to 0$, $\widetilde{B}_\varepsilon^d \overset{\text{Law}}{\to} N \sim \mathcal{N}(0,1)$.

Proof. We denote by

$$\widetilde{\chi}_j(d, \varepsilon), \quad j = 1, 2, ...,$$

the sequence of the cumulants of the random variable $\widetilde{B}_\varepsilon^d$. We deal separately with the cases $d = 1$ and $d \geq 2$.

(Case $d = 1$) As already observed, in this case, $\mathbf{W}$ is a standard Brownian motion on $[0,1]$, so that $\widetilde{B}_\varepsilon^1$ takes the form $\widetilde{B}_\varepsilon^1 = I_2(f_\varepsilon)$, where I_2 indicates a double Wiener-Itô integral with respect to $\mathbf{W}$, and

$$f_\varepsilon(x,y) = (4 \log 1/\varepsilon)^{-1/2} [(x \vee y \vee \varepsilon)^{-1} - 1]. \tag{11.8}$$

The conclusion now follows from Proposition 11.1 and (2.12).

(Case $d \geq 2$) In this case, $\widetilde{B}_\varepsilon^d$ has the form $\widetilde{B}_\varepsilon^d = I_2(f_\varepsilon^d)$, with

$$f_\varepsilon^d(x_1, ..., x_d; y_1, ..., y_d) = (4 \log 1/\varepsilon)^{-d/2} \prod_{j=1}^d [(x_j \vee y_j \vee \varepsilon)^{-1} - 1]. \tag{11.9}$$

By using an appropriate modification of (2.11), one sees that the following relation holds

$$(2^{j-1}(j-1)!)^{-1} \times \widetilde{\chi}_j(d, \varepsilon) = [(2^{j-1}(j-1)!)^{-1} \times \widetilde{\chi}_j(1, \varepsilon)]^d,$$

so that the conclusion derives once again from Proposition 11.1 and (2.12). ■

Remark. The example developed in this section shows how the Malliavin/Stein approach can overcome most of the difficulties **D1** – **D5**, that were pointed out at the end of Sections 2.2 and 2.3 in connection with the method of cumulants and with random time-changes. In particular, one has that

- Stein's method allows in this case to deduce *exact* rates of convergence in the sense of the total variation (but also Kolmogorov and Wasserstein) distance. This successfully addresses **D1** and **D5**.
- The convergence $\widetilde{B}_\varepsilon^d \overset{\text{Law}}{\to} N$ is now implied by the simple condition $\chi_4^{(n)} \longrightarrow 0$. Moreover, to obtain lower bounds one must merely verify the three relations at (11.2) and (11.3). This eliminates the difficulty pointed out in **D2.**
- Finally, our techniques allow to deal *directly* with quadratic functionals of a Brownian sheet, without making use of *any* underlying martingale structure. This overcomes the drawbacks of random time-changes described at **D4**.

In the next section, we will describe a situation involving non-quadratic transformations (thus adressing point **D3** in the above quoted list).

11.2 Hermite functionals of Ornstein-Uhlenbeck sheets

Let $d \geq 1$. Let G be a centered Gaussian measure over $\mathbb{R}^d$, with control given by the Lebesgue measure $\mu(dx_1, ..., dx_d) = dx_1 \cdots dx_d$. Fix $\lambda > 0$, and, for every $\mathbf{t} = (t_1, ..., t_d) \in \mathbb{R}_+^d$, define the *d-variate Ornstein-Uhlenbeck kernel*

$$f_{\mathbf{t}}(\mathbf{x}) = (2\lambda)^{\frac{d}{2}} \prod_{j=1}^{d} \exp\{-\lambda(t_i - x_i)\} \mathbf{1}_{\{x_i \leq t_i\}}, \ \ \mathbf{x} = (x_1, ..., x_d) \in \mathbb{R}^d. \tag{11.10}$$

