Malliavin–Stein method: a survey of some recent developments

Ehsan Azmoodeh[a], **Giovanni Peccati**[b], **Xiaochuan Yang**[c,*]

[a]*Department of Mathematical Sciences, University of Liverpool, UK*
[b]*DMATH, Université du Luxembourg, Luxembourg*
[c]*Department of Mathematical Sciences, University of Bath, UK*

ehsan.azmoodeh@liverpool.ac.uk (E. Azmoodeh), giovanni.peccati@gmail.com (G. Peccati),
xiaochuan.j.yang@gmail.com (X. Yang)

Abstract Initiated around the year 2007, the Malliavin–Stein approach to probabilistic approximations combines Stein's method with infinite-dimensional integration by parts formulae based on the use of Malliavin-type operators. In the last decade, Malliavin–Stein techniques have allowed researchers to establish new quantitative limit theorems in a variety of domains of theoretical and applied stochastic analysis. The aim of this survey is to illustrate some of the latest developments of the Malliavin–Stein method, with specific emphasis on extensions and generalizations in the framework of Markov semigroups and of random point measures.

Keywords Limit theorems, Stein's method, Malliavin calculus, Wiener space, Poisson space, multiple integral, Markov triple, Markov generator, eigenspace, eigenfunction, spectrum, functional Γ-calculus, weak convergence, fourth moment theorems, Berry–Essen bounds, probability metrics

MSC2020 MSC 60F05, 60B10, 28C20, 60H07, 47D07, 34L10, 47A10

1 Introduction and overview

The **Malliavin–Stein method** for probabilistic approximations was initiated in the paper [64], with the aim of providing a quantitative counterpart to the (one- and multi-

*Corresponding author.

VTeX

dimensional) central limit theorems for random variables living in the Wiener chaos of a general separable Gaussian field. As formally discussed in the sections to follow, the basic idea of the approach initiated in [64] is that, in order to assess the discrepancy between some target law (Normal or Gamma, for instance), and the distribution of a nonlinear functional of a Gaussian field, one can fruitfully apply infinite-dimensional integration by parts formulae from the **Malliavin calculus of variations** [57, 66, 77, 78] to the general bounds associated with the so-called **Stein's method** for probabilistic approximations [66, 23]. In particular, the Malliavin–Stein approach captures and amplifies the essence of [21], where Stein's method was combined with finite-dimensional integration by parts formulae for Gaussian vectors, in order to deduce **second order Poincaré inequalities** – as applied to random matrix models with Gaussian-subordinated entries (see also [70, 96]).

We recall that, as initiated by P. Malliavin in the path-breaking reference [56], the Malliavin calculus is an infinite-dimensional differential calculus, whose operators act on smooth nonlinear functionals of Gaussian fields (or of more general probabilistic objects). As vividly described in the classical references [57, 77], as well as in the more recent books [66, 78], since its inception such a theory has generated a staggering number of applications, ranging, e.g., from mathematical physics to stochastic differential equations, and from mathematical finance to stochastic geometry (in particular, models involving stabilization, but also hyperplane, flat or cylinder processes), analysis on manifolds and mathematical statistics. On the other hand, the similarly successful and popular Stein's method (as created by Ch. Stein in the classical reference [92] – see also the 1986 monograph [93]) is a collection of analytical techniques, allowing one to estimate the distance between the distributions of two random objects, by using characterizing differential operators (or difference operator in the case where the random variables of interest are discrete). The discovery in [64] that the two theories can be fruitfully combined has been a major breakthrough in the domain of probabilistic limit theorems and approximations.

Since the publication of [64], the Malliavin–Stein method has generated several hundreds of papers, with ramifications in many (often unexpected) directions, including functional inequalities, random matrix theory, stochastic geometry, noncommutative probability and computer sciences. Many of hese developments largely exceed the scope of the present survey, and we invite the interested reader to consult the following general references (i)–(iii) for a more detailed presentation: (i) the webpage [1] is a constantly updated resource, listing all existing papers written around the Malliavin–Stein method; (ii) the monograph [66], written in 2012, contains a self-contained presentation of Malliavin calculus and Stein's method, as applied to functionals of general Gaussian fields, with specific emphasis on random variables belonging to a fixed Wiener chaos; (iii) the text [81] is a collection of surveys, containing an in-depth presentation of variational techniques on the Poisson space (including the Malliavin–Stein method), together with their application to asymptotic problems arising in stochastic geometry. The following more specific references (a)–(c) point to some recent developments that we find particularly exciting and ripe for further developments: (a) the papers [58, 59, 68, 82, 85, 88, 94] provide a representative overview of applications of Malliavin–Stein techniques to the study of nodal sets associated with Gaussian random fields on two-dimensional manifolds; (b) the papers [62, 74]

– and many of the reference therein – display a pervasive use of Malliavin–Stein techniques to determine rates of convergence in total variation in the Breuer–Major Theorem; (c) references [19, 61] deal with the problem of tightness and functional convergence in the Breuer–Major theorem evoked at Point (b).

The aim of the present survey is twofold. On the one hand, we aim at presenting the essence of the Malliavin–Stein method for functionals of Gaussian fields, by discussing the crucial elements of Malliavin calculus and Stein's method together with their interaction (see Section 2 and Section 3). On the other hand, we aim at introducing the reader to some of the most recent developments of the theory, with specific focus on the general theory of Markov semigroups in a diffusive setting (following the seminal references [52, 5], as well as [73, 53, 54]), and on integration by parts formulae (and associated operators) in the context of functionals of a random point measure [37, 38, 55, 49, 48, 90]. This corresponds to the content of Section 4 and Section 5, respectively. Finally, Section 6 deals with some recent results (and open problems) concerning χ^2 approximations.

From now on, every random object will be defined on a suitable common probability space $(\Omega, \mathscr{F}, \mathbb{P})$, with $\mathbb{E}$ indicating mathematical expectation with respect to $\mathbb{P}$. Throughout the paper, the symbol $\mathscr{N}(\mu, \sigma^2)$ will be a shorthand for the one-dimensional Gaussian distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$. In particular, $X \sim \mathscr{N}(\mu, \sigma^2)$ if and only if

$$\mathbb{P}[X \in A] = \int_A e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi\sigma^2}},$$

for every Borel set $A \subset \mathbb{R}$.

2 Elements of Stein's method for normal approximations

In this section, we briefly introduce the main ingredients of **Stein's method for normal approximations** in dimension one. The approximation will be performed with respect to the **total variation** and **1-Wasserstein** distances between the distributions of two random variables; more detailed information about these distances can be found in [66, Appendix C] and the references therein.

The crucial intuition behind Stein's method lies in the following heuristic reasoning: *it is a well-known fact (see, e.g., Lemma 2.1-(e) below) that a random variable X has the standard $\mathscr{N}(0, 1)$ distribution if and only if*

$$\mathbb{E}[Xf(X) - f'(X)] = 0, \tag{2.1}$$

for every smooth mapping $f : \mathbb{R} \to \mathbb{R}$; heuristically, it follows that, if X is a random variable such that the quantity $\mathbb{E}[Xf(X) - f'(X)]$ is close to zero for a large class of test functions f, then the distribution of X should be close to Gaussian.

The fact that such a heuristic argument can be made rigorous and applied in a wide array of probabilistic models was the main discovery of Stein's original contribution [92], where the foundations of Stein's method were first laid. The reader is referred to Stein's monograph [93], as well as the books [23, 66], for an exhaustive presentation

of the theory and its applications (in particular, for extensions to multidimensional approximations).

We recall that the total variation distance, between the laws of two real-valued random variables F and G, is defined by

$$d_{TV}(F, G) := \sup_{B \in \mathcal{B}(\mathbb{R})} \left| \mathbb{P}[F \in B] - \mathbb{P}[G \in B] \right|. \tag{2.2}$$

One has to note that the topology induced by the distance d_{TV} – on the set of all probability measures on $\mathbb{R}$ – is stronger than the topology of convergence in distribution; one sometimes uses the following equivalent representation of d_{TV} (see, e.g., [66, p. 213]):

$$d_{TV}(F, G)$$
$$= \frac{1}{2} \sup \left\{ \left| \mathbb{E}[h(F)] - \mathbb{E}[h(G)] \right| : h \text{ is Borel measurable and } \|h\|_\infty \le 1 \right\}. \tag{2.3}$$

The 1-Wasserstein distance d_W, between the distributions of two real-valued integrable random variables F and G, is given by

$$d_W(F, G) := \sup_{h \in \mathrm{Lip}(1)} \left| \mathbb{E}[h(F)] - \mathbb{E}[h(G)] \right|, \tag{2.4}$$

where Lip(K), $K > 0$, stands for the class of all Lipschitz mappings $h : \mathbb{R} \to \mathbb{R}$ such that h has a Lipschitz constant $\le K$. As for total variation, the topology induced by d_W – on the set of all probability measures on $\mathbb{R}$ having a finite absolute first moment – is stronger than the topology of convergence in distribution; it is also interesting to recall the dual representation

$$d_W(F, G) = \inf \mathbb{E} \left| X - Y \right|, \tag{2.5}$$

where the infimum is taken over all couplings (X, Y) of F and G; see, e.g., [97, p. 95] for a discussion of this fact.

The following classical result, whose complete proof can be found, e.g., in [66, p. 64 and p. 67], contains all the elements of Stein's method that are needed for our discussion; as for many fundamental findings in the area, this result can be traced back to [92].

Lemma 2.1. *Let $N \sim \mathcal{N}(0, 1)$ be a standard Gaussian random variable.*

(a) *Fix $h : \mathbb{R} \to [0, 1]$, a Borel-measurable function. Define $f_h : \mathbb{R} \to \mathbb{R}$ as*

$$f_h(x) := e^{\frac{x^2}{2}} \int_{-\infty}^{x} \{h(y) - \mathbb{E}[h(N)]\} e^{-\frac{y^2}{2}} dy, \quad x \in \mathbb{R}. \tag{2.6}$$

Then, f_h is continuous on $\mathbb{R}$ with $\|f_h\|_\infty \le \sqrt{\frac{\pi}{2}}$ and $f_h \in \mathrm{Lip}(2)$. Moreover, there exists a version of f_h' verifying

$$f_h'(x) - x f_h(x) = h(x) - \mathbb{E}[h(N)], \quad \text{for all } x \in \mathbb{R}. \tag{2.7}$$

(b) *Consider $h : \mathbb{R} \to \mathbb{R} \in \text{Lip}(1)$, and define $f_h : \mathbb{R} \to \mathbb{R}$ as in (2.6). Then, f_h is of class C^1 on $\mathbb{R}$, with $\|f_h'\|_\infty \leq 1$ and $f_h' \in \text{Lip}(2)$, and f_h solves (2.7).*

(c) *Let X be an integrable random variable. Then*

$$d_{TV}(X, N) \leq \sup_f \left| \mathbb{E}\big[f(X)X - f'(X) \big] \right|$$

where the supremum is taken over all pairs (f, f') such that f is a Lipschitz function whose absolute value is bounded by $\sqrt{\frac{\pi}{2}}$, and f' is a version of the derivative of f satisfying $\|f'\|_\infty \leq 2$.

(d) *Let X be an integrable random variable. Then,*

$$d_W(X, N) \leq \sup_f \left| \mathbb{E}\big[f(X)X - f'(X) \big] \right|$$

where the supremum is taken over all C^1 functions $f : \mathbb{R} \to \mathbb{R}$ such that $\|f'\|_\infty \leq 2$ and $f' \in \text{Lip}(2)$.

(e) *Let X be a general random variable. Then $X \sim \mathcal{N}(0, 1)$ if and only if $\mathbb{E}[f'(X) - Xf(X)] = 0$ for every absolutely continuous function f such that $\mathbb{E}|f'(N)| < +\infty$.*

Sketch of the proof. Points (a) and (b) can be verified by a direct computation. Point (c) and Point (d) follow by plugging the left-hand side of (2.7) into (2.3) and (2.4), respectively. Finally, the fact that the relation $\mathbb{E}[f'(X) - Xf(X)] = 0$ implies that $X \sim \mathcal{N}(0, 1)$ is a direct consequence of Point (c), whereas the reverse implication follows by an integration by parts argument. □

3 Normal approximation with Stein's method and Malliavin calculus

The first part of the present section contains some elements of Gaussian analysis and Malliavin calculus. The reader can consult, for instance, the references [66, 77, 57, 78] for further details. In Section 3.2 we will shortly explore the connection between Malliavin calculus and the version of Stein's method presented in Section 2.

3.1 Isonormal processes, multiple integrals, and the Malliavin operators

Let $\mathfrak{H}$ be a real separable Hilbert space. For any $q \geq 1$, we write $\mathfrak{H}^{\otimes q}$ and $\mathfrak{H}^{\odot q}$ to indicate, respectively, the qth **tensor power** and the qth **symmetric tensor power** of $\mathfrak{H}$; we also set by convention $\mathfrak{H}^{\otimes 0} = \mathfrak{H}^{\odot 0} = \mathbb{R}$. When $\mathfrak{H} = L^2(A, \mathcal{A}, \mu) =: L^2(\mu)$, where μ is a σ-finite and nonatomic measure on the measurable space $(A, \mathcal{A})$, then $\mathfrak{H}^{\otimes q} \simeq L^2(A^q, \mathcal{A}^q, \mu^q) =: L^2(\mu^q)$, and $\mathfrak{H}^{\odot q} \simeq L_s^2(A^q, \mathcal{A}^q, \mu^q) := L_s^2(\mu^q)$, where $L_s^2(\mu^q)$ stands for the subspace of $L^2(\mu^q)$ composed of those functions that are μ^q-almost everywhere symmetric. We denote by $W = \{W(h) : h \in \mathfrak{H}\}$ an **isonormal Gaussian process** over $\mathfrak{H}$. This means that W is a centered Gaussian family with a covariance structure given by the relation $\mathbb{E}[W(h)W(g)] = \langle h, g \rangle_{\mathfrak{H}}$. Without loss of

generality, we can also assume that $\mathscr{F} = \sigma(W)$, that is, $\mathscr{F}$ is generated by W, and use the shorthand notation $L^2(\Omega) := L^2(\Omega, \mathscr{F}, \mathbb{P})$.

For every $q \geq 1$, the symbol C_q stands for the qth **Wiener chaos** of W, defined as the closed linear subspace of $L^2(\Omega)$ generated by the family $\{H_q(W(h)) : h \in \mathfrak{H}, \|h\|_{\mathfrak{H}} = 1\}$, where H_q is the qth **Hermite polynomial**, defined as follows:

$$H_q(x) = (-1)^q e^{\frac{x^2}{2}} \frac{d^q}{dx^q}\left(e^{-\frac{x^2}{2}}\right). \tag{3.1}$$

We write by convention $C_0 = \mathbb{R}$. For any $q \geq 1$, the mapping $I_q(h^{\otimes q}) = H_q(W(h))$ can be extended to a linear isometry between the symmetric tensor product $\mathfrak{H}^{\odot q}$ (equipped with the modified norm $\sqrt{q!}\, \|\cdot\|_{\mathfrak{H}^{\otimes q}}$) and the qth Wiener chaos C_q. For $q = 0$, we write by convention $I_0(c) = c$, $c \in \mathbb{R}$.

