Optimal Gamma Approximation on Wiener Space

Ehsan Azmoodeh, Peter Eichelsbacher and Lukas Knichel

Ehsan Azmoodeh,
Ruhr University Bochum,
Faculty of Mathematics, IB 2/115,
44780 Bochum, Germany.
E-mail address: ehsan.azmoodeh@rub.de
URL: https://www.ruhr-uni-bochum.de/ffm/Lehrstuehle/stochastik/azmoodeh.html

Peter Eichelsbacher,
Ruhr University Bochum,
Faculty of Mathematics, IB 2/101,
44780 Bochum, Germany.
E-mail address: peter.eichelsbacher@rub.de
URL: https://www.ruhr-uni-bochum.de/ffm/Lehrstuehle/stochastik/eichelsbacher.html

Lukas Knichel,
Ruhr University Bochum,
Faculty of Mathematics, IB 2/95,
44780 Bochum, Germany.
E-mail address: lukas.knichel@rub.de

Abstract. Nourdin and Peccati (2009a) established a neat characterization of
Gamma approximation on a fixed Wiener chaos in terms of convergence of only
the third and fourth cumulants. In this paper, we provide an optimal rate of
convergence in the d_2-distance in terms of the maximum of the third and fourth
cumulants analogous to the result for normal approximation in Nourdin and Peccati
(2015). In order to achieve our goal, we introduce a novel operator theory approach
to Stein's method. The recent development in Stein's method for the Gamma distribution of Döbler and Peccati (2018) plays a pivotal role in our analysis. Several
examples in the context of quadratic forms are considered to illustrate our optimal
bound.

1. Introduction and Main Result

Let $X = \{X(h) : h \in \mathfrak{H}\}$ be an isonormal Gaussian process over a separable Hilbert space $\mathfrak{H}$ on a suitable probability space $(\Omega, \mathscr{F}, P)$. In the landmark article Nualart and Peccati (2005), Nualart and Peccati discovered an astonishing central limit theorem (CLT) known nowadays as the *fourth moment theorem* for a sequence of normalized random variables inside a fixed Wiener chaos associated to X. It states that the convergence in distribution towards a standard Gaussian distribution is equivalent to the sole requirement that the fourth moments converge to 3. A few years later, their findings have created a fertile line of research, culminating in the popular article Nourdin and Peccati (2009b), introducing the so called *Malliavin-Stein* approach, an elegant combination of two probabilistic techniques namely Stein method (Stein, 1972; Chen et al., 2011) and Malliavin calculus (Nualart, 2006; Nualart and Nualart, 2018) in order to quantify the probability distance between a square integrable Wiener functional and a normal distribution. The reader may consult the excellent monograph Nourdin and Peccati (2012), as well as the constantly updated web resource https://sites.google.com/site/malliavinstein/home for a huge amount of applications and generalizations of the aforementioned results. Our study is mainly inspired by the following discovery (item (b) of the forthcoming theorem), which presents an optimal version of the fourth moment theorem. For every real-valued random variable F the quantity $\kappa_r(F)$ stands for the rth cumulant of F, see section 2.3.

Theorem 1.1 (**(Optimal) fourth moment theorem**, Nualart and Peccati, 2005; Nourdin and Peccati, 2009b, 2015). *Fix $q \geq 2$. Let $\{F_n : n \geq 1\}$ be a sequence of random variables in the qth Wiener chaos associated to X such that $\mathbb{E}[F_n^2] = 1$ for every $n \in \mathbb{N}$. Then*

(a) *$F_n \to N \sim \mathscr{N}(0,1)$ in distribution if and only if $\mathbb{E}[F_n^4] \to 3$. Also, the following quantitative estimate is in order: for $n \geq 1$,*

$$d_{TV}(F_n, N) \leq 2\sqrt{\frac{q-1}{3q}} \sqrt{\kappa_4(F_n)}. \tag{1.1}$$

(b) *Under the assumptions of item (a) there exist two constants C_1 and C_2 (independent of n) such that the following optimal rate of convergence in total variation distance holds:*

$$C_1 \max\{|\kappa_3(F_n)|, \kappa_4(F_n)\} \leq d_{TV}(F_n, N) \leq C_2 \max\{|\kappa_3(F_n)|, \kappa_4(F_n)\}.$$

Fix a parameter $\nu > 0$. In this paper, the target distribution of interest is the so called *centered Gamma* distribution denoted by $G(\nu) \sim CenteredGamma(\nu)$. This means that $G(\nu) = 2\,\widehat{G}(\nu/2) - \nu$, where $\widehat{G}(\nu/2)$ is a Gamma random variable with density $\widehat{g}(x) = x^{\frac{\nu}{2}-1}\, e^{-x}\, \Gamma(\frac{\nu}{2})^{-1}\, \mathbb{1}_{(0,\infty)}(x)$. Here $\Gamma(\nu) := \int_0^{+\infty} x^{\nu-1} e^{-x} dx$ denotes the *Euler Gamma function*. The centered Gamma distribution frequently appears as a natural limiting distribution in the context of the fourth moment theorems in several studies, see for example Azmoodeh et al. (2014, 2016, 2015); Kusuoka and Tudor (2018, 2012); Arras and Swan (2017); Eichelsbacher and Thäle (2015); Ledoux (2012); Nourdin and Rosiński (2014); Nourdin and Poly (2012); Arras et al. (2017); Eden and Víquez (2015). Our principal goal is to provide an optimal rate (analogous to that of item (b) Theorem 1.1) for the Gamma approximation on a fixed Wiener chaos. The statement of the next result is an up-to-date significant

improvement over the years of the findings in Nourdin and Peccati (2009a,b); Nourdin et al. (2010); Döbler and Peccati (2018), and it can be directly obtained, for example, by combining Döbler and Peccati (2018, Theorem 7.1) and Nourdin et al. (2010, Theorem 3.6).