For every fixed $q \geq 2$, we consider the qth *tensor power* of $f_{\mathbf{t}}$, denoted by $f_{\mathbf{t}}^{\otimes q}$, which is a function on $\mathbb{R}^{dq}$. Now write

$$Z_{\mathbf{t}}(1, d) = I_1(f_{\mathbf{t}}), \quad \mathbf{t} \in \mathbb{R}_+^d$$

Note that, for every $\mathbf{t} = (t_1, ..., t_d), \mathbf{s} = (s_1, ..., s_d)$, one has that $\mathbb{E}\left[Z_{\mathbf{t}}(1, d)^2\right] = \int f_t^2(\mathbf{x}) d\mathbf{x} =$ 1 and, more generally,

$$\mathbb{E}\left[Z_{\mathbf{t}}(1, d) Z_{\mathbf{s}}(1, d)\right] = \prod_{j=1}^{d} \exp\{-\lambda |t_j - s_j|\}. \tag{11.11}$$

In view of (11.11), the process $\mathbf{t} \mapsto Z_{\mathbf{t}}(d, 1)$ is called an *Ornstein-Uhlenbeck sheet* with d parameters. It follows form (5.11) that

$$Z_{\mathbf{t}}(q, d) \triangleq I_q(f_{\mathbf{t}}^{\otimes q}) = H_q(Z_{\mathbf{t}}(d, 1)),$$

where H_q is the qth Hermite polynomial. The main result of this section is the following CLT for linear functionals of $Z(q, d)$

Theorem 11.1 *Fix $\lambda > 0$ and $q \geq 2$, and define the positive constant $c = c(q, \lambda, d) := [2(q-1)!/\lambda]^d$. Then, one has that, as $T \to \infty$,*

$$M_T(q,d) = \frac{1}{\sqrt{cT^d}} \int_0^T \cdots \int_0^T Z_{\mathbf{t}}(q,d) d\mathbf{t} \xrightarrow{\text{Law}} N \sim \mathcal{N}(0,1), \tag{11.12}$$

where $\mathbf{t} = (t_1, ..., t_d)$ and $d\mathbf{t} = dt_1...dt_d$, and there exists a finite constant $\rho = \rho(\lambda, q, \nu, d) > 0$ such that, for every $T > 0$,

$$d_W(M_T(q,d),\, N) \leq \frac{\rho}{\sqrt{T^d}}. \tag{11.13}$$

Proof. By an argument similar to the one concluding the proof of Proposition 11.2, it is enough to prove the theorem in the case $d = 1$. The crucial fact is that, for each T, the random variable $M_T(1,q)$ has the form of a multiple integral, that is, $M_T(1,q) = I_q(F_T)$, where $F_T \in L_s^2(\mu^q)$ is given by

$$F_T(x_1; ...; x_q) = \frac{1}{\sqrt{cT}} \int_0^T f_t^{\otimes q}(x_1; ...; x_q) dt.$$

According to (10.3) and Proposition , both claims (11.12) and (11.13) are proved, once we show that, as $T \to \infty$, one has that

$$|1 - \mathbb{E}(M_T(q,d))^2| \sim 1/T, \tag{11.14}$$

and also that

$$\|F_T \otimes_r F_T\|_{L^2(\mu^{2q-2r})} = O(1/T), \ \forall r = 1, ..., q-1. \tag{11.15}$$

In order to prove (11.14) and (11.15), for every $t_1, t_2 \geq 0$ we introduce the notation

$$\langle f_{t_1}, f_{t_2} \rangle = \int_{\mathbb{R}} f_{t_1}(x) f_{t_2}(x)\, dx = e^{-\lambda(t_1+t_2)} e^{2\lambda(t_1 \wedge t_2)}. \tag{11.16}$$

To prove (11.14), one uses the relation (11.16) to get

$$\mathbb{E}\left[M_T(1,q)^2\right] = \frac{q!}{cT} \int_0^T \int_0^T \langle f_{t_1}, f_{t_2} \rangle^q \, dt_1 dt_2 = 1 - \frac{1}{Tq\lambda}(1 - e^{-\lambda q T}).$$