It is well known that $L^2(\Omega)$ can be decomposed into the infinite orthogonal sum of the spaces C_q: this means that any square-integrable random variable $F \in L^2(\Omega)$ admits the following **Wiener–Itô chaotic expansion**

$$F = \sum_{q=0}^{\infty} I_q(f_q), \tag{3.2}$$

where the series converges in $L^2(\Omega)$, $f_0 = E[F]$, and the kernels $f_q \in \mathfrak{H}^{\odot q}$, $q \geq 1$, are uniquely determined by F. For every $q \geq 0$, we denote by J_q the orthogonal projection operator on the qth Wiener chaos. In particular, if $F \in L^2(\Omega)$ has the form (3.2), then $J_q F = I_q(f_q)$ for every $q \geq 0$.

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

$$f \otimes_r g = \sum_{i_1,\ldots,i_r=1}^{\infty} \langle f, e_{i_1} \otimes \cdots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}} \otimes \langle g, e_{i_1} \otimes \cdots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}}. \tag{3.3}$$

Notice that the definition of $f \otimes_r g$ does not depend on the particular choice of $\{e_k, k \geq 1\}$, and that $f \otimes_r g$ is not necessarily symmetric; we denote its symmetrization by $f \widetilde{\otimes}_r g \in \mathfrak{H}^{\odot(p+q-2r)}$. Moreover, $f \otimes_0 g = f \otimes g$ equals the tensor product of f and g while, for $p = q$, $f \otimes_q g = \langle f, g \rangle_{\mathfrak{H}^{\otimes q}}$. When $\mathfrak{H} = L^2(A, \mathcal{A}, \mu)$ and $r = 1, \ldots, p \wedge q$, the contraction $f \otimes_r g$ is the element of $L^2(\mu^{p+q-2r})$ given by

$$f \otimes_r g(x_1, \ldots, x_{p+q-2r})$$
$$= \int_{A^r} f(x_1, \ldots, x_{p-r}, a_1, \ldots, a_r) \times$$
$$\times g(x_{p-r+1}, \ldots, x_{p+q-2r}, a_1, \ldots, a_r) d\mu(a_1)\ldots d\mu(a_r). \tag{3.4}$$

It is a standard fact of Gaussian analysis that the following **multiplication formula** holds: if $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$, then

$$I_p(f)I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r}\binom{q}{r} I_{p+q-2r}(f \widetilde{\otimes}_r g). \tag{3.5}$$

We now introduce some basic elements of the Malliavin calculus with respect to the isonormal Gaussian process W.

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

$$F = g\left(W(\varphi_1), \ldots, W(\varphi_n)\right), \tag{3.6}$$

where $n \geq 1$, $g : \mathbb{R}^n \to \mathbb{R}$ is an infinitely differentiable function such that its partial derivatives have polynomial growth, and $\varphi_i \in \mathfrak{H}, i = 1, \ldots, n$. The **Malliavin derivative** of F with respect to W is the element of $L^2(\Omega, \mathfrak{H})$ defined as

$$DF = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}\left(W(\varphi_1), \ldots, W(\varphi_n)\right) \varphi_i.$$

In particular, $DW(h) = h$ for every $h \in \mathfrak{H}$. By iteration, one can define the m**th derivative** $D^m F$, which is an element of $L^2(\Omega, \mathfrak{H}^{\odot m})$, for every $m \geq 2$. For $m \geq 1$ and $p \geq 1$, $\mathbb{D}^{m,p}$ denotes the closure of $\mathcal{S}$ with respect to the norm $\|\cdot\|_{m,p}$, defined by the relation

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

We often use the (canonical) notation $\mathbb{D}^\infty := \bigcap_{m \geq 1} \bigcap_{p \geq 1} \mathbb{D}^{m,p}$. For example, it is a well-known fact that any random variable F that is a finite linear combination of multiple Wiener–Itô integrals is an element of $\mathbb{D}^\infty$. The Malliavin derivative D obeys the following **chain rule**. If $\phi : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable with bounded partial derivatives and if $F = (F_1, \ldots, F_n)$ is a vector of elements of $\mathbb{D}^{1,2}$, then $\phi(F) \in \mathbb{D}^{1,2}$ and

$$D\phi(F) = \sum_{i=1}^{n} \frac{\partial \phi}{\partial x_i}(F) DF_i. \tag{3.7}$$

Note also that a random variable F as in (3.2) is in $\mathbb{D}^{1,2}$ if and only if $\sum_{q=1}^{\infty} q \|J_q F\|_{L^2(\Omega)}^2 < \infty$ and in this case one has the following explicit relation:

$$\mathbb{E}\left[\|DF\|_{\mathfrak{H}}^2\right] = \sum_{q=1}^{\infty} q \|J_q F\|_{L^2(\Omega)}^2.$$

If $\mathfrak{H} = L^2(A, \mathcal{A}, \mu)$ (with μ nonatomic), then the derivative of a random variable F as in (3.2) can be identified with the element of $L^2(A \times \Omega)$ given by

$$D_t F = \sum_{q=1}^{\infty} q I_{q-1}\left(f_q(\cdot, t)\right), \quad t \in A. \tag{3.8}$$

The operator **L**, defined as $\mathbf{L} = \sum_{q=0}^{\infty} -q J_q$, is the **infinitesimal generator of the Ornstein–Uhlenbeck semigroup**. The domain of **L** is

$$\mathrm{Dom}\mathbf{L} = \{F \in L^2(\Omega) : \sum_{q=1}^{\infty} q^2 \|J_q F\|_{L^2(\Omega)}^2 < \infty\} = \mathbb{D}^{2,2}.$$

For any $F \in L^2(\Omega)$, we define $\mathbf{L}^{-1}F = \sum_{q=1}^{\infty} -\frac{1}{q}J_q(F)$. The operator $\mathbf{L}^{-1}$ is called the *pseudoinverse* of $\mathbf{L}$. Indeed, for any $F \in L^2(\Omega)$, we have that $\mathbf{L}^{-1}F \in \mathrm{Dom}\mathbf{L} = \mathbb{D}^{2,2}$, and

$$\mathbf{L}\mathbf{L}^{-1}F = F - \mathbb{E}(F). \tag{3.9}$$

The following infinite dimensional Malliavin integration by parts formula plays a crucial role in the analysis (see, for instance, [66, Section 2.9] for a proof).

Lemma 3.1. *Suppose that $F \in \mathbb{D}^{1,2}$ and $G \in L^2(\Omega)$. Then, $\mathbf{L}^{-1}G \in \mathbb{D}^{2,2}$ and*

$$\mathbb{E}[FG] = \mathbb{E}[F]\mathbb{E}[G] + \mathbb{E}[\langle DF, -D\mathbf{L}^{-1}G\rangle_{\mathfrak{H}}]. \tag{3.10}$$

Inspired by the Malliavin integration by parts formula appearing in Lemma 3.1, we now introduce a class of **iterated Gamma operators**. We will need such operators in Section 6.

Definition 3.2 (See Chapter 8 in [66]). Let $F \in \mathbb{D}^{\infty}$; the sequence of random variables $\{\Gamma_i(F)\}_{i \geq 0} \subset \mathbb{D}^{\infty}$ is recursively defined as follows. Set $\Gamma_0(F) = F$ and, for every $i \geq 1$,

$$\Gamma_i(F) = \langle DF, -D\mathbf{L}^{-1}\Gamma_{i-1}(F)\rangle_{\mathfrak{H}}.$$

Definition 3.3 (Cumulants). Let F be a real-valued random variable such that $\mathbb{E}|F|^m < \infty$ for some integer $m \geq 1$, and write $\varphi_F(t) = \mathbb{E}[e^{itF}]$, $t \in \mathbb{R}$, for the characteristic function of F. Then, for $r = 1, \ldots, m$, the rth **cumulant** of F, denoted by $\kappa_r(F)$, is given by

$$\kappa_r(F) = (-i)^r \frac{d^r}{dt^r} \log \varphi_F(t)|_{t=0}. \tag{3.11}$$

Remark 3.4. When $\mathbb{E}(F) = 0$, then the first four cumulants of F are the following: $\kappa_1(F) = \mathbb{E}[F] = 0$, $\kappa_2(F) = \mathbb{E}[F^2] = \mathrm{Var}(F)$, $\kappa_3(F) = \mathbb{E}[F^3]$, and

$$\kappa_4(F) = \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2.$$

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

Proposition 3.5 (See Chapter 8 in [66]). *Let $F \in \mathbb{D}^{\infty}$. Then $\kappa_r(F) = (r-1)!\mathbb{E}[\Gamma_{r-1}(F)]$ for every $r \geq 1$.*

As announced, in the next subsection we show how to use the above Malliavin machinery in order to study the Stein's bounds presented in Section 2.

3.2 Connection with Stein's method

Let $F \in \mathbb{D}^{1,2}$ with $\mathbb{E}[F] = 0$ and $\mathbb{E}[F^2] = 1$. Take a C^1 function such that $\|f\| \leq \sqrt{\frac{\pi}{2}}$ and $\|f'\| \leq 2$. Using the Malliavin integration by parts formula stated in Lemma 3.1 together with the chain rule (3.7), we can write

$$\left|\mathbb{E}[f'(F) - Ff(F)]\right| = \left|\mathbb{E}[f'(F)\left(1 - \langle DF, -D\mathbf{L}^{-1}F\rangle_{\mathfrak{H}}\right)]\right| \\ \leq 2\mathbb{E}\left|1 - \langle DF, -D\mathbf{L}^{-1}F\rangle_{\mathfrak{H}}\right|. \tag{3.12}$$

If we furthermore assume that $F \in \mathbb{D}^{1,4}$, then the random variable $1 - \langle DF, -D\mathbf{L}^{-1}F \rangle_{\mathfrak{H}}$ is square-integrable, using the Cauchy–Schwarz inequality we infer that

$$\left| \mathbb{E}[f'(F) - Ff(F)] \right| \leq 2\sqrt{\mathrm{Var}\left(\langle DF, -D\mathbf{L}^{-1}F \rangle_{\mathfrak{H}} \right)}.$$

Note that in above we used the fact that $\mathbb{E}[\langle DF, -D\mathbf{L}^{-1}F \rangle_{\mathfrak{H}}] = \mathbb{E}[F^2] = 1$. The above arguments combined with Lemma 2.1 yield immediately[2] the next crucial statement, originally proved in [64].

Theorem 3.6. *Let $F \in \mathbb{D}^{1,2}$ be a generic random element with $\mathbb{E}[F] = 0$ and $\mathbb{E}[F^2] = 1$. Let $N \sim \mathcal{N}(0, 1)$. Assume further that F has a density with respect to the Lebesgue measure. Then,*

$$d_{TV}(F, N) \leq 2\,\mathbb{E}\left| 1 - \langle DF, -D\mathbf{L}^{-1}F \rangle_{\mathfrak{H}} \right|.$$

Moreover, assume that $F \in \mathbb{D}^{1,4}$, then

$$d_{TV}(F, N) \leq 2\sqrt{\mathrm{Var}\left(\langle DF, -D\mathbf{L}^{-1}F \rangle_{\mathfrak{H}} \right)}.$$

In particular case, if $F = I_q(f)$ belongs to the Wiener chaos of order $q \geq 2$, then

$$d_{TV}(F, N) \leq 2\sqrt{\frac{q-1}{3q}\left(\mathbb{E}[F^4] - 3 \right)}. \tag{3.13}$$

Note that, by virtue of Lemma 2.1, similar bounds can be immediately obtained for the Wasserstein distance d_W (and many more – see [66, Chapter 5]). In particular, the previous statement allows one to recover the following central limit theorem for chaotic random variables, first proved in [80].

Corollary 3.7 (Fourth Moment Theorem). *Let $\{F_n\}_{n \geq 1} = \{I_q(f_n)\}_{n \geq 1}$ be a sequence of random elements in a fixed Wiener chaos of order $q \geq 2$ such that $\mathbb{E}[F_n^2] = q!\| f_n \|^2 = 1$. Assume that $N \sim \mathcal{N}(0, 1)$. Then, as n tends to infinity, the following assertions are equivalent.*

(I) $F_n \longrightarrow N$ *in distribution.*

(II) $\mathbb{E}[F_n^4] \longrightarrow 3 \,(= \mathbb{E}[N^4])$.

As demonstrated by the webpage [1], the 'fourth moment theorem' stated in Corollary 3.7 has been the starting point of a very active line of research, composed of several hundred papers connected with disparate applications. In the next section, we will implicitly provide a general version of Theorem 3.6 (with the 1-Wasserstein distance replacing the total variation distance), whose proof relies only on the spectral properties of the Ornstein–Uhlenbeck generator $\mathbf{L}$ and on the so-called Γ calculus (see, e.g., [18]).

---

[2]This is not completely accurate: attention has indeed to be paid to the fact that the function f_h in (2.7) is only almost everywhere differentiable, and F does not necessarily have a density – see [60, Theorem 5.2] for a detailed proof based on the Lusin theorem.

4   The Markov triple approach

In this section, we introduce a general framework for studying and generalizing the fourth moment phenomenon appearing in the statement of Corollary 3.7. The forthcoming approach was first introduced in [52] by M. Ledoux, and then further developed and generalised in [5, 8].

4.1   Diffusive fourth moment structures

We start with definition of our general setup.

Definition 4.1. A **diffusive fourth moment structure** is a triple $(E, \mu, \mathbf{L})$ such that:

(a) (E, μ) is a probability space;

(b) $\mathbf{L}$ is a symmetric unbounded operator defined on some dense subset of $L^2(E, \mu)$, that we denote by $\mathcal{D}(\mathbf{L})$ (the set $\mathcal{D}(\mathbf{L})$ is called the **domain** of $\mathbf{L}$);

(c) the associated **carré-du-champ operator** Γ is a symmetric bilinear operator, and is defined by

$$2\Gamma[X, Y] := \mathbf{L}[XY] - X\mathbf{L}[Y] - Y\mathbf{L}[X]; \qquad (4.1)$$

(d) the operator $\mathbf{L}$ is **diffusive**, meaning that, for any $\mathcal{C}_b^2$ function $\varphi \colon \mathbb{R} \to \mathbb{R}$, any $X \in \mathcal{D}(\mathbf{L})$, it holds that $\varphi(X) \in \mathcal{D}(\mathbf{L})$ and

$$\mathbf{L}[\varphi(X)] = \varphi'(X)\mathbf{L}[X] + \varphi''(X)\Gamma[X, X]; \qquad (4.2)$$

Note that $\mathbf{L}[1] = 0$ (by taking $\varphi = 1 \in \mathcal{C}_b^2$). The latter property is equivalent to say that the operator Γ satisfies the chain rule:

$$\Gamma[\varphi(X), X] = \varphi'(X)\Gamma[X, X];$$

(e) the operator $-\mathbf{L}$ diagonalizes the space $L^2(E, \mu)$ with $\mathbf{sp}(-\mathbf{L}) = \mathbb{N}$, meaning that

$$L^2(E, \mu) = \bigoplus_{i=0}^{\infty} \mathbf{Ker}(\mathbf{L} + i\mathbf{Id});$$

(f) for any pair of eigenfunctions (X, Y) of the operator $-\mathbf{L}$ associated with the eigenvalues (p_1, p_2),

$$XY \in \bigoplus_{i \leq p_1+p_2} \mathbf{Ker}(\mathbf{L} + i\mathbf{Id}). \qquad (4.3)$$

In this context, we usually write $\Gamma[X]$ instead of $\Gamma[X, X]$ and $\mathbb{E}$ denotes the integration against probability measure μ.