Theorem 1.2. *Let $\nu > 0$. Fix $q \geq 2$ an even number (see Nourdin and Peccati, 2009a, Remark 1.3, item 3 when q is odd). Assume $F = I_q(f)$ is an element in the qth Wiener chaos such that $\mathbb{E}[F^2] = 2\nu$. Then*

$$d_1(F, G(\nu)) \leq \left(\max\{2, \frac{4}{\nu}\} \sqrt{\frac{q-1}{q}} \right) \sqrt{\mathbf{M}(F)} \tag{1.2}$$

where

$$\mathbf{M}(F) := \max \left\{ \left| \kappa_4(F) - \kappa_4(G(\nu)) \right|, \left| \kappa_3(F) - \kappa_3(G(\nu)) \right| \right\}. \tag{1.3}$$

Here d_1 stands for the so called 1-Wasserstein metric (see below for definition). As a consequence, for a sequence $\{F_n : n \geq 1\}$ of random variables in the qth Wiener chaos such that $\mathbb{E}[F_n^2] = 2\nu$ for every $n \in \mathbb{N}$, the following remarkable equivalence of asymptotic statements are in order:

(a): $F_n \to G(\nu)$ *in distribution.*
(b): $\kappa_3(F_n) \to 8\nu$, *and* $\kappa_4(F_n) \to 48\nu$.

Note that $\kappa_3(G(\nu)), \kappa_4(G(\nu)) \neq 0$ unlike the case of normal approximation. We also recall the following natural generalization of the 1-Wasserstein metric d_1 that we will make use of throughout the paper. Let X and Y be two real-valued random variables. For $k \geq 2$, define

$$d_k(X, Y) := \sup_{h \in \mathcal{H}_k} \left| \mathbb{E}[h(X)] - \mathbb{E}[h(Y)] \right|$$

where the class of the test functions is

$$\mathcal{H}_k := \{ h \in C^{k-1}(\mathbb{R}) : \|h^{(1)}\|_\infty \leq 1, \ldots, \|h^{(k)}\|_\infty \leq 1 \}.$$

Here, we adapt the following convention that for every function $f : \mathbb{R} \to \mathbb{R}$ the quantity $\|f'\|_\infty$ stands for the smallest Lipschitz constant, i.e.

$$\|f'\|_\infty = \sup_{\substack{x,y \in \mathbb{R} \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|} \in \mathbb{R} \cup \{+\infty\}. \tag{1.4}$$

It is worth pointing out that $\|f'\|_\infty$ coincides with the uniform norm of the derivative of f whenever f is differentiable. A significant and also very challenging question, which we will deal with in this paper, is whether one can either provide an optimal rate or improve the rate (1.2) available in Theorem 1.2. For a general sequence $\{F_n : n \geq 1\}$ and a suitable probability metric d (often we assume that the topology induced by metric d is stronger than convergence in distribution), following Nourdin and Peccati (2012, Definition 9.2.1), we say that a numerical sequence $\{\rho(n) : n \in \mathbb{N}\}$ of strictly positive real numbers, decreasing to 0, yields an *optimal rate* with respect to the metric d, if there exist two constants C_1 and C_2 (independent of n) such that

$$C_1 \leq \frac{d(F_n, G(\nu))}{\rho(n)} \leq C_2, \, \forall \, n \in \mathbb{N}.$$

Our main result is the following non asymptotic optimal Gamma approximation in terms of the maximum of the third and fourth cumulants within the second Wiener

chaos in full generality that improves upon the rate (1.2) by a notable square power. It presents an analogous counterpart to the same phenomenon in the case of normal approximation, see Nourdin and Peccati (2015, Theorem 1.2) or Theorem 1.1 item (b).

Theorem 1.3 (**Non asymptotic optimal Gamma approximation**). *Let $\nu > 0$, and $G(\nu) \sim CenteredGamma(\nu)$. Then there exist two general constants $0 < C_1 < C_2$ (possibly depending on the parameter ν) such that for every random variable F in the second Wiener chaos associated with X with $\mathbb{E}[F^2] = 2\nu$, it holds that*

$$C_1 \, \mathbf{M}(F) \leq d_2(F, G(\nu)) \leq C_2 \, \mathbf{M}(F), \tag{1.5}$$

where the quantity $\mathbf{M}(F)$ is given by (1.3).

Remark 1.4.   (a) A significant feature of the optimal rate (1.5) is that it is non asymptotic and a priori does not assume the law of the chaotic random variable F to be close to that of $G(\nu)$ unlike the one in item (b) of Theorem 1.1 in the normal approximation case. For the upper bound, the starting point is an adaption of the technique developed in Nourdin and Peccati (2015). However, in order to achieve the optimal upper bound we introduce a novel technique within Stein's method to split test functions relying on tools from operator theory. This is the topic of section 3. Our methodology to obtain the optimal lower bound is based on complex analysis and differs from that in Nourdin and Peccati (2015). Up to our knowledge this method is new.

(b) Theorem 1.3 has to be seen as a full generalization of the main findings of Azmoodeh et al. (2018), where we assumed some additional technical conditions.

(c) Unlike the normal approximation, due to existence of singularity in the derivative of solution of the Gamma Stein equation associated to non smooth test functions (such as indicator functions) improvement of the optimal rate (1.5) to non smooth probability metrics such as the *Kolmogorov* distance or *total variation* distance or even d_1 metric is out of the scope of the techniques developed in this paper. The main reason is that we need to employ at least twice Malliavin integration by parts formula (2.6) in order to obtain the optimal upper bound in terms of Malliavin Gamma operators Γ_j defined as (2.7). This issue also point out in Döbler and Peccati (2018, Remark 1.3, item (b)).

(d) We also shortly comment on a natural thought relating to the generalization of the optimal rate (1.5) to higher order Wiener chaoses. In fact, at least for the upper bound, our investigations imply that such an extension would come at the cost of very complicated computations involving norms of *contraction operators* to verify estimate (4.10) (possibly with a different constant) for general chaotic random variables. Furthermore, our method to achieve the optimal lower bound, relying on complex analysis, cannot be used anymore in higher order chaoses, and hence one requires the introduction of new ideas.

(e) In Gaunt et al. (2017, Theorem 3.3) the authors provide an upper optimal rate of order n^{-1} for non centered *chi-square* approximation of the *Pearson's statistic*. Although their result allows for a wide class of random variables (not necessary the Gaussian framework), however imposing the

normal distribution to the entires of the partial sums leads to a special sub-class – the so called *quadratic forms* in Gaussian random variables – of the second Wiener chaos. Furthermore, they require more smooth probability distance.

The outline of our paper is as follows: In section 2, we give a brief introduction to Malliavin calculus on the Wiener space and specify the notation used in the paper. Section 3 gathers the essential ingredients of Stein's method for the centered Gamma distribution, developed recently in Döbler and Peccati (2018). We also present our novel operator theory viewpoint inside this section. Section 4 contains the main theoretical findings of this paper – an upper bound for the d_2 distance between a general element F living in a finite sum of Wiener chaoses and the target distribution $G(\nu)$ in terms of iterated Gamma operators, as well as the optimal Gamma approximation rate. The end of this section is devoted to applications of our main findings. Lastly, we close the paper with an appendix section with focus on the newly introduced Gamma operators.