In the remaining of the proof, we will write κ in order to indicate a strictly positive finite constant independent of T, that may change from line to line. To deal with (11.15), fix $r = 1, ..., q-1$ and use the fact

$$\begin{aligned} & F_T \otimes_r F_T(w_1, ..., w_{q-r}, z_1, ..., z_{q-r}) \\ = & \frac{1}{cT} \int_0^T \int_0^T \\ & \qquad \times \left(f_{t_1}(w_1) \cdots f_{t_1}(w_{q-r}) f_{t_2}(z_1) \cdots f_{t_2}(z_{q-r})\right) \langle f_{t_1} f_{t_2} \rangle^r \ dt_1 dt_2, \end{aligned}$$

and therefore

$$\begin{aligned} & \|F_T \otimes_r F_T\|^2_{L^2(\mu^{2q-2r})} \\ = & \frac{\kappa}{T^2} \int_0^T \int_0^T \int_0^T \int_0^T \langle f_{t_1} f_{t_3} \rangle^{q-r} \langle f_{t_2} f_{t_4} \rangle^{q-r} \langle f_{t_1} f_{t_2} \rangle^r \langle f_{t_3} f_{t_4} \rangle^r \ dt_1 dt_2 dt_3 dt_4 \\ \leq & \frac{\kappa}{T}, \end{aligned}$$

where the last relation is obtained by resorting to the explicit representation (11.16), and then by evaluating the restriction of the quadruple integral to each simplex of the type $\{t_{\pi(1)} > t_{\pi(2)} > t_{\pi(3)} > t_{\pi(4)}\}$, where π is a permutation of the set $\{1, 2, 3, 4\}$. ■

12 Further readings

The content of the previous sections is mainly related to the papers [57] and [60], dealing with one- and multi-dimensional upper bounds in the Gaussian and Gamma approximations of functionals of Gaussian fields. In the following list, we shall provide a short description of some further works that have been written on related subjects.

- In [58] it is described how one can once again combine Stein's method with Malliavin calculus, in order to detect optimal rates of convergence for sequences of functionals of Gaussian fields. Given a sequence $\{F_n\}$ of such functionals and given $N \sim \mathcal{N}(0, 1)$, we say that a sequence of positive numbers $\varphi(n) \searrow 0$ provides an optimal rate of convergence, if there exists a constant $0 < c < 1$, such that

  $$c < \frac{d(F_n, N)}{\varphi(n)} \leq 1$$

  for n large enough (where d is some suitable distance between the law of F_n and the law of N). Proposition 11.1 gives an example of such a situation.

- In [64] one can find applications of (7.30) to the estimation of densities and tail probabilities associated with smooth functionals of Gaussian processes. In particular, a new formula for the density of a regular functional "with full support" is derived. Part of the computations performed in this paper are related to the theory developed by Ch. Stein in [89, Chapter VI].

- The paper [97] contains new estimates for tail bounds based on (7.30). The results are also related to polymer models.

- The paper [62] contains the proof of new second-order Poincaré inequalities, involving the operator norm of the second derivative of a given smooth random variable. This gives a generalization of a class of inequalities proved by Chatterjee in [9].

- In [70] one can find an extension of the theory developed in [57] to the case of the Gaussian approximation of functionals of Poisson random measures. This is based on an appropriate version of Malliavin calculus on the Poisson space, and is related to the work by Decreusefond and Savy [17].

- The paper [61] provides an extension of the Malliavin/Stein approach to deal with functionals of infinite Rademacher sequences. The necessary discrete Malliavin operators are discussed in [79].

References

[1] P. Baldi, G. Kerkyacharian, D. Marinucci and D. Picard (2006). High-frequency asymptotics for wavelet-based tests for Gaussianity and isotropy on the torus. To appear in: *J. Mult. Anal.*

[2] A. D. Barbour (1990). Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84**(3), 297–322.