Remark 4.2.   (1) Property (d) together with symmetric property of the operator $\mathbf{L}$ determine a functional calculus through the following fundamental **integration by parts formula**: for any X, Y in $\mathcal{D}(\mathbf{L})$ and $\varphi \in \mathcal{C}_b^2$,

$$\mathbb{E}[\varphi'(X)\Gamma[X, Y]] = -\mathbb{E}[\varphi(X)\mathbf{L}[Y]] = -\mathbb{E}[Y\mathbf{L}[\varphi(X)]]. \qquad (4.4)$$

(2) The results in this section can be stated under the weaker assumption that $\mathbf{sp}(-\mathbf{L}) = \{0 = \lambda_0 < \lambda_1, \ldots, \lambda_k < \cdots\} \subset \mathbb{R}_+$ is discrete. However, to keep a transparent presentation, we restrict ourselves to the assumption $\mathbf{sp}(-\mathbf{L}) = \mathbb{N}$. The reader is referred to [5] for further details.

(3) We point out that, by a recursive argument, assumption (4.3) yields that for any $X \in \mathbf{Ker}(\mathbf{L} + p\mathbf{Id})$ and any polynomial P of degree m, we have

$$P(X) \in \bigoplus_{i \leq mp} \mathbf{Ker}\,(\mathbf{L} + i\mathbf{Id})\,. \tag{4.5}$$

(4) The eigenspaces of a diffusive fourth moment structure are **hypercontractive** (see [10] for details and sufficient conditions), that is, there exists a constant $C(M, k)$ such that for any $X \in \bigoplus_{i \leq M} \mathbf{Ker}\,(\mathbf{L} + i\mathbf{Id})$:

$$\mathbb{E}(X^{2k}) \leq C(M, k)\, \mathbb{E}(X^2)^k. \tag{4.6}$$

(5) Property (f) in the previous definition roughly implies that eigenfunctions of $\mathbf{L}$ in a diffusive fourth moment structure behave like orthogonal polynomials with respect to multiplication.

For further details on our setup, we refer the reader to [18] as well as [5, 8]. The next example describes some diffusive fourth moment structures. The reader can consult [8, Section 2.2] for two classical methods for building further diffusive fourth moment structures starting from known ones.

Example 4.3.   (a) **Finite-Dimensional Gaussian Structures:** Let $d \geq 1$ and denote by γ_d the d-dimensional standard Gaussian measure on $\mathbb{R}^d$. It is well known (see, for example, [18]), that γ_d is the invariant measure of the Ornstein–Uhlenbeck generator, defined for any test function φ by

$$\mathbf{L}\varphi(x) = \Delta\varphi - \sum_{i=1}^{d} x_i \partial_i \varphi(x). \tag{4.7}$$

Its spectrum is given by $-\mathbb{N}_0$ and the eigenspaces are of the form

$$\mathbf{Ker}(\mathbf{L} + k\mathbf{Id}) = \left\{ \sum_{i_1+i_2+\cdots+i_d=k} \alpha(i_1, \ldots, i_d) \prod_{j=1}^{d} H_{i_j}(x_j) \right\},$$

where H_n denotes the Hermite polynomial of order n. Since, eigenfunctions of $\mathbf{L}$ are multivariate polynomials so it is straightforward to see that assumption (f) is also verified.

(b) **Wiener space and isonormal processes:** Letting $d \to \infty$ in the setup of the previous item (a) one recovers the infinite dimensional generator of the Ornstein–Uhlenbeck semigroup for isonormal processes, as defined in Section 3.1. It is easily verified in particular, by using (3.5), that $(\Omega, \mathscr{F}, \mathbf{L})$ is also a diffusive fourth moment structure.

(c) **Laguerre Structure:** Let $\nu \geq -1$, and $\pi_{1,\nu}(dx) = x^{\nu-1}\frac{e^{-x}}{\Gamma(\nu)}\mathbf{1}_{(0,\infty)}dx$ be the Gamma distribution with parameter ν on $\mathbb{R}_+$. The associated Laguerre generator is defined for any test function φ (in dimension one) by

$$\mathbf{L}_{1,\nu}(\varphi) = x\varphi''(x) + (\nu + 1 - x)\varphi'(x). \qquad (4.8)$$

By a classical tensorization procedure, we obtain the Laguerre generator in dimension d associated with the measure

$$\pi_{d,\nu}(dx) = \pi_{1,\nu}(dx_1)\pi_{1,\nu}(dx_2)\cdots\pi_{1,\nu}(dx_d),$$

where $x = (x_1, x_2, \ldots, x_d)$:

$$\mathbf{L}_{d,\nu}(\varphi) = \sum_{i=1}^{d}\left(x_i\partial_{i,i}\varphi + (\nu + 1 - x_i)\partial_i\varphi\right). \qquad (4.9)$$

It is also classical that (see, for example, [18]) the spectrum of $\mathbf{L}_{d,\nu}$ is given by $-\mathbb{N}_0$ and moreover that

$$\mathbf{Ker}(\mathbf{L}_{d,p} + k\mathbf{Id}) = \left\{\sum_{i_1+i_2+\cdots+i_d=k}\alpha(i_1,\ldots,i_d)\prod_{j=1}^{d}L_{i_j}^{(\nu)}(x_j)\right\}, \qquad (4.10)$$

where $L_n^{(\nu)}$ stands for the Laguerre polynomial of order n with parameter ν which is defined by

$$L_n^{(\nu)}(x) = \frac{x^{-\nu}e^x}{n!}\frac{d^n}{dx^n}\left(e^{-x}x^{n+\nu}\right).$$

In the next subsection, we demonstrate how a diffusive fourth moment structure can be combined with the tools of Γ calculus, in order to deduce substantial generalizations of Theorem 3.6.

4.2 Connection with Γ calculus

Throughout this section, we assume that $(E, \mu, \mathbf{L})$ is a diffusive fourth moment structure. Our principal aim is to prove a fourth moment criterion analogous to that of (3.13) for eigenfunctions of the operator $\mathbf{L}$. To do this, we assume that $X \in \mathbf{Ker}(\mathbf{L} + q\mathbf{Id})$ for some $q \geq 1$ with $\mathbb{E}[X^2] = 1$. The arguments implemented in the proof will clearly demonstrate that requirements (d) and (f) in Definition 4.1 are the most crucial elements in order to establish our estimates.

Proposition 4.4. *Let $q \geq 1$. Assume that $X \in \mathbf{Ker}(\mathbf{L} + q\mathbf{Id})$ with $\mathbb{E}[X^2] = 1$. Then,*

$$\mathrm{Var}\left(\Gamma[X]\right) \leq \frac{q^2}{3}\left\{\mathbb{E}[X^4] - 3\right\}.$$

Proof. First note that by using integration by parts formula (4.4), we have $\mathbb{E}[\Gamma[X]] = -\mathbb{E}[XLX] = q\mathbb{E}[X^2] = q$. Secondly, by using the definition of the carré-du-champ operator Γ and the fact that $\mathbf{L}X = -qX$, one easily verifies that

$$\Gamma[X] - q = \frac{1}{2} (\mathbf{L} + 2q\mathbf{Id}) (X^2 - 1).$$

Next, taking into account properties (f) and (g) we can conclude that

$$X^2 - 1 \in \bigoplus_{1 \le i \le 2q} \mathbf{Ker}\,(\mathbf{L} + i\mathbf{Id}).$$

For the rest of the proof, we use the notation J_i to denote the projection of a square-integrable element X onto the eigenspace $\mathbf{Ker}\,(\mathbf{L} + i\mathbf{Id})$. Now,

$$\mathrm{Var}\,(\Gamma[X])$$

$$= \mathbb{E}\left[(\Gamma[X] - q)^2\right] = \frac{1}{4} \mathbb{E}\left[(\mathbf{L} + 2q\mathbf{Id}) (X^2 - 1) \times (\mathbf{L} + 2q\mathbf{Id}) (X^2 - 1)\right]$$

$$= \frac{1}{4} \mathbb{E}\left[\mathbf{L}(X^2 - 1) (\mathbf{L} + 2q\mathbf{Id}) (X^2 - 1)\right] + \frac{q}{2} \mathbb{E}\left[(X^2 - 1) (\mathbf{L} + 2q\mathbf{Id}) (X^2 - 1)\right]$$

$$= \frac{1}{4} \sum_{1 \le i \le 2q} (-i)(2q - i) \mathbb{E}\left[\left(J_i (X^2 - 1)\right)^2\right] + \frac{q}{2} \mathbb{E}\left[(X^2 - 1) (\mathbf{L} + 2q\mathbf{Id}) (X^2 - 1)\right]$$

$$\le \frac{q}{2} \mathbb{E}\left[(X^2 - 1) (\mathbf{L} + 2q\mathbf{Id}) (X^2 - 1)\right]$$

$$= q\mathbb{E}\left[(X^2 - 1)(\Gamma[X] - q)\right] = q\mathbb{E}\left[(X^2 - 1)\Gamma[X]\right]$$

$$= q\mathbb{E}\left[\Gamma[\frac{X^3}{3} - X, X]\right] = -q\mathbb{E}\left[\left(\frac{X^3}{3} - X\right) \mathbf{L}X\right]$$

$$= q^2 \mathbb{E}\left[X \left(\frac{X^3}{3} - X\right)\right] = q^2 \mathbb{E}\left[\frac{X^4}{3} - X^2\right]$$

$$= \frac{q^2}{3} \left\{\mathbb{E}[X^4] - 3\right\},$$

thus yielding the desired conclusion. $\qquad\square$

In order to avoid some technicalities, we now present a quantitative bound in the 1-Wasserstein distance d_W (and not in the more challenging total variation distance d_{TV}) for eigenfunctions of the operator $\mathbf{L}$. This requires to adapt the Stein's method machinery presented in Section 2 to our setting, as a direct application of the integration by part formula (4.4). The arguments below are borrowed in particular from [52, Proposition 1].

Proposition 4.5. *Let $(E, \mu, \mathbf{L})$ be a diffusive fourth moment structure. Assume that $X \in \mathbf{Ker}(\mathbf{L} + q\mathbf{Id})$ for some $q \ge 1$ with $\mathbb{E}[X^2] = 1$. Let $N \sim \mathcal{N}(0, 1)$. Then,*

$$d_W(X, N) \le \frac{2}{q} \mathrm{Var}\,(\Gamma[X])^{\frac{1}{2}}.$$

Proof. For every function f of class C^1 on $\mathbb{R}$, with $\|f'\|_\infty \leq 1$ and $f' \in \text{Lip}(2)$ according to Part (b) in Lemma 2.1, it is enough to show that

$$\left| \mathbb{E}\left[f'(X) - Xf(X) \right] \right| \leq \frac{2}{q} \, \text{Var}\,(\Gamma[X])^{\frac{1}{2}}.$$

Since $\mathbf{L}X = -qX$, and diffusivity of the operator Γ together with integration by parts formula (4.4), one can write that

$$\mathbb{E}\left[f'(X) - Xf(X) \right] = \mathbb{E}\left[f'(X) + \frac{1}{q}\mathbf{L}(X)f(X) \right] = \mathbb{E}\left[f'(X) - \frac{1}{q}\Gamma[f(X), X] \right]$$

$$= \mathbb{E}\left[f'(X) - \frac{1}{q}f'(X)\Gamma[X] \right]$$

$$= \frac{1}{q}\mathbb{E}\left[f'(X)\,(q - \Gamma[X]) \right].$$

Now, the claim follows at once by using the Cauchy–Schwarz inequality and noting that $\mathbb{E}[\Gamma[X]] = q\,\mathbb{E}[X^2] = q$. $\qquad\qquad\square$

We end this section with the following general version of the fourth moment theorem for eigenfunctions of the operator $\mathbf{L}$, obtained by combining Propositions 4.4 and 4.5.

Theorem 4.6. *Let $(E, \mu, \mathbf{L})$ be a diffusive fourth moment structure. Assume that $X \in \mathbf{Ker}(\mathbf{L} + q\mathbf{Id})$ for some $q \geq 1$ with $\mathbb{E}[X^2] = 1$. Let $N \sim \mathcal{N}(0, 1)$. Then,*

$$d_W(X, N) \leq \frac{2}{\sqrt{3}}\sqrt{\mathbb{E}[X^4] - 3}.$$

It follows that, if $\{X_n\}_{n \geq 1}$ is a sequence of eigenfunctions in a fixed eigenspace $\mathbf{Ker}(\mathbf{L} + q\mathbf{Id})$ where $q \geq 1$ and $\mathbb{E}[X_n^2] = 1$ for all $n \geq 1$, then the following implication holds: $\mathbb{E}[X_n^4] \to 3$ if and only if X_n converges in distribution towards the standard Gaussian random variable N.

Remark 4.7. The fact that the condition $\mathbb{E}[X_n^4] \to 3$ is necessary for convergence to the Gaussian random variable is a direct consequence of the hypercontractive estimate (4.6).

4.3   Transport distances, Stein discrepancy and Γ calculus

The general setting of the Markov triple together with Γ calculus provide a suitable framework to study **functional inequalities** such as the classical **logarithmic Sobolev inequality** or the celebrated **Talagrand quadratic transportation cost inequality**. For simplicity, here we restrict ourselves to the setting of Wiener structure and the Gaussian measure to be our reference measure. The reader may consult references [53, 54] for a presentation of the general setting, and [72, 73] for some previous references connecting fourth moment theorems and entropic estimates.

Let $d \geq 1$, and $d\gamma(x) = (2\pi)^{-\frac{d}{2}}e^{-\frac{|x|}{2}}dx$ be the standard Gaussian measure on $\mathbb{R}^d$. Assume that $d\nu = h\,d\gamma$ is a probability measure on $\mathbb{R}^d$ with a (smooth) density function $h : \mathbb{R}^d \to \mathbb{R}_+$ with respect to the Gaussian measure γ. Inspired

from Gaussian integration by parts formula we introduce first the crucial notion of a **Stein kernel** τ_ν associated with the probability measure ν and, then, the concept of **Stein discrepancy**.