2. Preliminaries: Gaussian Analysis and Malliavin Calculus

In this section, we provide a brief introduction to Malliavin calculus and define some of the operators used in this framework. For more details, see for example the textbooks Nourdin and Peccati (2012); Nualart (2006); Nualart and Nualart (2018).

2.1. *Isonormal Gaussian Processes and Wiener Chaos.* Let $\mathfrak{H}$ be a real separable Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathfrak{H}}$, and $X = \{X(h) : h \in \mathfrak{H}\}$ be an isonormal Gaussian process, defined on some probability space $(\Omega, \mathscr{F}, P)$. This means that X is a family of centered, jointly Gaussian random variables with covariance structure $\mathbb{E}[X(g)X(h)] = \langle g, h \rangle_{\mathfrak{H}}$. We assume that $\mathscr{F}$ is the σ-algebra generated by X. For an integer $q \geq 1$, we will write $\mathfrak{H}^{\otimes q}$ or $\mathfrak{H}^{\odot q}$ to denote the q-th tensor product of $\mathfrak{H}$, or its symmetric q-th tensor product, respectively. If $H_q(x) = (-1)^q e^{x^2/2} \frac{d^q}{dx^n} e^{-x^2/2}$ is the q-th Hermite polynomial, then the closed linear subspace of $L^2(\Omega)$ generated by the family $\{H_q(X(h)) : h \in \mathfrak{H}, \|h\|_{\mathfrak{H}} = 1\}$ is called the q-th *Wiener chaos* of X and will be denoted by $\mathscr{H}_q$. For $f \in \mathfrak{H}^{\odot q}$, let $I_q(f)$ be the q-th multiple Wiener-Itô integral of f (see Nourdin and Peccati, 2012, Definition 2.7.1). An important observation is that for any $f \in \mathfrak{H}$ with $\|f\|_{\mathfrak{H}} = 1$ we have that $H_q(X(f)) = I_q(f^{\otimes q})$. As a consequence I_q provides a modified isometry from $\mathfrak{H}^{\odot q}$ onto the q-th Wiener chaos $\mathscr{H}_q$ of X. It is a well-known fact, called the *Wiener-Itô chaotic decomposition*, that any element $F \in L^2(\Omega)$ admits the expansion

$$F = \sum_{q=0}^{\infty} I_q(f_q), \tag{2.1}$$

where $f_0 = \mathbb{E}[F]$ and the $f_q \in \mathfrak{H}^{\odot q}$, $q \geq 1$ are uniquely determined.

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

$$f \otimes_r g = \sum_{i_1,\ldots,i_r=1}^{\infty} \langle f, e_{i_1} \otimes \ldots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}} \otimes \langle g, e_{i_1} \otimes \ldots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}}. \tag{2.2}$$

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 $p = q$, $f \otimes_q g = \langle f, g \rangle_{\mathfrak{H}^{\otimes q}}$. Contractions appear naturally in the *product formula* for multiple integrals: if $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$, then

$$I_p(f)I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r}\binom{q}{r} I_{p+q-2r}(f \widetilde{\otimes}_r g). \tag{2.3}$$

An important result is the following isometry property of multiple integrals. Let $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$, where $1 \leq q \leq p$. Then

$$\mathbb{E}[I_p(f)I_q(g)] = \begin{cases} p! \, \langle f, g \rangle_{\mathfrak{H}^{\otimes p}} & \text{if } p = q \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

2.2. The Malliavin Operators. We denote by $\mathscr{S}$ the set of *smooth* random variables, i.e. all random variables of the form $F = g(X(\varphi_1), \ldots, X(\varphi_n))$, where $n \geq 1$, $\varphi_1, \ldots, \varphi_n \in \mathfrak{H}$ and $g : \mathbb{R}^n \to \mathbb{R}$ is a C^∞-function, whose partial derivatives have at most polynomial growth. For these random variables, we define the *Malliavin derivative* of F with respect to X as the $\mathfrak{H}$-valued random element $DF \in L^2(\Omega, \mathfrak{H})$ defined as

$$DF = \sum_{i=1}^{\infty} \frac{\partial g}{\partial x_i} \big(X(\varphi_1), \ldots, X(\varphi_n)\big) \, \varphi_i.$$

The set $\mathscr{S}$ is dense in $L^2(\Omega)$ and using a closure argument, we can extend the domain of D to $\mathbb{D}^{1,2}$, which is the closure of $\mathscr{S}$ in $L^2(\Omega)$ with respect to the norm $\|F\|_{\mathbb{D}^{1,2}}^2 := \mathbb{E}[F^2] + \mathbb{E}[\|DF\|_{\mathfrak{H}}^2]$. See Nourdin and Peccati (2012) for a more general definition of higher order Malliavin derivatives and spaces $\mathbb{D}^{p,q}$. The Malliavin derivative satisfies the following chain-rule. If $\phi : \mathbb{R}^m \to \mathbb{R}$ is a continuously differentiable function with bounded partial derivatives and $F = (F_1, \ldots, F_m)$ is a vector of elements of $\mathbb{D}^{1,q}$ for some q, then $\phi(F) \in \mathbb{D}^{1,q}$ and

$$D\phi(F) = \sum_{i=1}^{m} \frac{\partial \phi}{\partial x_i}(F) \, DF_i. \tag{2.5}$$

Note that the conditions on ϕ are not optimal and can be weakened. For $F \in L^2(\Omega)$, with chaotic expansion as in (2.1), we define the *pseudo-inverse* of the infinitesimal generator of the Ornstein-Uhlenbeck semigroup as

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

The following integration by parts formula is one of the main ingredients to proving the main theorem of section 4.1. Let $F, G \in \mathbb{D}^{1,2}$. Then

$$\mathbb{E}[FG] = \mathbb{E}[F]\mathbb{E}[G] + \mathbb{E}[\langle DG, -DL^{-1}F \rangle_{\mathfrak{H}}]. \tag{2.6}$$