[3] O. Barndorff-Nielsen, J. Corcuera, M. Podolskij and J. Woerner (2008). Bipower variations for Gaussian processes with stationary increments. Preprint.

[4] J.C. Breton and I. Nourdin (2008). Error bounds on the non-normal approximation of Hermite power variations of fractional Brownian motion. *Electron. Comm. Probab.* **13**, 482-493 (Electronic)

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

[6] T. Cacoullos, V. Papahanasiou and S.A. Utev (1994). Variational inequalities with examples and an application to the central limit theorem. *Ann. Probab.* **22**(3), 1607–1618.

[7] D. Chambers et E. Slud (1989). Central limit theorems for nonlinear functionals of stationary Gaussian processes. *Probab. Theory Related Fields* **80**, 323-349.

[8] S. Chatterjee (2007). A new method of normal approximation. *Ann. Probab.* **36**(4), 1584-1610

[9] S. Chatterjee (2009). Fluctuation of eigenvalues and second order Poincaré inequalities. *Probab. Theory Related Fields* **143**, 1–40.

[10] S. Chatterjee and E. Meckes (2008). Multivariate normal approximation using exchangeable pairs. *ALEA* **4**, 257-283 (Electronic)

[11] L.H.Y. Chen (1985). Poincaré-type inequalities via stochastic integrals. *Z. Wahrsch. verw. Gebiete* **69**, 251–277.

[12] L.H.Y. Chen (1987). Characterization of probability distributions by Poincaré-type inequalities. *Ann. Inst. H. Poincaré Probab. Statist.* **23**(1), 91–110.

[13] L.H.Y. Chen (1988). The central limit theorem and Poincaré-type inequalities. *Ann. Probab.* **16**(1), 300–304.

[14] L. Chen and Q.-M. Shao (2005). Stein's method for normal approximation. In: *An introduction to Stein's method*, 1–59. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. **4**, Singapore Univ. Press, Singapore, 2005.

[15] H. Chernoff (1981). A note on an inequality involving the normal distribution. *Ann. Probab.* **9**(3), 533–535.

[16] J.M. Corcuera, D. Nualart and J.H.C. Woerner (2006). Power variation of some integral long memory process. *Bernoulli* **12**(4), 713-735.

[17] L. Decreusefond and N. Savy (2007). Rubinstein distance on configuration spaces. Preprint.

[18] L. Decreusefond and A. Üstünel (1999). Stochastic analysis of the fractional Brownian motion. *Potential Anal.* **10**(2), 177–214

[19] C. Dellacherie, B. Maisonneuve and P.-A. Meyer (1992). *Probabilités et Potentiel (Chapitres XVII à XXIV).* Herman, Paris.

[20] P. Diaconis and S. Zabell (1991). Closed form summation for classical distributions: variations on a theme of de Moivre. *Stat. Science* **6**(3), 284-302

[21] P. Deheuvels, G. Peccati et M. Yor (2006) On quadratic functionals of the Brownian sheet and related processes. *Stoch. Proc. Appl.* **116**, 493-538.

[22] R.M. Dudley (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Func. Anal.* **1**, 290–330.

[23] R.M. Dudley (2003). *Real Analysis and Probability* (2^{nd} Edition). Cambridge University Press, Cambridge.

[24] D.D. Engel (1982). The multiple stochastic integral. *Memoirs of the AMS* **38**, 1–82.

[25] R. Fox et M.S. Taqqu (1987). Multiple stochastic integrals with dependent integrators. *Journal ,of Multivariate Analysis* **21**(1), 105–127.

[26] M.S. Ginovyan and A.A. Sahakyan (2007). Limit theorems for Toeplitz quadratic functionals of continuous-time stationary processes. *Probab. Theory Related Fields* **138** (3-4), 551-579

[27] L. Giraitis and D. Surgailis (1985). CLT and other limit theorems for functionals of Gaussian processes. *Z. Wahrsch. verw. Gebiete* **70**, 191–212.