Definition 4.8. (a) A measurable matrix-valued map τ_ν on $\mathbb{R}^d$ is called a **Stein kernel** for the centered probability measure ν if for every smooth test function $\phi : \mathbb{R}^d \to \mathbb{R}$,

$$\int_{\mathbb{R}^d} x \cdot \nabla\phi d\nu = \int_{\mathbb{R}^d} \langle \tau_\nu, \mathrm{Hess}(\phi) \rangle_{\mathrm{HS}} d\nu,$$

where $\mathrm{Hess}(\phi)$ stands for the Hessian of ϕ, and $\langle\,,\,\rangle_{\mathrm{HS}}$, and $\|\,,\,\|_{\mathrm{HS}}$ denote the usual Hilbert–Schmidt scalar product and norm, respectively.

(b) The **Stein discrepancy** of ν with respect to γ is defined as

$$S(\nu, \gamma) = \inf \left(\int_{\mathbb{R}^d} \|\tau_\nu - \mathbf{Id}\|_{\mathrm{HS}}^2 d\nu \right)^{\frac{1}{2}}$$

where the infimum is taken over all Stein kernels of ν, and takes the value $+\infty$ if a Stein kernel for ν does not exist.

We recall that the Stein kernel τ_ν is uniquely defined in dimension $d = 1$, and that unicity may fail in higher dimensions $d \geq 2$, see [73, Appendix A]. Also, $\tau_\gamma = \mathbf{Id}_{d\times d}$ is the identity matrix. We further refer to [40, 25] for existence of the Stein kernel in general settings. The interest of the Stein's discrepancy comes, e.g., from the fact that – as a simple application of Stein's method –

$$d_{TV}(\nu, \gamma) \leq 2 \int_{\mathbb{R}} |\tau_\nu - 1| d\nu \leq 2 \left(\int_{\mathbb{R}} |\tau_\nu - 1|^2 d\nu \right)^{\frac{1}{2}},$$

yielding that $d_{TV}(\nu, \gamma) \leq 2 \, S(\nu, \gamma)$; see [53] for further details.

Next, we need the notion of Wasserstein distance. Let $p \geq 1$. Given two probability measures ν and μ on the Borel sets of $\mathbb{R}^d$, whose marginals have finite moments of order p, we define the p**-Wasserstein distance** between ν and μ as

$$W_p(\nu, \mu) = \inf_{\pi} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\pi(x, y) \right)^{\frac{1}{p}}$$

where the infimum is taken over all probability measures π of $\mathbb{R}^d \times \mathbb{R}^d$ with marginals ν and μ; note that $W_1 = d_W$, as defined in Section 2.

We recall that, for a measure $\nu = h\gamma$ with a smooth density function h on $\mathbb{R}^d$,

$$H(\nu, \gamma) := \int_{\mathbb{R}^d} h \log h d\gamma = \mathrm{Ent}_\gamma(h)$$

is the **relative entropy** of the measure ν with respect to γ, and

$$I(\nu, \gamma) := \int_{\mathbb{R}^d} \frac{|\nabla h|^2}{h} d\gamma$$

is the **Fisher information** of ν with respect to γ. After having established these notions, we can state two popular probabilistic/entropic functional inequalities:

(i) [**Logarithmic Sobolev inequality**]: $\mathrm{H}(\nu, \gamma) \leq \frac{1}{2} \mathrm{I}(\nu, \gamma)$.

(ii) [**Talagrand quadratic transportation cost inequality**]:

$$\mathrm{W}_2^2(\nu, \gamma) \leq 2 \, \mathrm{H}(\nu, \gamma).$$

The next theorem is borrowed from [53], and represents a significant improvement of the previous logarithmic Sobolev and Talagrand inequalities based on the use of Stein discrepancies: the techniques used in the proof are based on an interpolation argument along the Ornstein–Uhlenbeck semigroup. The theorem establishes connections between the relative entropy H, the Stein discrepancy S, the Fisher information I, and the Wasserstein distance W, customarily called the HSI and the WSH inequalities. The reader is also referred to the recent works [40, 25, 89] for related estimates of the Stein discrepancy based on the use of **Poincaré inequalities**, as well as on **optimal transport techniques**. See [15] for a further amplification of the approach of [53], with applications to the quantitative multidimensional CLT in the 2-Wasserstein distance. See also [33].

Theorem 4.9. *Let $d\nu = h d\gamma$ be a centered probability measure on $\mathbb{R}^d$ with smooth density function h with respect to the standard Gaussian measure γ.*

(1) *Then the following **Gaussian** HSI **inequality** holds:*

$$\mathrm{H}(\nu, \gamma) \leq \frac{1}{2} \, \mathrm{S}^2(\nu, \gamma) \log \left(1 + \frac{\mathrm{I}(\nu, \gamma)}{\mathrm{S}^2(\nu, \gamma)} \right).$$

(2) *Assume further that $\mathrm{S}(\nu, \gamma)$ and $\mathrm{H}(\nu, \gamma)$ are both positive and finite. Then, the following **Gaussian** WSH **inequality** holds:*

$$\mathrm{W}_2(\nu, \gamma) \leq \mathrm{S}(\nu, \gamma) \arccos \left(e^{-\frac{\mathrm{H}(\nu, \gamma)}{\mathrm{S}^2(\nu, \gamma)}} \right).$$

The next subsection deals with the challenging problem of quantitative probabilistic approximations in infinite dimension.

4.4 Functional approximations and Dirichlet structures

Although Stein's method is already successfully used for quantifying functional limit theorems of the Donsker type (see [11, 12], as well as [34, 35, 45, 91] for a discussion of recent developments), the general problem of assessing the discrepancy between probability distributions on infinite-dimensional spaces (like, e.g., on classes of smooth functions or on the Skorohod space) is essentially open.

In the last years a new direction of research has emerged, where the ideas behind the Malliavin–Stein approach are applied in the framework of **Dirichlet structures**, in order to deal with quantitative estimates on the probabilistic approximation of Hilbert space-valued random variables. A general (and impressive!) contribution on the matter is the recent work by Bourguin and Campese [17], where the authors are able to retrieve several Hilbert space counterparts of the finite-dimensional results discussed in Section 3 above. Bourguin and Campese's approach (whose discussion

requires preliminaries that go beyond the scope of our survey) represents a substantial addition to a line of investigation intiated by L. Coutin and L. Decreusefond in the seminal works [26, 29, 27, 28, 30].

As a quick illustration, we conclude the section with two representative statements, taken from [26, 30] and [29], respectively.

Theorem 4.10 (See [26] and Section 3.2 in [30]). *Let $(N_\lambda(t) : t \geq 0)$ be a Poisson process with intensity λ. Then, as $\lambda \to \infty$,*

$$\left(\frac{N_\lambda(t) - \lambda t}{\sqrt{\lambda}} : t \geq 0 \right) \implies (B(t) : t \geq 0)$$

where the convergence takes place weakly in the Skorohkod space. Moreover, for every $\beta < \frac{1}{2}$ consider the so-called Besov–Liouville space $I_{\beta,2}$,

$$I_{\beta,2} = \left\{ f : \exists \dot{f}, \ f(x) = \frac{1}{\Gamma(\beta)} \int_0^x (x - t)^{\beta - 1} \dot{f}(t) dt \right\}.$$

Let μ_β denote the Wiener measure on the space $I_{\beta,2}$, and Q_λ be the probability measure induced by $(N_\lambda(t) : t \geq 0)$. Then, there exists a constant c_β such that

$$\sup_{\|F\|_{C_b^2(I_{\beta,2}, \mathbb{R})} \leq 1} \left| \int F dQ_\lambda - \int F d\mu_\beta \right| \leq \frac{c_\beta}{\sqrt{\lambda}}$$

where $C_b^2(I_{\beta,2}, \mathbb{R})$ is the set of twice Fréchet differentiable functionals on $I^{\beta,2}$.

The next result aims to provide a rate of convergence in the Donsker theorem in Wasserstein distance. Let $\eta \in (0, 1)$, $p \geq 1$. Define the fractional Sobolev space $W_{\eta,p}$ as the closure of the space C^1 w.r.t. norm

$$\|f\|_{\eta,p}^p := \int_0^1 |f(t)|^p dt + \int_0^1 \int_0^1 \frac{|f(t) - f(s)|^p}{|t - s|^{1+p\eta}} ds dt.$$

Also, for $n \geq 1$, define $\mathcal{A}^n = \{(k, j) : 1 \leq k \leq d, 0 \leq j \leq n - 1\}$, and let

$$S^n = \sum_{(k,j) \in \mathcal{A}^n} X_{(k,j)} h^n_{(k,j)}, \quad h^n_{(k,j)}(t) = \sqrt{n} \int_0^t \mathbf{1}_{[j/n, (j+1)/n]}(s) ds \, e_k$$

where $(e_k) : 1 \leq k \leq d$ is the canonical basis of $\mathbb{R}^d$, and $(X_{(k,j)}, (k, j) \in \mathcal{A}^n)$ is a family of independent identically distributed, $\mathbb{R}^d$-valued, random variables with $\mathbb{E}[X] = 0$, and $\mathbb{E}\|X\|_{\mathbb{R}^d}^2 = 1$, where X is a random variable which has their common distribution.

Theorem 4.11 (See Section 3 in [29]). *Let $W = W_{\eta,p}([0, 1], \mathbb{R}^d)$, and $\mu_{\eta,p}$ be the law of the d-dimensional Brownian motion B on the space W. Then, there exists a constant c such that for $X \in L^p(W; \mathbb{R}^d, \mu_{\eta,p})$ with $p \geq 3$,*

$$\sup_{F \in Lip_1(W_{\eta,p})} \left| \mathbb{E}[F(S^n)] - \mathbb{E}[F(B)] \right| \leq c \, \|X\|_{L^p}^p \, n^{-\frac{1}{6} + \frac{\eta}{3}} \ln n$$

where

$$Lip_1(W_{\eta,p}) := \left\{ F : W_{\eta,p} \to \mathbb{R}^d : \|F(x) - F(y)\|_{\mathbb{R}^d} \le \|x - y\|_{\eta,p}, \, \forall x, y \in W_{\eta,p} \right\}.$$

Further applications of the Malliavin–Stein techniques in the framework of Dirichlet structures are contained in [32, 31]. The next section focuses on a discrete Markov structure for which exact fourth moment estimates are available.

5 Bounds on the Poisson space: fourth moments, second-order Poincaré estimates and two-scale stabilization

We will now describe a nondiffusive Markov triple for which a fourth moment result analogous to Proposition 4.5 holds. Such a Markov triple is associated with the space of square-integrable functionals of a Poisson measure on a general pair $(Z, \mathscr{Z})$, where Z is a Polish space and $\mathscr{Z}$ is the associated Borel σ-field. The requirement that Z is Polish – together with several other assumptions adopted in the present section – is made in order to simplify the discussion; the reader is referred to [37, 38] for statements and proofs in the most general setting. See also [50, 51] for an exhaustive presentation of tools of stochastic analysis for functionals of Poisson processes, as well as [81] for a discussion of the relevance of variational techniques in the framework of modern stochastic geometry.

5.1 Setup

Let μ be a nonatomic σ-finite measure on $(Z, \mathscr{Z})$, and set $\mathscr{Z}_\mu := \{B \in \mathscr{Z} : \mu(B) < \infty\}$. In what follows, we will denote by

$$\eta = \{\eta(B) : B \in \mathscr{Z}\}$$

a **Poisson measure** on $(Z, \mathscr{Z})$ with **control** (or **intensity**) μ. This means that η is a random field indexed by the elements of $\mathscr{Z}$, satisfying the following two properties: (i) for every finite collection $B_1, \ldots, B_m \in \mathscr{Z}$ of pairwise disjoint sets, the random variables $\eta(B_1), \ldots, \eta(B_m)$ are stochastically independent, and (ii) for every $B \in \mathscr{Z}$, the random variable $\eta(B)$ has the Poisson distribution with mean $\mu(B)$.[3] Whenever $B \in \mathscr{Z}_\mu$, we also write $\hat{\eta}(B) := \eta(B) - \mu(B)$ and denote by

$$\hat{\eta} = \{\hat{\eta}(B) : B \in \mathscr{Z}_\mu\}$$

the **compensated Poisson measure** associated with η. Throughout this section, we assume that $\mathscr{F} = \sigma(\eta)$.

It is a well-known fact that one can regard the Poisson measure η as a random element taking values in the space $\mathbf{N}_\sigma = \mathbf{N}_\sigma(Z)$ of all σ-finite point measures χ on $(Z, \mathscr{Z})$ that satisfy $\chi(B) \in \mathbb{N}_0 \cup \{+\infty\}$ for all $B \in \mathscr{Z}$. Such a space is equipped with the smallest σ-field $\mathscr{N}_\sigma := \mathscr{N}_\sigma(Z)$ such that, for each $B \in \mathscr{Z}$, the mapping

---

[3]Here, we adopt the usual convention of identifying a Poisson random variable with mean zero (resp. with infinite mean) with an a.s. zero (resp. infinite) random variable.

$\mathbf{N}_\sigma \ni \chi \mapsto \chi(B) \in [0, +\infty]$ is measurable. In view of our assumptions on Z and following, e.g., [51, Section 6.1], throughout the paper we can assume without loss of generality that η is **proper**, in the sense that η can be P-a.s. represented in the form

$$\eta = \sum_{n=1}^{\eta(Z)} \delta_{X_n}, \tag{5.1}$$

where $\{X_n : n \geq 1\}$ is a countable collection of random elements with values in $\mathcal{Z}$ and where we write δ_z for the *Dirac measure* at z. Since we assume μ to be nonatomic, one has that $X_k \neq X_n$ for every $k \neq n$, P-a.s.

Now denote by $\mathbf{F}(\mathbf{N}_\sigma)$ the class of all measurable functions $\mathfrak{f} : \mathbf{N}_\sigma \to \mathbb{R}$ and by $\mathcal{L}^0(\Omega) := \mathcal{L}^0(\Omega, \mathscr{F})$ the class of real-valued, measurable functions F on Ω. Note that, as $\mathscr{F} = \sigma(\eta)$, each $F \in \mathcal{L}^0(\Omega)$ has the form $F = \mathfrak{f}(\eta)$ for some measurable function $\mathfrak{f}$. This $\mathfrak{f}$, called a **representative** of F, is P_η-a.s. uniquely defined, where $P_\eta = P \circ \eta^{-1}$ is the image measure of P under η. Using a representative $\mathfrak{f}$ of F, one can introduce the **add-one-cost operator** $D^+ = (D_z^+)_{z \in \mathcal{Z}}$ on $\mathcal{L}^0(\Omega)$ as follows:

$$D_z^+ F := \mathfrak{f}(\eta + \delta_z) - \mathfrak{f}(\eta), \quad z \in \mathcal{Z}. \tag{5.2}$$

Similarly, we define D^- on $\mathcal{L}^0(\Omega)$ as

$$D_z^- F := \mathfrak{f}(\eta) - \mathfrak{f}(\eta - \delta_z), \text{ if } z \in \text{supp}(\eta), \text{ and } D_z^- F := 0, \text{ otherwise}, \tag{5.3}$$

where $\text{supp}(\chi) := \{z \in \mathcal{Z} : \text{ for all } A \in \mathscr{Z} \text{ s.t.} z \in A : \chi(A) \geq 1\}$ is the support of the measure $\chi \in \mathbf{N}_\sigma$. We call $-D^-$ the **remove-one-cost operator** associated with η. We stress that the definitions of $D^+ F$ and $D^- F$ are, respectively, $P \otimes \mu$-a.e. and P-a.s. independent of the choice of the representative $\mathfrak{f}$ – see, e.g., the discussion in [37, Section 2] and the references therein. Note that the operator D^+ can be straightforwardly iterated as follows: set $D^{(1)} := D^+$ and, for $n \geq 2$ and $z_1, \ldots, z_n \in Z$ and $F \in \mathcal{L}^0(\Omega)$, recursively define

$$D_{z_1,\ldots,z_n}^{(n)} F := D_{z_1}^+ \left(D_{z_2,\ldots,z_n}^{(n-1)} F \right).$$

5.2 L^1 *integration by parts*

One of the most fundamental formulae in the theory of Poisson processes is the so-called **Mecke formula** stating that, for each measurable function $h : \mathbf{N}_\sigma \times Z \to [0, +\infty]$, the identity

$$\mathbb{E}\left[\int_Z h(\eta + \delta_z, z)\mu(dz)\right] = \mathbb{E}\left[\int_Z h(\eta, z)\eta(dz)\right] \tag{5.4}$$

holds true. In fact, the equation (5.4) characterizes the Poisson process, see [51, Chapter 4] for a detailed discussion. Such a formula can be used in order to define an (approximate) integration by parts formula on the Poisson space.