2.3. *Gamma Operators and Cumulants.* Let F be a random variable with characteristic function $\phi_F(t) = \mathbb{E}[e^{itF}]$. We define its n-th cumulant, denoted by $\kappa_n(F)$, as

$$\kappa_n(F) = \frac{1}{i^n} \frac{\partial^n}{\partial t^n} \log \phi_F(t) \Big|_{t=0},$$

provided that $F \in L^n(\Omega)$. Let F be a random variable with a finite chaos expansion. We define the operators Γ_i, $i \in \mathbb{N}_0$ via $\Gamma_0(F) := F$ and

$$\Gamma_{i+1}(F) := \langle D\Gamma_i(F), -DL^{-1}F \rangle_{\mathfrak{H}}, \quad \text{for } i \geq 0. \tag{2.7}$$

We also denote the centered versions of the Gamma-operators, i.e.

$$\overline{\Gamma}_j(F) := \Gamma_j(F) - \mathbb{E}[\Gamma_j(F)].$$

Note that there is also an alternative definition, which can be found in most other papers in this framework, see for example Definition 8.4.1 in Nourdin and Peccati (2012) or Definition 3.6 in Biermé et al. (2012). For the sake of completeness, we also mention the classical Gamma operators, which we call *alternative* Gamma operators, which we shall denote by Γ^{alt}. These are defined via

$$\Gamma_0^{alt}(F) := F \quad \text{and} \quad \Gamma_{i+1}^{alt}(F) := \langle DF, -DL^{-1}\Gamma_i^{alt}(F) \rangle_{\mathfrak{H}}, \quad \text{for } i \geq 0. \tag{2.8}$$

The classical Gamma operators are related to the cumulants of F by the following identity from Nourdin and Peccati (2010): For all $j \geq 0$, we have

$$\mathbb{E}[\Gamma_j^{alt}(F)] = \frac{1}{j!} \kappa_{j+1}(F).$$

If $j \geq 3$, this does not hold anymore for our new Gamma operators. Instead, in our next result, we will list some useful relations between the classical and the new Gamma operators.

Proposition 2.1. *Let F be a centered random variable admitting a finite chaos expansion. Then*

(a) $\Gamma_1(F) = \Gamma_1^{alt}(F)$,
(b) $\mathbb{E}[\Gamma_j(F)] = \mathbb{E}[\Gamma_j^{alt}(F)] = \frac{1}{j!} \kappa_{j+1}(F)$ *for $j = 1, 2$.*
(c) $\mathbb{E}[\Gamma_3(F)] = 2\mathbb{E}[\Gamma_3^{alt}(F)] - \text{Var}(\Gamma_1(F)) = \frac{1}{3}\kappa_4(F) - \text{Var}(\Gamma_1(F))$,
(d) *When $F = I_2(f)$, for some $f \in \mathfrak{H}^{\odot 2}$, is an element of the second Wiener chaos, then*
$$\Gamma_j(F) = \Gamma_j^{alt}(F) \quad \text{for all } j \geq 1.$$

The proofs of these statements can be found in the appendix along with an explicit representation of the Gamma operators in terms of contractions.

2.4. *Useful facts on Second Wiener Chaos.* Let $F = I_2(f)$, for some $f \in \mathfrak{H}^{\odot 2}$ be a generic element in the second Wiener chaos. It is a classical result (see Nourdin and Peccati, 2012, section 2.7.4) that these kind of random variables can be analyzed through the associated *Hilbert-Schmidt operator* $A_f : \mathfrak{H} \to \mathfrak{H}$ that maps $g \mapsto f \otimes_1 g$. Denote by $\{c_{f,i} : i \in \mathbb{N}\}$ the set of eigenvalues of A_f. We also introduce the following sequence of auxiliary kernels $\left\{ f \otimes_1^{(p)} f : p \geq 1 \right\} \subset \mathfrak{H}^{\odot 2}$, defined recursively as $f \otimes_1^{(1)} f = f$, and, for $p \geq 2$, $f \otimes_1^{(p)} f = \left(f \otimes_1^{(p-1)} f \right) \otimes_1 f$.

Proposition 2.2. *(see e.g. Nourdin and Peccati, 2012, p. 43)*

(1) *The random element F admits the representation*

$$F = \sum_{i=1}^{\infty} c_{f,i}\left(N_i^2 - 1\right), \tag{2.9}$$

where the (N_i) are i.i.d. $\mathcal{N}(0,1)$ and the series converges in $L^2(\Omega)$ and almost surely.

(2) *For every $p \geq 2$*

$$\kappa_p(F) = 2^{p-1}(p-1)! \sum_{i=1}^{\infty} c_{f,i}^p = 2^{p-1}(p-1)! \langle f, f \otimes_1^{(p-1)} f \rangle_{\mathfrak{H}}$$

$$= 2^{p-1}(p-1)! \operatorname{Tr}\left(A_f^p\right) \tag{2.10}$$

where $\operatorname{Tr}(A_f^p)$ stands for the trace of the pth power of operator A_f.

It is known that when ν is an integer, $G(\nu) \sim \bar{\chi}^2$ is a centered chi-squared random variable with ν degrees of freedom, and (2.9) shows that $G(\nu)$ is itself an element of the second Wiener chaos, where ν-many of the eigenvalues are 1 and the remaining ones are 0. Hence, in this case, we deduce from (2.10) that $\kappa_p(G(\nu)) = 2^{p-1}(p-1)!\,\nu$. Perhaps not surprisingly, this is also the case when ν is any positive real number.

Lemma 2.3. *Let $\nu > 0$ and $G(\nu) \sim CenteredGamma(\nu)$. Then*

$$\kappa_p(G(\nu)) = \begin{cases} 0, & p = 1; \\ 2^{p-1}(p-1)!\,\nu, & p \geq 2. \end{cases} \tag{2.11}$$

Proof: Since the cumulant generating function of a Gamma random variable is well-known, we can easily compute that of $G(\nu)$ to be $K(t) = \frac{\nu}{2}\log\left(\frac{1}{1-2t}\right) - \nu t$. By simple induction over p, we obtain

$$\frac{\mathrm{d}^p K}{\mathrm{d}t^p}(t) = \begin{cases} -\nu + \dfrac{\nu}{1-2t} & , p = 1; \\ \dfrac{\nu}{2} \dfrac{2^p(p-1)!}{(1-2t)^{p+1}} & , p \geq 2. \end{cases}$$