[28] L. Goldstein and G. Reinert (1997). Stein's method and the zero bias trans- formation with application to simple random sampling. *Ann. Appl. Probab.* **7**(4), 935-952.

[29] F. Götze (1991). On the rate of convergence in the multivariate CLT. *Ann. Prob.* **19**(2), 724–739.

[30] M.S. Ginovyan and A.A. Sahakyan (2007). Limit theorems for Toeplitz quadratic functionals of continuous-time stationary processes. *Probab. Theory Related Fields* **138**, no. 3-4, 551-579.

[31] C. Houdré and V. Pérez-Abreu (1995). Covariance identities and inequalities for func- tionals on Wiener and Poisson spaces. *Ann. Probab.* **23**, 400–419.

[32] E.P. Hsu (2005). Characterization of Brownian motion on manifolds through integration by parts. In: *Stein's method and applications*, 195–208. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. **5**. Singapore Univ. Press, Singapore.

[33] Y. Hu and D. Nualart (2005). Renormalized self-intersection local time for fractional Brownian motion. *Ann. Probab.* **33**(3), 948-983.

[34] K. Itô (1951). Multiple Wiener integral. *J. Math. Soc. Japan* **3**, 157–169.

[35] S. Janson (1997). *Gaussian Hilbert Spaces.* Cambridge University Press.

[36] T. Jeulin (1980). *Semimartingales et Grossissement d'une Filtration*, LNM **833**. Springer-Verlag, Berlin Heidelberg New York.

[37] T. Jeulin and M. Yor (1979). Inégalité de Hardy, semi-martingales et faux-amis. Dans : *Séminaire de Probabilités XIII*, LNM **721**, Springer-Verlag, Berlin Heidelberg New York, 332–359.

[38] J.F.C. Kingman (1967). Completely random measures. *Pacific J. Math.* **21**, 59–78.

[39] S. Kwapień and W.A. Woyczyński (1992). *Random Series and Stochastic Integrals: Single and Multiple*. Birkhäuser, Basel.

[40] H.M. Luk (1994). *Stein's method for the Gamma distribution and related statistical applications*. Ph. D. Dissertation, University of Southern California, Los Angeles, USA.

[41] P. Major (1981). *Multiple Wiener-Itô integrals*. LNM **849**. Springer-Verlag, Berlin Heidelberg New York.

[42] P. Malliavin (1976). Stochastic calculus of variations and hypoelliptic operators. In: *Proc. Inter. Symp. on Stoch. Diff. Equations, Kyoto 1976*, Wiley 1978, 195–263.

[43] P. Malliavin (1997). *Stochastic Analysis*. Springer-Verlag, Berlin, Heidelberg, New York.

[44] P. Malliavin and A. Thalmaier (2004). *Stochastic calculus of variations in mathematical finance*. Springer-Verlag.

[45] M. Marcus and J. Rosen (2008). CLT for L^p moduli of continuity of Gaussian processes. *Stoch. Proc. Appl.* **118**, 1107–1135.

[46] D. Marinucci (2007). A Central Limit Theorem and Higher Order Results for the Angular Bispectrum. *Probab. Theory Related Fields* **141**, 389-409.

[47] D. Marinucci and G. Peccati (2007). High-frequency asymptotics for subordinated stationary fields on an Abelian compact group. *Stoch. Proc. Appl.* **118**(4), 585-613.

[48] D. Marinucci and G. Peccati (2007). Group representations and high-frequency central limit theorems for subordinated random fields on a sphere. Preprint.

[49] P. R. Masani (1995). The homogeneous chaos from the standpoint of vector measures, *Phil. Trans. R. Soc. Lond. A* **355**, 1099–1258

[50] H. P. McKean (1973). Geometry of differential space. *Ann. Probab.* **1**, 197–206

[51] P.-A. Meyer (1982). Note sur les processus de Ornstein-Uhlenbeck. In: *Séminaire de Probabilités XVI* , 95-132. Lecture Notes in Math. 920, Springer-Verlag, Berlin.