For random variables $F, G \in \mathcal{L}^0(\Omega)$ such that $D^+ F D^+ G \in L^1(P \otimes \mu)$, we define

$$\Gamma_0(F, G) := \frac{1}{2} \left\{ \int_Z (D_z^+ F D_z^+ G)\,\mu(dz) + \int_Z (D_z^- F D_z^- G)\,\eta(dz) \right\} \tag{5.5}$$

which verifies $\mathbb{E}[|\Gamma_0(F, G)|] < \infty$, and $\mathbb{E}[\Gamma_0(F, G)] = \mathbb{E}[\int_Z (D_z^+ F D_z^+ G)\, \mu(dz)]$, in view of the Mecke formula. The following statement, taken from [37], can be regarded as an integration by parts formula in the framework of Poisson random measures, playing a role similar to that of Lemma 3.1 in the setting of Gaussian fields. It is an almost direct consequence of (5.4).

Lemma 5.1 (L^1 integration by parts). *Let $G, H \in \mathcal{L}^0(\Omega)$ be such that*

$$GD^+H, \ D^+G\, D^+H \in L^1(P \otimes \mu).$$

Then,

$$\mathbb{E}\left[G\left(\int_Z D_z^+ H\, \mu(dz) - \int_Z D_z^- H\, \eta(dz)\right)\right] = -\mathbb{E}[\Gamma_0(G, H)]. \tag{5.6}$$

5.3 Multiple integrals

For an integer $p \geq 1$ we denote by $L^2(\mu^p)$ the Hilbert space of all square-integrable and real-valued functions on $\mathcal{Z}^p$ and we write $L_s^2(\mu^p)$ for the subspace of those functions in $L^2(\mu^p)$ which are μ^p-a.e. symmetric. Moreover, for ease of notation, we denote by $\|\cdot\|_2$ and $\langle \cdot, \cdot \rangle_2$ the usual norm and scalar product on $L^2(\mu^p)$ for whatever value of p. We further define $L^2(\mu^0) := \mathbb{R}$. For $f \in L^2(\mu^p)$, we denote by $I_p(f)$ the **multiple Wiener–Itô integral** of f with respect to $\hat{\eta}$. If $p = 0$, then, by convention, $I_0(c) := c$ for each $c \in \mathbb{R}$. Now let $p, q \geq 0$ be integers. The following basic properties are proved, e.g., in [50], and are analogous to the properties of multiple integrals in a Gaussian framework, as discussed in Section 3.1:

1. $I_p(f) = I_p(\tilde{f})$, where $\tilde{f}$ denotes the **canonical symmetrization** of $f \in L^2(\mu^p)$;

2. $I_p(f) \in L^2(P)$, and $\mathbb{E}\big[I_p(f)I_q(g)\big] = \delta_{p,q}\, p!\, \langle \tilde{f}, \tilde{g} \rangle_2$, where $\delta_{p,q}$ denotes the Kronecker delta symbol.

As in the Gaussian framework of Section 3.1, for $p \geq 0$ the Hilbert space consisting of all random variables $I_p(f)$, $f \in L^2(\mu^p)$, is called the *p*-**th Wiener chaos** associated with η, and is customarily denoted by C_p. It is a crucial fact that every $F \in L^2(P)$ admits a unique representation

$$F = \mathbb{E}[F] + \sum_{p=1}^{\infty} I_p(f_p), \tag{5.7}$$

where $f_p \in L_s^2(\mu^p)$, $p \geq 1$, are suitable symmetric kernel functions, and the series converges in $L^2(P)$. Identity (5.7) is the analogue of relation (3.2), and is once again referred to as the **chaotic decomposition** of the functional $F \in L^2(P)$.

The multiple integrals discussed in this section also enjoy multiplicative properties similar to formula (3.5) above – see, e.g., [50, Proposition 5] for a precise statement. One consequence of such product formulae is that, if $F \in C_p$ and $G \in C_q$ are such that FG is square-integrable, then

$$FG \in \bigoplus_{r=0}^{p+q} C_r, \tag{5.8}$$

which can be seen as a property analogous to (4.3).

5.4 Malliavin operators

We now briefly discuss Malliavin operators on the Poisson space.

1. The **domain** dom D of the **Malliavin derivative operator** D is the set of all $F \in L^2(P)$ such that the chaotic decomposition (5.7) of F satisfies $\sum_{p=1}^{\infty} p\, p!\|f_p\|_2^2 < \infty$. For such an F, the random function $Z \ni z \mapsto D_z F \in L^2(P)$ is defined via

$$D_z F = \sum_{p=1}^{\infty} p I_{p-1}\big(f_p(z, \cdot)\big), \qquad (5.9)$$

whenever z is such that the series is converging in $L^2(P)$ (this happens μ-a.e.), and set to zero otherwise; note that $f_p(z, \cdot)$ is an a.e. symmetric function on Z^{p-1}. Hence, $DF = (D_z F)_{z \in Z}$ is indeed an element of $L^2(P \otimes \mu)$. It is well-known that $F \in \text{dom}\, D$ if and only if $D^+F \in L^2(P \otimes \mu)$, and in this case

$$D_z F = D_z^+ F, \qquad P \otimes \mu\text{-a.e.} \qquad (5.10)$$

2. The domain dom $\mathbf{L}$ of the **Ornstein–Uhlenbeck generator** $\mathbf{L}$ is the set of those $F \in L^2(P)$ whose chaotic decomposition (5.7) verifies the condition $\sum_{p=1}^{\infty} p^2\, p!\|f_p\|_2^2 < \infty$ (so that dom $\mathbf{L} \subset$ dom D) and, for $F \in$ dom $\mathbf{L}$, one defines

$$\mathbf{L}F = -\sum_{p=1}^{\infty} p I_p(f_p). \qquad (5.11)$$

By definition, $\mathbb{E}[\mathbf{L}F] = 0$; also, from (5.11) it is easy to see that $\mathbf{L}$ is **symmetric**, in the sense that

$$\mathbb{E}\big[(\mathbf{L}F)G\big] = \mathbb{E}\big[F(\mathbf{L}G)\big]$$

for all $F, G \in$ dom $\mathbf{L}$. Note that, from (5.11), it is immediate that the spectrum of $-\mathbf{L}$ is given by the nonnegative integers and that $F \in$ dom $\mathbf{L}$ is an eigenfunction of $-\mathbf{L}$ with corresponding eigenvalue p if and only if $F = I_p(f_p)$ for some $f_p \in L_s^2(\mu^p)$, that is:

$$C_p = \text{Ker}(\mathbf{L} + pI).$$

The following identity corresponds to formula (65) in [50]: if $F \in$ dom $\mathbf{L}$ is such that $D^+F \in L^1(P \otimes \mu)$, then

$$\mathbf{L}F = \int_{\mathcal{Z}} \big(D_z^+ F\big)\mu(dz) - \int_{\mathcal{Z}} \big(D_z^- F\big)\eta(dz). \qquad (5.12)$$

Define for any $F \in L^2(P)$ the pseudoinverse $\mathbf{L}^{-1}$ by

$$\mathbf{L}^{-1}F = -\sum_{p=1}^{\infty} \frac{1}{p} I_p(f_p).$$

Recall [50, Section 8] the covariance identity

$$\text{Cov}(F, G) = -\int \mathbb{E}[D_z G D_z \mathbf{L}^{-1} F]\mu(dz). \qquad (5.13)$$

3. For suitable random variables $F, G \in \text{dom}\,\mathbf{L}$ such that $FG \in \text{dom}\,\mathbf{L}$, we introduce the **carré du champ operator** Γ associated with $\mathbf{L}$ by

$$\Gamma(F, G) := \frac{1}{2}\big(\mathbf{L}(FG) - F\mathbf{L}G - G\mathbf{L}F\big). \tag{5.14}$$

The symmetry of $\mathbf{L}$ implies immediately the crucial **integration by parts formula**

$$\mathbb{E}\big[(\mathbf{L}F)G\big] = \mathbb{E}\big[F(\mathbf{L}G)\big] = -\mathbb{E}\big[\Gamma(F, G)\big]; \tag{5.15}$$

we will see below that, for many random variables F, G, relation (5.15) is indeed the same as identity appearing in Lemma 5.1.

The following result – proved in [37] – provides an explicit representation of the carré-du-champ operator Γ in terms of Γ_0, as introduced in (5.5).

Proposition 5.2. *For all $F, G \in \text{dom}\,\mathbf{L}$ such that $FG \in \text{dom}\,\mathbf{L}$ and*

$$DF,\ DG,\ FDG,\ GDF \in L^1(P \otimes \mu),$$

we have that $DF = D^+F$, $DG = D^+G$, in such a way that $DF\,DG = D^+F\,D^+G \in L^1(P \otimes \mu)$, and

$$\Gamma(F, G) = \Gamma_0(F, G), \tag{5.16}$$

where Γ_0 is defined in (5.5).

One crucial consequence of this result is that the operator Γ *is not diffusive*, in the sense that the triple $(\Omega, P, \mathbf{L})$ *is not a diffusive fourth moment structure*, as introduced in Definition 4.1; it follows in particular that the machinery of Section 4 cannot be directly applied.

5.5   Fourth moment theorems

Starting at least from the reference [83] (where Malliavin calculus and Stein's method were first combined on the Poisson space), the establishing a fourth moment bound similar to Theorem 4.6 on the Poisson space has been an open problem for several years. As recalled above, the main difficulty in achieving such a result is the discrete nature of add-one-cost and remove-one-cost operators, preventing in particular the triple $(\Omega, P, \mathbf{L})$ from enjoying a diffusive property.

The next statement contains one of the main bounds proved in [38], and shows that a quantitative fourth moment bound is available on the Poisson space. Such a bound (which also has a multidimensional extension) is proved by a clever combination of Malliavin-type techniques with an infinitesimal version of the **exchangeable pairs approach** toward Stein's method – see, e.g., [23].

Theorem 5.3. *For $q \geq 2$, let $F = I_q(f_q)$ be a multiple integral of order q with respect to $\hat{\eta}$, and assume that $\mathbb{E}[F^2] = 1$. Then,*

$$d_W(F, N) \leq \left(\sqrt{\frac{2}{\pi}} + \frac{4}{3}\right)\sqrt{\mathbb{E}[F^4] - 3}.$$

One should notice that the first bound of this type was proved in [37] under slightly more restrictive assumptions; also, reference [37] contains analogous bounds in the Kolmogorov distance, that are not achievable by using exchangeable pairs. In particular, one of the key estimates used in [37] is the following remarkable equality and bound

$$\frac{1}{2q}\int_Z \mathbb{E}\big[|D_z^+ F|^4\big]\mu(dz) = \frac{3}{q}\mathbb{E}\big[F^2\Gamma(F,F)\big] - \mathbb{E}\big[F^4\big] \le \frac{4q-3}{2q}\Big(\mathbb{E}\big[F^4\big] - 3\mathbb{E}[F^2]^2\Big),$$

that are valid for every $F \in C_q$, $q \ge 2$, such that the mapping $z \mapsto D_z^+ F$ verifies some minimal integrability conditions.

5.6 Second-order Poincaré estimates

What one calls **second-order Poincaré inequalities** is a collection of analytic estimates (first established on the Poisson space in [55]) where the Wasserstein and Kolmogorov distances, between a given function of η and a Gaussian random variable, are bounded by integrated moments of iterated add-one-cost operators on the Poisson space. The rationality behind such a name is the following. Just as the Poincaré inequality

$$\operatorname{Var}(F) \le \int_Z \mathbb{E}[(D_z^+ F)^2]\mu(dz), \tag{5.17}$$

controls the variance of a random variable F by means of integrated moments of the add-one cost (see [51, Section 18.3]), the integrated moments of second-order add-one-cost $D_x^+ D_y^+ F := D_{z,y}^2 F$ controll the discrepancy between the distribution of F and that of a Gaussian random variable – a phenomenon already observed in the Gaussian setting [21, 70, 96], where gradients typically replace add-one-cost operators.

For the rest of the section, we exclusively consider square-integrable random variables F such that $F \in \operatorname{dom} D$, in such a way that $D^+ F = DF$ (up to negligible sets). The starting point for proving second-order Poincaré estimates is the covariance identity (5.13), which can be proved as in the Gaussian setting by means of chaos expansions. When one combines Stein's method with such a formula, it is however not possible to deduce the existence of a Stein kernel as in the Gaussian setting (see (3.12)), since Malliavin operators on a Poisson space *do not* enjoy an exact chain rule such as (3.7). Indeed, we have that, for sufficiently smooth mapping $f : \mathbb{R} \to \mathbb{R}$,

$$\operatorname{Cov}(F, f(F)) = -\int \mathbb{E}[D_z(f(F))D_z \mathbf{L}^{-1} F]\mu(dz)$$

$$=: -\int \mathbb{E}[f'(F)D_z F D_z \mathbf{L}^{-1} F]\mu(dz) + R$$

where we approximate $D_z(f(F)) = f(F + D_z F) - f(F)$ by $f'(F)D_z F$ with the error term

$$D_z F \int_0^1 [f'(F + t D_z F) - f'(F)]dt$$

appearing in the implicit definition of R; notice that, in general, $R \neq 0$, thus the previous computations do not yield the existence of a Stein kernel. Selecting f as in Lemma 2.1-(d), one can bound the error term in the aforementioned calculation by $|D_z F|^2$. Therefore, for F such that $\mathbb{E}[F] = 0$ and $\text{Var}[F] = 1$, one has the bound

$$d_W(F, N) \leq \sqrt{\text{Var}\left[\int D_z F D_z \mathbf{L}^{-1} F] \mu(dz)\right]} + \int \mathbb{E}[|D_z F|^2 |D_z \mathbf{L}^{-1} F|] \mu(dz).$$

Applying the Poincaré inequality (5.17) to the variance term, as well as the contraction bound [55, Lemma 3.4] for the add-one-cost

$$\mathbb{E}[|D_z \mathbf{L}^{-1} F|^p] \leq \mathbb{E}[|D_z F|^p], \quad p \geq 1,$$

and analogous estimates for the iterated add-one-cost, leads to the following theorem.