The result now follows by letting $t = 0$. $\qquad\square$

Lemma 2.4. *Let $F = I_2(f)$ for some $f \in \mathfrak{H}^{\odot 2}$, and denote by A_f the corresponding Hilbert-Schmidt operator with eigenvalues $\{c_{f,i} : i \geq 1\}$. Then for every $r \geq 1$,*

$$\operatorname{Var}\left(\Gamma_r(F) - 2\Gamma_{r-1}(F)\right) = 2^{2r+1}\sum_{i=1}^{\infty} c_{f,i}^{2r}(c_{f,i} - 1)^2$$

$$= \frac{1}{(2r+1)!}\kappa_{2r+2}(F) - \frac{4}{(2r)!}\kappa_{2r+1}(F) + \frac{4}{(2r-1)!}\kappa_{2r}(F).$$

Proof: From Azmoodeh et al. (2015) equation (24), which follows by induction on r, we have the representation

$$\overline{\Gamma}_r(F) = 2^r I_2\left(f \otimes_1^{(r+1)} f\right). \tag{2.12}$$

Fix $s \geq 1$. Let's us first to show that for every $r \geq 1$ it holds that

$$\langle f \otimes_1^{(r+1)} f, f \otimes_1^{(s)} f \rangle_{\mathfrak{H}^{\otimes 2}} = \langle f, f \otimes_1^{(r+s)} f \rangle_{\mathfrak{H}^{\otimes 2}}. \tag{2.13}$$

To do this, first without loss of generality, we can assume that $\mathfrak{H} = L^2(A, \mathcal{A}, \mu)$, where μ is a σ- finite and non-atomic measure on the measurable space $(A, \mathcal{A})$. Using Fubini Theorem, and taking into account the symmetry of the kernel f, one can write

$$\langle f \otimes_1^{(r+1)} f, f \otimes_1^{(s)} f \rangle_{\mathfrak{H}^{\otimes 2}} = \int_{A^2} (f \otimes_1^{(r+1)} f)(x, y)(f \otimes_1^{(s)} f)(x, y)\mu(dx)\mu(dy)$$

$$= \int_{A^2} \left((f \otimes_1^{(r)} f) \otimes_1 f \right)(x, y)(f \otimes_1^{(s)} f)(x, y)\mu(dx)\mu(dy)$$

$$= \int_{A^2} \left[\int_A (f \otimes_1^{(r)} f)(z, x) f(z, y)\mu(dz) \right] (f \otimes_1^{(s)} f)(x, y)\mu(dx)\mu(dy)$$

$$= \int_{A^2} f(z, y) \left[\int_A (f \otimes_1^{(r)} f)(z, x)(f \otimes_1^{(s)} f)(x, y)\mu(dx) \right] \mu(dz)\mu(dy)$$

$$= \int_{A^2} f(z, y) \left((f \otimes_1^{(r)} f) \otimes_1 (f \otimes_1^{(s)} f) \right)(z, y)\mu(dz)\mu(dy)$$

$$= \int_{A^2} f(z, y)(f \otimes_1^{(r+s)} f)(z, y)\mu(dz)\mu(dy)$$

$$= \langle f, f \otimes_1^{(r+s)} f \rangle_{\mathfrak{H}^{\otimes 2}}.$$

Now, using relation (2.12), the isometry property (2.4), and relation (2.13), we obtain

$$\text{Var}\left(\Gamma_r(F) - 2\Gamma_{r-1}(F) \right) = 2^{2r+1} \| f \otimes_1^{(r+1)} f - f \otimes_1^{(r)} f \|_{\mathfrak{H}^{\otimes 2}}^2$$

$$= 2^{2r+1} \left(\| f \otimes_1^{(r+1)} f \|_{\mathfrak{H}^{\otimes 2}}^2 - 2 \langle f, f \otimes_1^{(2r)} f \rangle_{\mathfrak{H}^{\otimes 2}} + \| f \otimes_1^{(r)} f \|_{\mathfrak{H}^{\otimes 2}}^2 \right)$$

$$= 2^{2r+1} \left(\langle f, f \otimes_1^{(2r+1)} f \rangle_{\mathfrak{H}^{\otimes 2}} - 2 \langle f, f \otimes_1^{(2r)} f \rangle_{\mathfrak{H}^{\otimes 2}} + \langle f, f \otimes_1^{(2r-1)} f \rangle_{\mathfrak{H}^{\otimes 2}} \right)$$

$$= 2^{2r+1} \text{Tr}\left(A_f^{2r+2} - 2 A_f^{2r+1} + A_f^{2r} \right).$$

The result now follows with (2.10). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3. Stein's Method for the centered Gamma distribution

Let $X_r \sim \Gamma(r, 1)$ be distributed according to a Gamma distribution with *shape parameter* $r > 0$. It means that random variable X_r admits the density

$$p_r(x) = \begin{cases} \frac{1}{\Gamma(r)} x^{r-1} e^{-x}, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

Consider the centered Gamma random variable $G(\nu) = 2 X_{\nu/2} - \nu \sim CenteredGamma(\nu)$. Stein's method for $X_{\nu/2}$ has first been studied in Luk (1994) and then later been refined in Pickett (2004). It is known (see e.g. Döbler and Peccati, 2018, equation (24)) that the Stein equation for the centered Gamma random variable $G(\nu)$ associated to the test function h is given by the following first order ODE with polynomial coefficients

$$2(x + \nu)f'(x) - xf(x) = h(x) - \mathbb{E}\left[h(G(\nu)) \right], \tag{3.2}$$

where $h : \mathbb{R} \to \mathbb{R}$ is measurable and $\mathbb{E}|h(G(\nu))| < \infty$. The following result is taken from Döbler and Peccati (2018, Theorem 2.3) and plays a crucial role in our

analysis. For the reader's convenience we restate it here. We recall our convention that for every function $f : \mathbb{R} \to \mathbb{R}$ the quantity $\|f'\|_\infty$ stands for the smallest Lipschitz constant, i.e.

$$\|f'\|_\infty = \sup_{\substack{x,y \in \mathbb{R} \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|} \in \mathbb{R} \cup \{+\infty\}. \tag{3.3}$$

Again one has to note that $\|f'\|_\infty$ coincides with the uniform norm of the derivative of f whenever f is differentiable.