[52] J. Nash (1956). Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.* **80**, 931–954.

[53] I. Nourdin (2008). Integration by parts on Wiener space and applications. Lecture notes available at: http: // www.proba.jussieu.fr/ ~nourdin/ibp-wienerspace.pdf

[54] I. Nourdin, D. Nualart and C.A. Tudor (2007). Central and non-central limit theorems for weighted power variations of fractional Brownian motion. Preprint.

[55] I. Nourdin and G. Peccati (2008). Weighted power variations of iterated Brownian motion. *The Electronic Journal of Probability.* **13**, n. 43, 1229-1256 (Electronic).

[56] I. Nourdin and G. Peccati (2008). Non-central convergence of multiple integrals. *Ann. Probab.*, to appear.

[57] I. Nourdin and G. Peccati (2008). Stein's method on Wiener chaos. *Probab. Theory Rel. Fields*, to appear.

[58] I. Nourdin and G. Peccati (2008). Stein's method and exact Berry-Essén bounds for functionals of Gaussian fields. Preprint.

[59] I. Nourdin and G. Peccati (2009). *Limit theorems using Stein's method and Malliavin calculus.* In preparation.

[60] I. Nourdin, G. Peccati and A. Réveillac (2008). Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Inst. H. Poincaré Probab. Statist.*, to appear.

[61] I. Nourdin, G. Peccati and G. Reinert (2008). Stein's method and stochastic analysis of Rademacher functionals. Preprint.

[62] I. Nourdin, G. Peccati and G. Reinert (2008). Second order Poincaré inequalities and CLTs on Wiener space. *J. Func. Anal.*, to appear.

[63] I. Nourdin and A. Réveillac (2008). Asymptotic behavior of weighted quadratic variations of fractional Brownian motion: the critical case $H = 1/4$. Preprint.

[64] I. Nourdin and F.G. Viens (2008). Density estimates and concentration inequalities with Malliavin calculus. Preprint.

[65] D. Nualart (2006). *The Malliavin calculus and related topics*. Springer-Verlag, Berlin, 2nd edition.

[66] D. Nualart and S. Ortiz-Latorre (2008). Central limit theorem for multiple stochastic integrals and Malliavin calculus. *Stoch. Proc. Appl.* **118** (4), 614–628.

[67] D. Nualart and G. Peccati (2005). Central limit theorems for sequence of multiple stochastic integrals. *Ann. Probab.* **33** (1), 177–193.

[68] D. Nualart and W. Schoutens (2000). Chaotic and predictable representation for Lévy processes. *Stoch. Proc. Appl.* **90**, 109-122.

[69] G. Peccati (2007). Gaussian approximations of multiple integrals. *Electronic Communications in Probability* **12**, 350-364 (electronic).

[70] G. Peccati, J.-L. Solé, M.S. Taqqu and F. Utzet (2008). Stein's method and normal approximation of Poisson functionals. Preprint.

[71] G. Peccati and M.S. Taqqu (2007). Stable convergence of multiple Wiener-Itô integrals. *J. Theo. Probab.* **21**(3), 527-570

[72] G. Peccati and M.S. Taqqu (2008). Moments, cumulants and diagram formulae for non-linear functionals of random measures. Preprint available at: http://fr.arxiv.org/abs/0811.1726.

[73] G. Peccati and C.A. Tudor (2005). Gaussian limits for vector-valued multiple stochastic integrals. In: Séminaire de *Probabilités XXXVIII* , 247–262. Lecture Notes in Math. **1857**, Springer-Verlag, Berlin.

[74] G. Peccati and M. Yor (2004). Hardy's inequality in $L^2([0,1])$ and principal values of Brownian local times. *Asymptotic Methods in Stochastics*, AMS, Fields Institute Communications Series, 49-74.

[75] G. Peccati and M. Yor (2004). Four limit theorems for quadratic functionals of Brownian motion and Brownian bridge. *Asymptotic Methods in Stochastics*, AMS, Fields Institute Communication Series, 75-87.