Theorem 5.4 (Second-order Poincaré estimates [55]). *Let $F \in \text{dom } D$ be such that $\mathbb{E}[F] = 0$ and $\text{Var}[F] = 1$, and let N be a standard Gaussian random variable. Then,*

$$d_W(F, N) \leq \gamma_1 + \gamma_2 + \gamma_3,$$

where

$$\gamma_1 := 2\left[\iiint \mathbb{E}[(D_x F D_y F)^2]^{1/2} \mathbb{E}[(D_{x,z}^2 F D_{y,z}^2 F)^2]^{1/2} \mu^3(dxdydz)\right]^{1/2},$$

$$\gamma_2 := \left[\iiint \mathbb{E}[(D_{x,z}^2 F D_{y,z}^2 F)^2] \mu^3(dxdydz)\right]^{1/2},$$

$$\gamma_3 := \int \mathbb{E}[|D_x F|^3] \mu(dx).$$

As mentioned above, second-order Poincaré techniques are equally useful for obtaining bounds in the Kolmogorov distance – see [55], as well as [90] for a powerful extension to the framework of multivariate normal approximations.

An example of a successful application of second-order Poincaré estimates from [55] (to which we refer the reader for a discussion of the associated literature) is the derivation of presumably optimal Berry–Esseen bounds for the total edge length of the Poisson-based nearest neighbor graph. More precisely, let η_t be a Poisson point process with intensity $t > 0$ on a convex compact set $H \subset \mathbb{R}^d$. We consider the graph with vertex set $\text{supp }\eta_t$ and edge set formed by $\{x, y\} \subset \text{supp }\eta_t$ when either x is the nearest neighbor of y or the other way around. Consider the total edge length of the graph so obtained, denoted by L_t. Then we have

$$d_W\left(\frac{L_t - \mathbb{E}[L_t]}{\sqrt{\text{Var}[L_t]}}, N\right) \leq \frac{C}{\sqrt{t}},$$

where C depends only on H. We refer the reader to [55, Theorem 7.1] for a far more general statement, and to [49] for a collection of presumably optimal bounds on the normal approximation of **exponentially stabilizing** random variables (see the next subsection).

5.7   Stabilization theory and two-scale bounds

While the second-order Poincaré estimates can provide sharp Berry–Esseen bounds, they are not always applicable. This is the case, for instance, for certain combinatorial optimization statistics or connectivity functionals of the underlying Poisson process. The problem is typically that the iterated add-one-cost of the functionals, although well-defined almost surely, are not computationally tractable, e.g., for obtaining moment estimates.

In this section, we present an alternative collection of analytic inequalities, called the **two-scale stabilization bounds**, which avoid the use of iterated add-one-cost – they are one of the main findings from [48]; see also [22] for several related estimates obtained by a discretization procedure. As their name suggests, these bounds are closely related to the stabilization theory of Penrose and Yukich [87, 86]. Such a theory originated from the ground-breaking central limit theorem of Kesten and Lee [46] for the total edge weight M_n of Euclidean minimal spanning trees (MST) with stationary Poisson points η_n in a ball of radius $n \in \mathbb{N}$. Recall that the MST is the connected graph over the vertex set η_n that minimizes its total length. Without referring to the stochastic analysis on the Poisson space, Kesten and Lee already performed a fine study of the add-one-cost of M_n (and *not* of the iterated add-one-cost) implying some moment estimates of $D_x M_n$. Penrose and Yukich [87] extrapolated the high level ideas from [46] and transformed them into a general theory applicable to (non-quantitative) central limit theorems for a plethora of problems in stochastic geometry. The theory was further extended to multivariate normal approximation by Penrose [86]. A variant of the theory using score functionals was put forward by Baryshnikov and Yukich [13].

We now define properly the notions of strong and weak stabilization. We assume for concreteness that the ambient space is $\mathbb{R}^d$ and η is a Poisson process of unit intensity. A Poisson functional $F = F(\eta)$ is **strongly stabilizing** if there exists an almost surely finite random variable R, called the **stabilization radius**, such that

$$D_0 F(\eta|_{B_R}) = D_0 F(\eta),$$

where B_R stands for a ball with radius R centered at the origin. Here is a simple example. Fix $r > 0$ and make an edge between two points in $\eta_n := \eta|_{B_n}$ within distance r. The graph $G(\eta_n, r)$ so obtained is known as the **Gilbert graph** or the **random geometric graph**. Then, the number $F(\eta_n)$ of edges within a finite window containing the origin has stabilization radius $R = r$ almost surely, since $D_0 F(\eta)$ is the number of edges incident to the origin in $G(\eta + \delta_0, r)$. Proving strong stabilization often relies on combinatorial and geometric arguments in many problems of stochastic geometry, see [87] for a list of examples. In general situations, R is genuinely random in contrast to the simple example given above.

To obtain central limit theorems, it actually suffices to show a weaker version of stabilization. We say that F is **weakly stabilizing** if for *any* sequence of measurable sets E_n satisfying $\liminf E_n = \mathbb{R}^d$, we have the almost sure convergence

$$D_0 F(\eta|_{E_n}) \to \Delta$$

where Δ is a random variable. It is clear that a strongly stabilizing functional is also weakly stabilizing with $\Delta = D_0 F(\eta)$.

Theorem 5.5 (See [87, Theorem 3.1]). *Suppose that F is weakly stabilizing and satisfies the moment condition*

$$\sup_A \mathbb{E}[|D_0 F(\eta|_A)|^4] < \infty,$$

where the supremum is taken for all balls A that contain 0. Then there exists $\sigma^2 \geq 0$, such that

$$\frac{1}{\sqrt{\mathrm{Vol}(B_n)}}(F(\eta_n) - \mathbb{E}[F(\eta_n)]) \xrightarrow{d} N(0, \sigma^2).$$

It is remarkable how few assumptions one needs in order to obtain a CLT. Notice that the limiting variance σ^2 could be 0. In [87], it was shown that $\sigma^2 > 0$ whenever Δ is not a constant. Theorem 5.5 was proved by a martingale method and does not offer insights on how fast the normalized sequence converges to normal. The latter question was addressed in a recent preprint by Lachièze-Rey, Peccati and Yang [48]. Under slightly strengthened conditions on the functionals, they assessed the rate of normal approximation in Theorem 5.5. To state one of the bounds that can be deduced from [48], we consider again the ball B_n of radius n centered at the origin, and introduce the key quantity

$$\psi_n = \psi_n(A_{n,.}) := \sup_{x \in B_n} \mathbb{E}[|D_x F(\eta|_{B_n}) - D_x F(\eta|_{A_{n,x}})|], \quad n \geq 1,$$

where $A_{n,x}$ is any measurable set indexed by n and x. In practice, we take $A_{n,x} = B_{b_n}(x) = \{y : |x - y| \leq b_n\}$ with $1 \ll b_n \ll n$ which is a local window of x compared to the scale of B_n. In what follows, we accept this choice and call ψ_n a **two-scale discrepancy** in view of this interpretation. The following result, taken from [48], can be applied in many concrete problems in stochastic geometry.

Theorem 5.6 ([48, Corollary 1.3]). *Let $\hat{F}_n = \mathrm{Var}[F(\eta_n)]^{-1/2}(F(\eta_n) - \mathbb{E}[F(\eta_n)])$ with $\eta_n = \eta|_{B_n}$ as before. Suppose that*

$$\sup_{n \in \mathbb{N}, x \in B_n} \mathbb{E}[|D_x F(\eta_n)|^p] < \infty$$

for some $p > 4$ and also that there exists an absolute constant $b > 0$ such that $\mathrm{Var}[F(\eta_n)] \geq b|B_n|$. Then there exists a finite positive constant c such that

$$\frac{1}{c} d_W(\hat{F}_n, N(0, 1)) \leq \psi_n^{\frac{1}{2}(1-\frac{4}{p})} + \left(\frac{b_n}{n}\right)^{\frac{d}{2}}.$$

This theorem simplifies and extends some arguments in the proof of a quantitative CLT for the minimal spanning trees by Chatterjee and Sen [22]. Analogous Kolmogorov bounds for univariate normal approximation, and bounds for multivariate normal approximation are also considered in [48]. More remarks are in order.

Remark 5.7.   i) The sequence (b_n) serves as a free parameter in the bound. One should keep track of the dependence of ψ_n on b_n and make an optimization in the end.

ii) For any fixed $x \in \mathbb{R}^d$, applying the weak stabilization condition for F with two sequences (B_n) and $(B_{b_n}(x))$ (together with the translation invariance of η and the moment assumption for the add-one-cost) yields the following convergence

$$\mathbb{E}[|D_x F(\eta|_{B_n}) - D_x F(\eta|_{B_{b_n}(x)})|] \to 0.$$

As such, Theorem 5.6 quantifies Theorem 5.5 after uniformly strengthening the assumptions of Theorem 5.5.

iii) When the functional is strongly stabilizing, this bound takes an even simpler form. More precisely, we say R_x is a stabilization radius at x if

$$D_x F(\eta|_{B_R(x)}) = D_x F(\eta).$$

Then, applying Hölder's inequality and the uniform moment condition for the add-one-cost leads to the existence a positive finite c such that

$$\psi_n \leq c \sup_{x \in B_n} \mathbb{P}[R_x \geq b_n]^{1 - \frac{1}{p}}.$$

Hence, the upper tail of R_x is relevant in the rate of normal approximation. One may further classify the stabilization condition with regards to the decay of the upper tail. For instance, we say that the funcitonal F is **exponentially stabilizing** if R_x has a sub-exponential upper tail.

iv) There are some general methods for obtaining lower bounds of variance. For example, one can partition the space into nonoverlapping cubes of appropriate size then use projection method for functions of independent random variables such as Hoeffding decomposition. Another method via chaos expansion was given in [55, Section 5]

We mention one application where the second order Poincaré estimates do not apply but the two-scale stabilization bounds do. Fix $r > 0$ and consider the number K_n of components in the Gilbert graph $G(\eta_n, 2r)$ (or equivalently the Boolean model $O_{r,n} = \cup_{x \in \eta_n} B(x, r)$) as $n \to \infty$. This corresponds to the so-called thermodynamic regime, where the family of random sets $O_r = \cup_{x \in \eta} B(x, r)$ (unbounded analogue of $O_{r,n}$) indexed by r exhibits a phase transition at some $r^* \in (0, \infty)$ defined as

$$r^* = \inf\{r : \mathbb{P}[0 \text{ is connected to infinity in } O_r] > 0\}.$$

We stress that the analysis of K_n is relatively involved in the critical phase due to the co-existence of the unbounded occupied component and the unbounded vacant component (in O_r^c). However, the following estimate was obtained in [48] for all $r > 0$ in dimension 2 using the strong stabilization bound:

$$d_W((K_n - \mathbb{E}[K_n])/\sqrt{\mathrm{Var}[K_n]}, N(0, 1)) \leq \frac{C}{n^\beta},$$

where C and β are finite positive constants. In $d \geq 3$, a polylogarithmic rate was obtained. The bottleneck of these estimates are the two-arm exponents of the critical Boolean models which are hard to improve.

More generally, when one considers higher dimensional topological statistics of the Boolean model such as the Betti numbers, it may occur that strong stabilization does not hold [98, 95, 20]. In such case, the two-scale weak stabilization bound might be well suited for obtaining quantitative CLT's.

6 Malliavin–Stein method for targets in the second Wiener chaos

In this section, we present a short overview of some recent developments of the Mallaivin–Stein approach for target distributions in the second Gaussian Wiener chaos. We also formulate some complementary conjectures. We adopt the same notation as in Section 3.1 above. Let W stand for an isonormal Gaussian process on a separable Hilbert space $\mathfrak{H}$. Recall that the elements in the second Wiener chaos are random variables having the general form $F = I_2(f)$, with $f \in \mathfrak{H}^{\odot 2}$. With any kernel $f \in \mathfrak{H}^{\odot 2}$, we associate the following **Hilbert–Schmidt** operator

$$A_f : \mathfrak{H} \mapsto \mathfrak{H}; \quad g \mapsto f \otimes_1 g.$$

We also write $\{\alpha_{f,k}\}_{k\geq 1}$ and $\{e_{f,k}\}_{k\geq 1}$, respectively, to indicate the (not necessarily distinct) eigenvalues of A_f and the corresponding eigenvectors. The next proposition gathers together some relevant properties of the elements of the second Wiener chaos associated with W.

Proposition 6.1 (See Section 2.7.4 in [66]). *Let $F = I_2(f)$, $f \in \mathfrak{H}^{\odot 2}$, be a generic element of the second Wiener chaos of W.*

1. *The following equality holds: $F = \sum_{k\geq 1} \alpha_{f,k}(N_k^2 - 1)$, where $\{N_k\}_{k\geq 1}$ is a sequence of i.i.d. $\mathcal{N}(0, 1)$ random variables that are elements of the isonormal process W, and the series converges in $L^2(\Omega)$ and almost surely.*

2. *For any $r \geq 2$,*
$$\kappa_r(F) = 2^{r-1}(r-1)! \sum_{k\geq 1} \alpha_{f,k}^r.$$

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

For the rest of the section, to avoid unnecessary complication, we consider target distributions in the second Wiener chaos of the form

$$F_\infty = \sum_{i=1}^{d} \alpha_{\infty,i} (N_i^2 - 1) \tag{6.1}$$

where $N_i \sim \mathcal{N}(0, 1)$ are i.i.d, and the coefficients $(\alpha_{\infty,i} : i = 1, \ldots, d)$ are distinct, and $\alpha_{\infty,i} = 0$ for $i \geq d + 1$. We also work under the normalization assumption $\mathbb{E}[F_\infty^2] = 1$. We highlight the following particular cases: (i) $\alpha_{\infty,i} = 1$ for $i = 1, \ldots, d$, for which the target random variable F_∞ reduces to a centered **chi-squared distribution with d degree of freedom** (here, the Malliavin–Stein method has been successfully implemented in a series of papers [63, 36, 64, 71, 6]); (ii) $d = 2$,

and $\alpha_{\infty,1} \times \alpha_{\infty,2} < 0$, in which case the target random variable F_∞ belongs to the so-called **Variance–Gamma** class of probability distributions. We refer to [41–43, 39, 7] for development of Stein and Malliavin–Stein methods for the Variance–Gamma distributions.

To any target distribution F_∞ of the form (6.1) we attach the following polynomial

$$Q(x) = \left(P(x)\right)^2 := \left(x \prod_{i=1}^{d} (x - \alpha_{\infty,i})\right)^2. \tag{6.2}$$

It turns out that polynomials P and Q plays a major role in quantitative limit theorems in this setup. The next result provides a (suitable) Stein operator for target distributions F_∞ in the second Wiener chaos. Also, the stability phenomenon of the weak convergence of the sequences in the second Wiener chaos is studied in [69] using tools from complex analysis.