Theorem 3.1. *(Döbler and Peccati, 2018, Theorem 2.3) (a) Let h be a Lipschitz-continuous function on the whole real line $\mathbb{R}$. Then there exists a unique bounded Lipschitz-continuous solution $S(h)$ to the equation (3.2) on the whole real line $\mathbb{R}$ satisfying the bounds*

$$\big\|S(h)\big\|_\infty \leq \|h'\|_\infty, \quad and \quad \big\|S(h)'\big\|_\infty \leq c_\nu \|h'\|_\infty,$$

where the constant $c_\nu = \max\{1, \frac{2}{\nu}\}$.
(b) Suppose that the function h is continuously differentiable on $\mathbb{R}$ such that both h and h' are Lipschitz-continuous. Then there is a continuously differentiable solution $S(h)$ of equation (3.2) on $\mathbb{R}$ whose derivative $S(h)'$ is Lipschitz-continuous, and moreover

$$\big\|S(h)''\big\|_\infty \leq c_\nu \|h'\|_\infty + \|h''\|_\infty.$$

3.1. *Explicit Formula for the Solution of the Stein Equation.* This section is entirely based on Döbler and Peccati (2018). It is known that a Stein equation for the $\Gamma(r, 1)$ distribution is given by

$$xf'(x) + (r - x)f(x) = h(x) - \mathbb{E}[h(X_r)], \tag{3.4}$$

where $h : \mathbb{R} \to \mathbb{R}$ is a measurable test function with $\mathbb{E}|h(X_r)| < +\infty$. Döbler and Peccati (2018, p. 3406) showed that if $h \in \mathrm{Lip}(\mathbb{R})$, then there exists a unique Lipschitz-continuous function f_h on $\mathbb{R}$ solving (3.4), given by

$$f_h(x) = \begin{cases} f_h^-(x), & x < 0, \\ f_h^+(x), & x > 0, \end{cases}$$

where for $x < 0$,

$$f_h^-(x) = \frac{1}{xq_l(x)} \int_0^x \Big(h(t) - \mathbb{E}\big[h(X_r)\big] \Big) q_l(t) dt$$

$$= -\frac{1}{xq_l(x)} \int_x^0 \Big(h(t) - \mathbb{E}\big[h(X_r)\big] \Big) q_l(t) dt$$

and $q_l(x) = -(-x)^{r-1}e^{-x}$. Also $f_h^+(x) = \frac{1}{xp_r(x)} \int_0^x \Big(h(t) - \mathbb{E}\big[h(X_r)\big] \Big) p_r(t) dt$ for $x > 0$. Furthermore, one can extend f_h^- and f_h^+ continuously by setting $f_h^-(0) = f_h^+(0) := \frac{h(0) - \mathbb{E}[h(X_r)]}{r}$. Now, for a given test function $h : \mathbb{R} \to \mathbb{R}$, set $h_1(x) := h(2x - \nu)$. Following Döbler and Peccati (2018, p. 3399), if f_h is the solution of (3.4) (with $r = \nu/2$), where h is replaced by h_1, then $S(h)(x) := \frac{1}{2} f_h\left(\frac{x+\nu}{2}\right)$ solves (3.2). Therefore, the unique bounded solution $S(h)$ of the Stein equation (3.2) admits the following explicit representation

$$S(h)(x) = \int_{-\nu}^x \left(\frac{\hat{q}(t)}{2(x+\nu)\hat{q}(x)} \mathbb{1}_{\{x \leq -\nu\}}(x) + \frac{\hat{p}_\nu(t)}{2(x+\nu)\hat{p}_\nu(x)} \mathbb{1}_{\{x > -\nu\}}(x) \right) \tag{3.5}$$

$$\Big(h(t) - \mathbb{E}\big[h(G(\nu))\big]\Big)\,dt,$$

where $\hat{p}_\nu$ is the density of the centered Gamma distribution $G(\nu)$ given by

$$\hat{p}_\nu(x) = \frac{1}{2}\,p_{\nu/2}\left(\frac{x+\nu}{2}\right) = \begin{cases} 2^{-\frac{\nu}{2}}\,\Gamma\left(\frac{\nu}{2}\right)^{-1}\,(x+\nu)^{\frac{\nu}{2}-1}\,e^{-\frac{x+\nu}{2}}, & x > -\nu \\ 0, & x \le -\nu; \end{cases}$$

and $\hat{q}(x) := \frac{1}{2}\,q_l\left(\frac{x+\nu}{2}\right) = -2^{-\frac{\nu}{2}}\big(-(x+\nu)\big)^{\frac{\nu}{2}-1}\,e^{-\frac{x+\nu}{2}}$. Also note that

$$S(h)(-\nu) = \frac{h(-\nu) - \mathbb{E}\big[h(G(\nu))\big]}{\nu}. \tag{3.6}$$

The following lemma will be used in the proof of Proposition 3.7. Using a simple adaptation, a similar statement also holds for the solution $S(h)$ corresponding to the Stein equation (3.2) of the centered Gamma distribution $G(\nu)$.

Lemma 3.2. *Let $X_r \sim \Gamma(r,1)$ with cumulative distribution function F_r, and h be a Lipschitz-continuous function. Then there exist two non-negative bounded functions U^+ on $(0,+\infty)$, and U^- on $(-\infty,0]$ such that $U^\pm \downarrow 0$ as $x \to \pm\infty$, and the following estimates are in order:*

(a) for $x > 0$ it holds that $\left|f_h'(x)\right| \le 2\|h'\|_\infty U^+(x)$,

(b) for $x < 0$ it holds that $\left|f_h'(x)\right| \le 2\|h'\|_\infty U^-(x)$.

Proof: Let $Q_l(x) := \int_x^0 (-q_l(y))dy$. Consider

$$V^+(x) := \frac{\int_0^x F_r(y)dy \int_x^\infty (1 - F_r(y))\,dy}{x^2 p_r(x)}, \text{ and } V^-(x) := \frac{(r-x)\int_x^0 Q_l(y)dy}{-x^2 q_l(x)}.$$

It is known that both estimates in parts (a) and (b) take place with $V^\pm$ instead of $U^\pm$ (see Döbler, 2015, Corollary 3.15. Part (b), and Döbler and Peccati (2018, relation (35), page 4304)). Moreover, for $x > r$, the function V^+ satisfies

$$0 \le V^+(x) \le U^+(x) := \frac{\int_x^\infty (1 - F_r(y))\,dy}{x p_r(x)} \le 1.$$

Also, it is straightforward to check that as $x \to +\infty$, the function U^+ is decreasing to 0. (It is also true that $0 \le U^+(x) \le 1$ for $0 < x \le r$ Döbler and Peccati (2018, see the top of page 3403)). Part (b) is similar. $\quad\square$