[76] V. Pipiras and M.S. Taqqu (2000). Integration questions related to fractional Brownian motion. *Probab. Theory Related Fields* **118**(2), 251-291.

[77] V. Pipiras and M.S. Taqqu (2001). Are classes of deterministic integrands for fractional Brownian motion complete? *Bernoulli* **7**(6), 873-897

[78] V. Pipiras and M.S. Taqqu (2003). Fractional calculus and its connection to fractional Brownian motion. In: *Long Range Dependence*, 166-201, Birkhäuser, Basel.

[79] N. Privault (2008). Stochastic analysis of Bernoulli processes. *Probab. Surveys* **5**, 405-483 (Electronic)

[80] G. Reinert (2005). Three general approaches to Stein's method. In: *An introduction to Stein's method*, 183–221. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. **4**, Singapore Univ. Press, Singapore.

[81] G. Reinert and A. Röllin (2007). Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. Preprint available at http://arxiv.org/abs/0711.1082

[82] D. Revuz and M. Yor (1999). *Continuous martingales and Brownian motion.* Springer-Verlag, Berlin Heidelberg New York.

[83] G.-C. Rota et J. Shen (2000). On the combinatorics of cumulants. *J. Comb. Theory Series A* **91**, 283-304.

[84] G.-C. Rota et C. Wallstrom (1997). Stochastic integrals: a combinatorial approach. *Ann. Probab.* **25**(3), 1257-1283.

[85] W. Schoutens (2001). Orthogonal polynomials in Stein's method. *J. Math. Anal. Appl.* **253**(2), 515-531.

[86] M. Schreiber (1969). Fermeture en probabilité de certains sous-espaces d'un espace L^2. *Zeitschrift Warsch. verw. Gebiete* **14**, 36-48.

[87] E.V. Slud (1993). The moment problem for polynomial forms in normal random variables. *Ann. Probab.* **21**(4), 2200-2214.

[88] Ch. Stein (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory, 583–602. Univ. California Press, Berkeley, Calif..

[89] Ch. Stein (1986). Approximate computation of expectations. Institute of Mathematical Statistics Lecture Notes, Monograph Series, 7. Institute of Mathematical Statistics, Hayward, CA.

[90] A. V. Skorohod (1975). On a generalization of a stochastic integral. *Theory Probab. Appl.* **20**, 219–233.

[91] D.W. Stroock (1987). Homogeneous chaos revisited. *Séminaire de Probabilités XXI*, LNM **1247**, Springer-Verlag, Berlin Heidelberg New York, 1–8.

[92] D. Surgailis (2000). CLTs for Polynomials of Linear Sequences: Diagram Formulae with Applications. Dans : *Long Range Dependence*. Birkhäuser, Basel, pp. 111-128.

[93] D. Surgailis (2000). Non-CLT's: U-Statistics, Multinomial Formula and Approximations of Multiple Wiener-Itô integrals. Dans : *Long Range Dependence*. Birkhäuser, Basel, pp. 129-142.

[94] A.H. Pickett (2004). *Rates of convergence in χ^2 approximations via Stein's method*. Ph. D. Dissertation, Oxford University. Oxford, UK.

[95] A.N. Tikhomirov (1980). Convergence rate in the central limit theorem for weakly dependent random variables. *Theory Probab. Appl.* **25**, 790–809

[96] H. van Zanten (2000). A multivariate central limit theorem for continuous local martingales. *Stat. Probab. Lett.*, **50**, 229-235.

[97] F. Viens (2009). Stein's lemma, Malliavin calculus, and tail bounds, with application to polymer fluctuation exponent. Preprint.

[98] W. Whitt (1980). Some useful functions for functional limit theorems, *Math. Op. Res.* **5** (1), 67-85

[99] N. Wiener (1938). The homogeneous chaos. *Amer. J. Math.* **60**, 879-936.