Theorem 6.2 (Stein characterization [3]). *Let F_∞ be an element of the second Wiener chaos of the form* (6.1). *Assume that F is a generic centered random variable living in a finite sum of Wiener chaoses (hence smooth in the sense of Malliavin calculus). Then, $F = F_\infty$ (equality in distribution) if and only if $\mathbb{E}[\mathcal{A}_\infty f(F)] = 0$ where the differential operator $\mathcal{A}_\infty$ is given by*

$$\mathcal{A}_\infty f(x) := \sum_{l=2}^{d+1} (b_l - a_{l-1}x) f^{(d+2-l)}(x) - a_{d+1} x f(x), \tag{6.3}$$

for all functions $f : \mathbb{R} \to \mathbb{R}$ such that $\mathcal{A}_\infty f(F) \in L^1(\Omega)$ and coefficients

$$a_l := \frac{P^{(l)}(0)}{l! 2^{l-1}}, \quad 1 \le l \le d+1, \tag{6.4}$$

$$b_l := \sum_{r=l}^{d+1} \frac{a_r}{(r-l+1)!} \kappa_{r-l+2}(F_\infty), \quad 2 \le l \le d+1. \tag{6.5}$$

The polynomials P and Q are given by relation (6.2).

The next conjecture puts forward a non-Gaussian counterpart to the Stein's Lemma 2.1.

Conjecture 6.3 (Stein Universality Lemma). *Let $\mathcal{H}$ denote an appropriate separating (see [66, Definition C.1.1]) class of test functions. For every given test function $h \in \mathcal{H}$ consider the associated Stein equation*

$$\mathcal{A}_\infty f(x) = h(x) - \mathbb{E}[h(F_\infty)]. \tag{6.6}$$

Then, equation (6.6) *admits a bounded d times differentiable solution f_h such that $\|f_h^{(r)}\|_\infty < +\infty$ for all $r = 1, \ldots, d$ and the bounds are independent of the given test function h.*

The rest of the section is devoted to several quantitative estimates involving target distributions in the second Wiener chaos. The first estimate is stated in terms of the 2-Wasserstein transport distance W_2 (see Section 4.3 for definition). See also [47] for several related results of a quantitative nature.

Theorem 6.4 ([2]). *Let* $(F_n : n \geq 1)$ *be a sequence of random variables belonging to the second Wiener chaos associated to the isonormal process W so that $\mathbb{E}[F_n^2] = 1$ for all $n \geq 1$. Assume that the target random variable F_∞ takes the form (6.1). Define*

$$\Delta(F_n) := \sum_{r=2}^{deg(Q)} \frac{Q^{(r)}(0)}{r!} \frac{\kappa_r(F_n)}{(r-1)!2^{r-1}}, \tag{6.7}$$

where the polynomial Q is given by (6.2). Then, there exists a constant $C > 0$ (possibly depending only on the target random variable F_∞ but independent of n) such that

$$W_2(F_n, F_\infty) \leq C \left(\sqrt{\Delta(F_n)} + \sum_{r=2}^{d+1} |\kappa_r(F_n) - \kappa_r(F_\infty)| \right). \tag{6.8}$$

Example 6.5. Consider the target random variable F_∞ of the form (6.1) with $d = 2$ and $\alpha_{\infty,1} = -\alpha_{\infty,2} = 1/2$. Hence, F_∞ ($= N_1 \times N_2$, where $N_1, N_2 \sim \mathcal{N}(0, 1)$ are independent and equality holds in law) belongs to the class of Variance–Gamma distributions $VG_c(r, \theta, \sigma)$ with parameters $r = \sigma = 1$ and $\theta = 0$. Then, [39, Corollary 5.10, part (a)] reads

$$d_W(F_n, F_\infty) \leq C \sqrt{\Delta(F_n) + 1/4\,\kappa_3^2(F_n)} \tag{6.9}$$

that is in line with the estimate (6.8). One has to note that $\kappa_3(F_\infty) = 0$.

The next result provides a quantitative bound in the **Kolmogorov distance**. The proof relies on the classical Berry–Essen estimate in terms of bounding the difference of the characteristic functions. We recall that for two real-valued random variables X and Y the Kolmogorov distance is defined as

$$d_{Kol}(X, Y) := \sup_{x \in \mathbb{R}} \left| \mathbb{P}[X \leq x] - \mathbb{P}[Y \leq x] \right|.$$

Theorem 6.6 ([4]). *Let the target random variable F_∞ in the second Wiener chaos be of the form (6.1). Assume that $(F_n : n \geq 1)$ be a sequence of centered random elements living in a finite sum of the Wiener chaoses. Then, there exists a constant C (possibly depending on the sequence (F_n), but not on n) such that*

$$d_{Kol}(F_n, F_\infty)$$

$$\leq C \sqrt{\mathbb{E}\left[\left| \sum_{r=1}^{d+1} a_r \left(\Gamma_{r-1}(F_n) - \mathbb{E}[\Gamma_{r-1}(F_n)] \right) \right| \right] + \sum_{r=2}^{d+1} |\kappa_r(F_n) - \kappa_r(F_\infty)|} \tag{6.10}$$

$$\leq C \sqrt{\sqrt{\mathrm{Var}\left(\sum_{r=1}^{d+1} a_r \Gamma_{r-1}(F_n) \right)} + \sum_{r=2}^{d+1} |\kappa_r(F_n) - \kappa_r(F_\infty)|}$$

where the coefficients $(a_r : r = 1, \ldots, d + 1)$ are given by relation (6.4). In the particular case, when the sequence $(F_n : n \geq 1)$ belongs to the second Wiener chaos, it holds that

$$\mathrm{Var}\left(\sum_{r=1}^{d+1} a_r \Gamma_{r-1}(F_n) \right) = \Delta(F_n)$$

where the quantity $\Delta(F_n)$ is as in Theorem 6.4, and the estimate (6.10) takes the form (compare with the estimate (6.8))

$$d_{Kol}(F_n, F_\infty) \leq C \sqrt{\sqrt{\Delta(F_n)} + \sum_{r=2}^{d+1} |\kappa_r(F_n) - \kappa_r(F_\infty)|}.$$

We end the section with the following conjecture, whose object is the control of the iterated Gamma operators of Malliavin calculus appearing in the RHS of the estimate (6.10) by means of finitely many cumulants. Lastly, we point out that the forthcoming estimate (6.12) has to be compared with the famous estimate $\mathrm{Var}(\Gamma_1(F)) \leq C \, \kappa_4(F)$ in the normal approximation setting, when F is a chaotic random variable.

Conjecture 6.7. *Let F_∞ be the target random variable in the second Wiener chaos of the form (6.1). Assume that $F = I_q(f)$ is a chaotic random variable in the q-th Wiener chaos with $q \geq 2$. Then, there exists a general constant C (possibly depending on q and d) such that*

$$\mathrm{Var}\left(\sum_{r=1}^{d+1} \frac{P^{(r)}(0)}{r! 2^{r-1}} \Gamma_{r-1}(F)\right) \leq C \sum_{r=2}^{deg(Q)} \frac{Q^{(r)}(0)}{r!} \frac{\kappa_r(F)}{(r-1)! 2^{r-1}} \tag{6.11}$$

where the polynomials P and Q are given by equation (6.2). In the particular case of the normal product target distribution, i.e., $d = 2$, and $\alpha_{\infty,1} = -\alpha_{\infty,2} = 1/2$, the estimate (6.11) boils down to

$$\mathrm{Var}\left(\Gamma_2(F) - F\right) \leq C \left\{\frac{\kappa_6(F)}{5!} - 2\frac{\kappa_4(F)}{3!} + \kappa_2(F)\right\}, \tag{6.12}$$

where C is an absolute constant.

Funding

References

[1] A webpage about Stein's method and Malliavin calculus (by I. Nourdin). https://sites.google.com/site/malliavinstein/home

[2] Arras, B., Azmoodeh, E., Poly, G., Swan, Y.: A bound on the 2-Wasserstein distance between linear combinations of independent random variables. Stoch. Process. Appl. **129**(7), 2341–2375 (2019). MR3958435. https://doi.org/10.1016/j.spa.2018.07.009

[3] Arras, B., Azmoodeh, E., Poly, G., Swan, Y.: Stein characterizations for linear combinations of gamma random variables. Braz. J. Probab. Stat. **34**(2), 394–413 (2020). MR4093265. https://doi.org/10.1214/18-BJPS420

[4] Arras, B., Mijoule, G., Poly, G., Swan, Y.: A new approach to the Stein-Tikhomirov method: with applications to the second Wiener chaos and Dickman convergence (2016). arXiv:1605.06819

[5] Azmoodeh, E., Campese, S., Poly, G.: Fourth Moment Theorems for Markov diffusion generators. J. Funct. Anal. **266**(4), 2341–2359 (2014). MR3150163. https://doi.org/10.1016/j.jfa.2013.10.014

[6] Azmoodeh, E., Eichelsbacher, P., Knichel, L.: Optimal gamma approximation on Wiener space. ALEA Lat. Am. J. Probab. Math. Stat. **17**(1), 101–132 (2020). MR4057185. https://doi.org/10.30757/alea.v17-05

[7] Azmoodeh, E., Gasbarra, D.: New moments criteria for convergence towards normal product/tetilla laws (2017). arXiv:1708.07681

[8] Azmoodeh, E., Malicet, D., Mijoule, G., Poly, G.: Generalization of the Nualart-Peccati criterion. Ann. Probab. **44**(2), 924–954 (2016). MR3474463. https://doi.org/10.1214/14-AOP992

[9] Azmoodeh, E., Peccati, G., Poly, G.: Convergence towards linear combinations of chi-squared random variables: a Malliavin-based approach. In: In memoriam Marc Yor—Séminaire de Probabilités XLVII. Lecture Notes in Math., vol. 2137, pp. 339–367. Springer, Cham (2015). MR3444306. https://doi.org/10.1007/978-3-319-18585-9_16

[10] Bakry, D.: L'hypercontractivité et son utilisation en théorie des semigroupes. In: Lectures on probability theory, Saint-Flour, 1992. Lecture Notes in Math., vol. 1581, pp. 1–114. Springer, Berlin (1994). MR1307413. https://doi.org/10.1007/BFb0073872

[11] Barbour, A.D.: Stein's method for diffusion approximations. Probab. Theory Relat. Fields **84**(3), 297–322 (1990). MR1035659. https://doi.org/10.1007/BF01197887

[12] Barbour, A.D., Janson, S.: A functional combinatorial central limit theorem. Electron. J. Probab. **14**, 2352–2370 (2009). MR2556014. https://doi.org/10.1214/EJP.v14-709

[13] Baryshnikov, Yu., Yukich, J.E.: Gaussian limits for random measures in geometric probability. Ann. Appl. Probab. **15**(1A), 213–253 (2005). MR2115042. https://doi.org/10.1214/105051604000000594

[14] Biermé, H., Bonami, A., Nourdin, I., Peccati, G.: Optimal Berry-Esseen rates on the Wiener space: the barrier of third and fourth cumulants. ALEA Lat. Am. J. Probab. Math. Stat. **9**(2), 473–500 (2012). MR3069374

[15] Bonis, Th.: Stein's method for normal approximation in Wasserstein distances with application to the multivariate central limit theorem. Probab. Theory Relat. Fields **178**, 827–860 (2020). MR4168389. https://doi.org/10.1007/s00440-020-00989-4

[16] Bourguin, S., Campese, S., Leonenko, N., Taqqu, M.S.: Four moments theorems on Markov chaos. Ann. Probab. **47**(3), 1417–1446 (2019). MR3945750. https://doi.org/10.1214/18-AOP1287

[17] Bourguin, S., Campese, S.: Approximation of Hilbert-Valued Gaussians on Dirichlet structures. Electron. J. Probab. **25**, 150 (2020), 30 pp. MR4193891. https://doi.org/10.1214/20-ejp551

[18] Bakry, D., Gentil, I., Ledoux, M.: Analysis and Geometry of Markov Diffusion Operators. Springer (2014). MR3155209. https://doi.org/10.1007/978-3-319-00227-9

[19] Campese, S., Nourdin, I., Nualart, D.: Continuous Breuer-Major theorem: tightness and non-stationarity. Ann. Probab. **48**(1), 147–177 (2020). MR4079433. https://doi.org/10.1214/19-AOP1357

[20] Can, V.H., Trinh, K.D.: Random connection models in the thermodynamic regime: central limit theorems for add-one cost stabilizing functionals arXiv:2004.06313 (2020). MR4049088. https://doi.org/10.1214/19-ecp279

[21] Chatterjee, S.: Fluctuations of eigenvalues and second order Poincaré inequalities. Probab. Theory Relat. Fields **143**, 1–40 (2009). MR2449121. https://doi.org/10.1007/

s00440-007-0118-6

[22] Chatterjee, S., Sen, S.: Minimal spanning trees and Stein's method. Ann. Appl. Probab. **27**(3), 1588–1645 (2017). MR3678480. https://doi.org/10.1214/16-AAP1239

[23] Chen, L.H.Y., Goldstein, L., Shao, Q.M.: Normal approximation by Stein's method. Springer (2010). MR2732624. https://doi.org/10.1007/978-3-642-15007-4

[24] Chen, L.H., Poly, G.: Stein's method, Malliavin calculus, Dirichlet forms and the fourth moment theorem. In: Festschrift of Masatoshi Fukushima. Interdisciplinary Mathematical Sciences, vol. 17, pp. 107–130 (2015). MR3379337. https://doi.org/10.1142/9789814596534_0006

[25] Courtade, T.A., Fathi, M., Pananjady, A.: Existence of Stein kernels under a spectral gap, and discrepancy bounds. Ann. Inst. Henri Poincaré Probab. Stat. **55**(2), 777–790 (2019). MR3949953. https://doi.org/10.1214/18-aihp898

[26] Coutin, L., Decreusefond, L.: Stein's method for Brownian approximations. Commun. Stoch. Anal. **7**(3), 349–372 (2014). MR3167403. https://doi.org/10.31390/cosa.7.3.01

[27] Coutin, L., Decreusefond, L.: Higher order approximations via Stein's method. Commun. Stoch. Anal. **8**(2), 155–168 (2014). MR3269842. https://doi.org/10.31390/cosa.8.2.02

[28] Coutin, L., Decreusefond, L.: Stein's method for rough paths. Potential Anal. **53**(2), 387–406 (2020). MR4125096. https://doi.org/10.1007/s11118-019-09773-z

[29] Coutin, L., Decreusefond, L.: Donsker's theorem in Wasserstein-1 distance. Electron. Commun. Probab. **25**, 27 (2020), 13 pp. MR4089734. https://doi.org/10.1214/20-ecp308

[30] Decreusefond, L.: The Stein-Dirichlet-Malliavin method. ESAIM Proc. Surveys, vol. 51. EDP Sci., Les Ulis (2015). MR3440790. https://doi.org/10.1051/proc/201551003

[31] Decreusefond, L., Halconruy, H.: Malliavin and Dirichlet structures for independent random variables. Stoch. Process. Appl. **129**(8), 2611–2653 (2019). MR3980139. https://doi.org/10.1016/j.spa.2018.07.019

[32] Decreusefond, L., Schulte, M., Thäle, Ch.: Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry. Ann. Probab. **44**(3), 2147–2197 (2016). MR3502603. https://doi.org/10.1214/15-AOP1020

[33] Döbler, C.: Normal approximation via non-linear exchangeable pairs (2020). arXiv:2008.02272

[34] Döbler, C., Kasprzak, M.: Stein's method of exchangeable pairs in multivariate functional approximations. Elec. J. Probab. (2021, to appear). MR4235479. https://doi.org/10.18287/2541-7525-2020-26-2-23-49

[35] Döbler, C., Kasprzak, M., Peccati, G.: Functional convergence of sequential U-processes with size-dependent kernels, arXiv:2008.02272 (2019).