3.2. An Operator Theory Approach.

Let $a, b \in \mathbb{R}^+ \cup \{\infty\}$. Define

$$\mathcal{B}_{a,b} := \Big\{ f : \mathbb{R} \to \mathbb{R}, \text{ Lipschitz-continuous} : \|f\|_\infty < a, \text{ and } \|f'\|_\infty < b \Big\}.$$

Lemma 3.3. *Let $\mathcal{B} := \mathcal{B}_{\infty,\infty}$. For every given $h \in \mathcal{B}$, define $\|f\|_\mathcal{B} := \|f\|_\infty + \|f'\|_\infty$. Then $\|\cdot\|_\mathcal{B}$ is a norm on the real vector space $\mathcal{B}$, and furthermore the pair $(\mathcal{B}, \|\cdot\|_\mathcal{B})$ is a Banach space, the so-called Lipschitz-space.*

Proof: It is straightforward to see that the pair $(\mathcal{B}, \|\cdot\|_\mathcal{B})$ is a normed space. Furthermore, it is a classical fact that it is a Banach space, see for example Weaver (1999, Proposition 6.1.2). $\quad\square$

Lemma 3.4. *Consider the mapping $S : \mathcal{B} \to \mathcal{B}$ such that for every $h \in \mathcal{B}$, the action $S(h)$ is defined as the unique bounded solution to the centered Gamma Stein equation (3.2), which is guaranteed to exist by Theorem 3.1 item (a). Then $S(h) \in \mathcal{B}$, and S is a bounded linear operator from the Banach space $\mathcal{B}$ to itself.*

Proof: Let $h \in \mathcal{B}$. Then a direct application of Theorem 3.1 item (a) yields that $S(h) \in \mathcal{B}$. To show linearity of S, take $h_1, h_2 \in \mathcal{B}$, and $\alpha \in \mathbb{R}$. Then using the Gamma Stein equation (3.2), together with the fact that $S(h)$ is the unique bounded solution to the latter, we infer that $S(h_1 + \alpha h_2) = S(h_1) + \alpha S(h_2)$. For the boundedness of $S : \mathcal{B} \to \mathcal{B}$ we apply Theorem 3.1 part (a) to obtain

$$\|S(h)\|_{\mathcal{B}} = \|S(h)\|_{\infty} + \|S(h)'\|_{\infty} \leq \|h'\|_{\infty} + c_{\nu}\|h'\|_{\infty} \leq (1 + c_{\nu})(\|h\|_{\infty} + \|h'\|_{\infty})$$
$$= (1 + c_{\nu})\|h\|_{\mathcal{B}}.$$

Hence $\|S\| \leq 1 + c_{\nu}$. $\hfill\square$

Proposition 3.5. *Consider the bounded linear operator $S : \mathcal{B} \to \mathcal{B}$ defined as in Lemma 3.4. Then the following statements are in order.*

    (a) *The operator S does not admit any non-zero eigenvalue, i.e. if $S(h) = \lambda h$ for some non-zero constant $\lambda \in \mathbb{R}$, then necessary $h = 0$.*

    (b) *For every non-zero scalar $\lambda \in \mathbb{R}$, the operator $I + \lambda S : \mathcal{B} \to \mathcal{B}$ is a one to one map, where $I : \mathcal{B} \to \mathcal{B}$ stands for the identity operator.*

Proof: (a) By contrary assume that there exists a non-zero scalar $\lambda \in \mathbb{R}$ such that

$$S(h) = \lambda h. \tag{3.7}$$

We claim that $h(-\nu) = 0$. Otherwise introduce the auxiliary test function $g = \frac{h}{h(-\nu)} - 1$. Then, obviously, $g \in \mathcal{B}$, and moreover by virtue of relation (3.7), we have $S(g) = \lambda(g + 1)$. Furthermore, we have $\mathbb{E}[g(G(\nu))] = -\lambda\nu$, because $S(g)(-\nu) = \lambda$. Therefore, the function g satisfies the first order non-homogeneous ode

$$2\lambda(x + \nu)g' - (\lambda x + 1)g = \lambda(x + \nu). \tag{3.8}$$

Then general solutions of the ode (3.8) on the interval $(-\nu, \infty)$ are given by

$$g(x) = e^{\frac{x}{2}}(x + \nu)^{\beta}\Big\{ C_3 + \frac{1}{2}\int_{-\nu}^{x} e^{-\frac{y}{2}}(y + \nu)^{-\beta}dy \Big\}, \tag{3.9}$$

where $\beta := \frac{1 - \lambda\nu}{2\lambda}$. Now, if $\beta < 1$, then as $x \to +\infty$, we have

$$\int_{-\nu}^{x} e^{-\frac{y}{2}}(y + \nu)^{-\beta}dy \to c_{\beta} < \infty.$$

This implies that $g(x) \to +\infty$ as $x \to +\infty$, which is a contradiction to the fact that g must be a bound function. When $\beta \geq 1$, i.e. $\tilde{\beta} := 1 - \beta \leq 0$ as $x \to +\infty$, we obtain that for some finite constant d_{β} that

$$\int_{-\nu}^{x} e^{-\frac{y}{2}}(y + \nu)^{-\beta}dy \to d_{\beta}\Gamma(\tilde{\beta}),$$

which is either an infinite number or a finite number depending on whether $\tilde{\beta} \in -\mathbb{N} \cup \{0\}$ is a negative integer or not. Therefore, in any case, we have obtained that $g(x) \to +\infty$ as $x \to +\infty$, which is a contradiction. Hence always $h(-\nu) = 0$.

This implies that $\mathbb{E}\left[h(G(\nu))\right] = 0$ by using (3.6). On the other hand, $S(h) = \lambda h$ satisfies the first order ode (3.2), and therefore

$$2\lambda(x + \nu)h' - (\lambda x + 1)h = 0. \tag{3.10}$$

The general solutions of the ordinary differential equation (3.10) on the interval $(-\nu, \infty)$ are given by

$$h(x) = C_1 e^{\frac{x}{2}} (x + \nu)^{\frac{1-\nu}{2\lambda}}, \tag{3.11}$$

for some constant C_1. If $C_1 \neq 0$, then this is a contradiction to the fact that $S(h)$ is a bounded function over the whole real line. Hence it must hold that $C_1 = 0$. Similarly, the general solutions of the ordinary differential equation (3.10) on the interval $(-\infty, -\nu)$ are given by

$$h(x) = C_2 e^{-\frac{x}{2}} (-x - \nu)^{\frac{1-\nu}{2\lambda}} \tag{3.12}$$

where C_2 is a general constant. Now if $C_2 \neq 0$, we infer that $S(h)$ is unbounded on the domain $(-\infty, -\nu)$, which leads to a contradiction. Therefore $C_2 = 0$, and as a direct consequence we get $h = 0$.