[36] Döbler, C., Peccati, G.: The Gamma Stein equation and non-central de Jong theorems. Bernoulli **24**(4B), 3384–3421 (2018). MR3788176. https://doi.org/10.3150/17-BEJ963

[37] Döbler, C., Peccati, G.: The fourth moment theorem on the Poisson space. Ann. Probab. **46**(4), 1878–1916 (2018). MR3813981. https://doi.org/10.1214/17-AOP1215

[38] Döbler, C., Vidotto, A., Zheng, G.: Fourth moment theorems on The Poisson space in any dimension. Electron. J. Probab. **23**, 36 (2018), 27 pp. MR3798246. https://doi.org/10.1214/18-EJP168

[39] Eichelsbacher, P., Thäle, C.: Malliavin-Stein method for Variance-Gamma approximation on Wiener space. Electron. J. Probab. **20**(123), 1–28 (2014). MR3425543. https://doi.org/10.1214/EJP.v20-4136

[40] Fathi, M.: Stein kernels and moment maps. Ann. Probab. **47**(4), 2172–2185 (2019). MR3980918. https://doi.org/10.1214/18-AOP1305

[41] Gaunt, R.E.: Variance-Gamma approximation via Stein's method. Electron. J. Probab. **19**(38), 1–33 (2014)

[42] Gaunt, R.E.: Wasserstein and Kolmogorov error bounds for variance-gamma approximation via Stein's method I. J. Theor. Probab. **33**, 465–505 (2020). https://doi.org/10.1007/s10959-018-0867-4

[43] Gaunt, R.E.: Stein factors for variance-gamma approximation in the Wasserstein and Kolmogorov distances (2020). arXiv:2008.06088. MR4064309. https://doi.org/10.1007/s10959-018-0867-4

[44] Gaunt, R.E., Mijoule, G., Swan, Y.: An algebra of Stein operators. J. Math. Anal. Appl. **469**(1), 260–279 (2019). MR3857522. https://doi.org/10.1016/j.jmaa.2018.09.015

[45] Kasprzak, M.: Stein's method for multivariate Brownian approximations of sums under dependence. Stoch. Process. Appl. **130**(8), 4927–4967 (2020). MR4108478. https://doi.org/10.1016/j.spa.2020.02.006

[46] Kesten, H., Lee, S.: The central limit theorem for weighted minimal spanning trees on random points. Ann. Appl. Probab. **6**(2), 495–527 (1996). MR1398055. https://doi.org/10.1214/aoap/1034968141

[47] Krein, Ch.: Weak convergence on Wiener space: targeting the first two chaoses. ALEA Lat. Am. J. Probab. Math. Stat. **16**(1), 85–139 (2019). MR3903026. https://doi.org/10.30757/ALEA.v16-05

[48] Lachièze-Rey, R., Peccati, G., Yang, X.: Quantitative two-scale stabilization on the Poisson space (2020). arXiv:2010.13362. MR4108865. https://doi.org/10.1090/proc/14964

[49] Lachièze-Rey, R., Schultei, M., Yukich, J.: Normal approximation for stabilizing functionals. Ann. Appl. Probab. **29**(2), 931–993 (2019). MR3910021. https://doi.org/10.1214/18-AAP1405

[50] Last, G.: Stochastic analysis for Poisson processes. In: Stochastic Analysis for Poisson Point Processes. Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry, pp. 1–36. Springer (2016). MR3585396. https://doi.org/10.1007/978-3-319-05233-5_1

[51] Last, G., Penrose, M.: Lectures on the Poisson Process. Cambridge University Press (2017). MR3791470

[52] Ledoux, M.: Chaos of a Markov operator and the fourth moment condition. Ann. Probab. **40**(6), 2439–2459 (2012). MR3050508. https://doi.org/10.1214/11-AOP685

[53] Ledoux, M., Nourdin, I., Peccati, G.: Stein's method, logarithmic Sobolev and transport inequalities. Geom. Funct. Anal. **25**, 256–306 (2015). MR3320893. https://doi.org/10.1007/s00039-015-0312-0

[54] Ledoux, M., Nourdin, I., Peccati, G.: A Stein deficit for the logarithmic Sobolev inequality. Sci. China Math. **60**(7), 1163–1180 (2016). MR3665794. https://doi.org/10.1007/s11425-016-0134-7

[55] Last, G., Peccati, G., Schulte, M.: Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. Probab. Theory Relat. Fields **165**(3–4), 667–723 (2016). MR3520016. https://doi.org/10.1007/s00440-015-0643-7

[56] Malliavin, P.: Stochastic calculus of variations and hypoelliptic operators. In: Proceedings of the International Symposium on Stochastic Differential Equations, Kyoto, pp. 195–263. Wiley, New York (1976). 1978. MR0536013

[57] Malliavin, P.: Stochastic Analysis. Springer (1997). MR1450093. https://doi.org/10.1007/978-3-642-15074-6

[58] Marinucci, D., Rossi, M., Vidotto, A.: Non-universal fluctuations of the empirical measure for isotropic stationary fields on $S^2 \times R$. Ann. App. Probab. (2021, to appear).

[59] Notarnicola, M.: Fluctuations of nodal sets on the 3-torus and general cancellation phenomena (2020). arXiv:2004.04990

[60] Nourdin, I.: In: Lectures on Gaussian approximations with Malliavin calculus. Sém. Probab., XLV, pp. 3–89 (2012). MR3185909. https://doi.org/10.1017/CBO9781139084659

[61] Nourdin, I., Nualart, D.: The functional Breuer-Major theorem. Probab. Theory Relat. Fields **176**, 203–218 (2020). MR4055189. https://doi.org/10.1007/s00440-019-00917-1

[62] Nourdin, I., Nualart, D., Peccati, G.: The Breuer-Major theorem in total variation: improved rates under minimal regularity. Stoch. Process. Appl. **131**, 1–20 (2021). MR4151212. https://doi.org/10.1016/j.spa.2020.08.007

[63] Nourdin, I., Peccati, G.: Noncentral convergence of multiple integrals. Ann. Probab. **37**(4), 1412–1426 (2009). MR2546749. https://doi.org/10.1214/08-AOP435

[64] Nourdin, I., Peccati, G.: Stein's method on Wiener chaos. Probab. Theory Relat. Fields **145**(1–2), 75–118 (2009). MR2520122. https://doi.org/10.1007/s00440-008-0162-x

[65] Nourdin, I., Peccati, G.: Cumulants on the Wiener space. J. Funct. Anal. **258**(11), 3775–3791 (2010). MR2606872. https://doi.org/10.1016/j.jfa.2009.10.024

[66] Nourdin, I., Peccati, G.: Normal Approximations with Malliavin Calculus: From Stein's Method to Universality. Cambridge Tracts in Mathematics. Cambridge University Press (2012). MR2962301. https://doi.org/10.1017/CBO9781139084659

[67] Nourdin, I., Peccati, G.: The optimal fourth moment theorem. Proc. Am. Math. Soc. **143**(7), 3123–3133 (2015). MR3336636. https://doi.org/10.1090/S0002-9939-2015-12417-3

[68] Nourdin, I., Peccati, G., Rossi, M.: Nodal statistics of planar random waves. Commun. Math. Phys. **369**(1), 99–151 (2019). MR3959555. https://doi.org/10.1007/s00220-019-03432-5

[69] Nourdin, I., Poly, G.: Convergence in law in the second Wiener/Wigner chaos. Electron. Commun. Probab. **17**(36), 1–12 (2012). MR2970700. https://doi.org/10.1214/ecp.v17-2023

[70] Nourdin, I., Peccati, G., Reinert, G.: Second order Poincaré inequalities and CLTs on Wiener space. J. Funct. Anal. **257**(4), 1005–1041 (2009). MR2527030. https://doi.org/10.1016/j.jfa.2008.12.017

[71] Nourdin, I., Peccati, G., Reinert, G.: Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. Ann. Probab. **38**(5), 1947–1985 (2010). MR2722791. https://doi.org/10.1214/10-AOP531

[72] Nourdin, I., Peccati, G., Swan, Y.: Entropy and the fourth moment phenomenon. J. Funct. Anal. **266**(5), 3170–3207 (2014). MR3158721. https://doi.org/10.1016/j.jfa.2013.09.017

[73] Nourdin, I., Peccati, G., Swan, Y.: Integration by parts and representation of information functionals. In: Proceedings of the 2014 IEEE International Symposium on Information Theory (ISIT), Honolulu, HI, pp. 2217–2221 (2015)

[74] Nourdin, I., Peccati, G., Yang, X.: Berry-Esseen bounds in the Breuer-Major CLT and Gebelein's inequality. Electron. Commun. Probab. **24**, 34 (2019). MR3978683. https://doi.org/10.1214/19-ECP241

[75] Nourdin, I., Peccati, G., Yang, X.:. Multivariate normal approximation on the Wiener space: new bounds in the convex distance (2020). arxiv:2001.02188. MR2606872. https://doi.org/10.1016/j.jfa.2009.10.024

[76] Nourdin, I., Rosinski, J.: Asymptotic independence of multiple Wiener-Ito integrals and the resulting limit laws. Ann. Probab. **42**(2), 497–526 (2014). MR3178465. https://doi.org/10.1214/12-AOP826

[77] Nualart, D.: The Malliavin Calculus and Related Topics. Probability and its Applications, 2 edn. Springer, Berlin and Heidelberg and New York (2006). MR2200233

[78] Nualart, D.: Malliavin Calculus and Its Applications. CBMS Regional Conference Series in Mathematics, vol. 110 (2009). MR2498953. https://doi.org/10.1090/cbms/110

[79] Nualart, D., Ortiz-Latorre, S.: Central limit theorems for multiple stochastic integrals and Malliavin calculus. Stoch. Process. Appl. **118**(4), 614–628 (2008). MR2394845. https://doi.org/10.1016/j.spa.2007.05.004

[80] Nualart, D., Peccati, G.: Central limit theorems for sequences of multiple stochastic integrals. Ann. Probab. **33**(1), 177–193 (2005). MR2118863. https://doi.org/10.1214/009117904000000621

[81] Peccati, G., Reitzner, M. (eds.): Stochastic Analysis for Poisson Point Processes. Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry Springer (2016). MR3444831. https://doi.org/10.1007/978-3-319-05233-5

[82] Peccati, G., Rossi, M.: Quantitative limit theorems for local functionals of arithmetic random waves. In: Combinatorics in Dynamics, Stochastics and Control, The Abel Symposium, Rosendal, Norway, August 2016, vol. 13, pp. 659–689 (2018). MR3967400. https://doi.org/10.1007/978-3-030-01593-0_23

[83] Peccati, G., Solé, J-., Utzet, F., Taqqu, M.S.: Stein's method and normal approximation of Poisson functionals. Ann. Probab. **38**, 443–478 (2010). MR2642882. https://doi.org/10.1214/10-AOP531

[84] Peccati, G., Taqqu, M.S.: Wiener Chaos: Moments, Cumulants and Diagrams. Springer (2010). MR2791919. https://doi.org/10.1007/978-88-470-1679-8

[85] Peccati, G., Vidotto, A.: Gaussian random measures generated by Berry's nodal sets. J. Stat. Phys. **178**(4), 443–478 (2020). MR4064212. https://doi.org/10.1007/s10955-019-02477-z

[86] Penrose, M.D.: Multivariate spatial central limit theorems with applications to percolation and spatial graphs. Ann. Probab. **33**(5), 1945–1991 (2005). MR2165584. https://doi.org/10.1214/009117905000000206

[87] Penrose, M.D., Yukich, J.E.: Central limit theorems for some graphs in computational geometry. Ann. Appl. Probab. **11**(4), 1005–1041 (2001). MR1878288. https://doi.org/10.1214/aoap/1015345393

[88] Rossi, M.: Random nodal lengths and Wiener chaos. In: Probabilistic Methods in Geometry, Topology and Spectral Theory. Probabilistic Methods in Geometry, Topology and Spectral Theory. Contemporary Mathematics Series, vol. 739, pp. 155–169 (2019). MR4033918. https://doi.org/10.1090/conm/739/14898

[89] Saumard, A.: Weighted Poincaré inequalities, concentration inequalities and tail bounds related to the behavior of the Stein kernel in dimension one. Bernoulli **25**(4B), 3978–4006 (2019). MR4010979. https://doi.org/10.3150/19-BEJ1117

[90] Schulte, M., Yukich, J.E.: Multivariate second order Poincaré inequalities for Poisson functionals. Electron. J. Probab. **24**, 130 (2019), 42 pp. MR4040990. https://doi.org/10.1214/19-ejp386

[91] Shih, H.-H.: On Stein's method for infinite-dimensional Gaussian approximation in abstract Wiener spaces. J. Funct. Anal. **261**(5), 1236–1283 (2011). MR2807099. https://doi.org/10.1016/j.jfa.2011.04.016

[92] Stein, C.: 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, II, pp. 583–602 (1972). MR0402873

[93] Stein, C.: Approximate computation of expectations. In: IMS. Lecture Notes-Monograph Series, vol. 7 (1986). MR0882007

[94] Todino, A.P.: A quantitative central limit theorem for the excursion area of random spherical harmonics over subdomains of S^2. J. Math. Phys. **60**(2), 023505 (2019). MR3916834. https://doi.org/10.1063/1.5048976

[95] Trinh, K.D.: On central limit theorems in stochastic geometry for add-one cost stabilizing functionals. Electron. Commun. Probab. **24**, 76 (2019), 15 pp. MR4049088. https://doi.org/10.1214/19-ecp279

[96] Vidotto, A.: An improved second order Poincaré inequality for functionals of Gaussian fields (2017). arXiv:1706.06985. MR4064306. https://doi.org/10.1007/s10959-019-00883-3

[97] Villani, C.: Optimal Transport. Old and New. Springer (2009). MR2459454. https://doi.org/10.1007/978-3-540-71050-9

[98] Yogeshwaran, D., Subag, E., Adler, R.J.: Random geometric complexes in the thermodynamic regime. Probab. Theory Relat. Fields **167**(1–2), 107–142 (2017). MR3602843. https://doi.org/10.1007/s00440-015-0678-9