(b) Assume that $\lambda \neq 0$ is a non-zero scalar. Then the mapping $I + \lambda S : \mathcal{B} \to \mathcal{B}$ is a linear operator. Hence, $I + \lambda S$ is a one to one map if and only if $\mathrm{Ker}\,(I + \lambda S) = \{0\}$, and the latter follows at once from part (a). □

Lemma 3.6. *Let $f_n : [a, b] \to \mathbb{R}$ be a sequence of L-Lipschitz continuous functions for every $n \in \mathbb{N}$: i.e. for all $x, y \in [a, b]$, and every n,*

$$\left| f_n(x) - f_n(y) \right| \leq L|x - y|.$$

Assume further that $f_n \to f$ pointwise as n tends to infinity. Then f is also an L-Lipschitz function and $f_n \to f$ uniformly.

Proof: It is elementary. □

Proposition 3.7. *The bounded linear operator $S : \mathcal{B} \to \mathcal{B}$ defined as in Lemma 3.4 is a compact operator.*

Proof: Let $U_{\mathcal{B}} := \{h \in \mathcal{B} : \|h\|_{\mathcal{B}} = \|h\|_\infty + \|h'\|_\infty \leq 1\}$ denote the unit ball of the Banach space $\mathcal{B}$. We need to show that the image $S(U_{\mathcal{B}})$ of the unit ball is a precompact set in $\mathcal{B}$, or equivalently, that every sequence $\{S(h_n)\}_{n \geq 1} \subseteq S(U_{\mathcal{B}})$ has a convergent subsequence in the topology of the Banach space $\mathcal{B}$. We divide the rest of the proof in three steps.

Step (1): First we show that there exists a subsequence $\{h_{n_k}\}_{k \geq 1}$ such that $h_{n_k} \to h$ pointwise for some $h \in U_{\mathcal{B}}$. Moreover $S(h_{n_k}) \to S(h)$, and $S(h_{n_k})' \to S(h)'$ pointwise. Note that $\{h_n\}_{n \geq 1} \subseteq U_{\mathcal{B}}$ is a bounded subset of $\mathcal{B}$. It is well known (see for example Weaver, 1999, Chapter 2 or Weaver, 2018, Theorem 2.4, and Proposition 2.1 as well as the survey Godefroy, 2015) that the Banach space $\mathcal{B}$ is a predual space, i.e. there exists a (unique) Banach space $\text{Æ}(\mathbb{R})$, the so called *Arens-Eells space*, such that $\text{Æ}(\mathbb{R})^* = \mathcal{B}$. On the other hand, the Banach-Alaoglu theorem implies that the unit ball $U_{\mathcal{B}}$ is weak-* compact. Moreover, $\mathbb{R}$ is a separable Banach space, so the Arens-Eells Banach space $\text{Æ}(\mathbb{R})$ is, too Godefroy (2015). Hence the weak-*topology on $U_{\mathcal{B}}$ is metrizable. Therefore, weak-* compact is the same as weak-* sequentially compact on the unit ball $U_{\mathcal{B}}$. It follows that the sequence $\{h_n\}_{n \geq 1}$ contains a subsequence that converges in the weak-* topology to an element $h \in U_{\mathcal{B}}$. Without loss of generality, we assume that the subsequence

is given by the sequence itself. Hence there exists an element $h \in U_\mathcal{B}$ such that $h_n \to h$ in the weak*-topology. Furthermore, the weak-* topology on the bounded subsets of $\mathcal{B}$ coincides with the topology of pointwise convergence, see Weaver (2018, Proposition 2.1). As a consequence, $h_n \to h$ pointwise (here one should not expect that $h_n \to h$ weakly; otherwise this implies that the unit ball is weakly sequentially compact, and therefore the Banach space $\mathcal{B}$ is reflexive which is a contradiction). An application of the Lebesgue dominated convergence theorem implies that $S(h_n) \to S(h)$ pointwise. Taking into account these observations together with the fact that for every $n \in \mathbb{N}$ we have

$$2(x+\nu)S(h_n)'(x) - xS(h_n)(x) = h_n(x) - \mathbb{E}\left[h_n(G(\nu))\right],$$

there exists a function f such that $S(h_n)' \to f$ pointwise. On the other hand, for every $x \in \mathbb{R}$ we have that

$$2(x+\nu)f(x) = h(x) - \mathbb{E}\left[h(G(\nu))\right] + xS(h)(x).$$

Recall that $h \in U_\mathcal{B}$. Hence, the function $S(h)$ satisfies the Gamma Stein equation

$$2(x+\nu)S(h)'(x) = h(x) - \mathbb{E}\left[h(G(\nu))\right] + xS(h)(x).$$

Hence $f = S(h)'$, and also $S(h_n)' \to S(h)'$ pointwise.

Step (2): In this step, we show that $S(U_\mathcal{B}) \subseteq C_0(\mathbb{R})$ is a family of functions having the *equivanishing at infinity* property, i.e. for every given $\epsilon > 0$, there exists a compact interval $K \subset \mathbb{R}$ such that $\left|f(x)\right| < \epsilon$ for all $f \in S(U_\mathcal{B})$ and for all $x \notin K$. To do this, we use the explicit integral representation (3.5). Note that since $\|h\|_\infty \leq 1$, we have $|h(t) - \mathbb{E}[h(G_\nu)]| \leq 2$ for all $t \in \mathbb{R}$. When $x > -\nu$, then (recall that $\hat{p}_\nu$ is the density of $G(\nu)$):

$$\left|S(h)(x)\right| = \left|\int_{-\nu}^{x} \frac{\hat{p}_\nu(t)}{2(x+\nu)\hat{p}_\nu(x)}\left(h(t) - \mathbb{E}\left[h(G_\nu)\right]\right)dt\right|$$

$$= \left|\int_{x}^{\infty} \frac{\hat{p}_\nu(t)}{2(x+\nu)\hat{p}_\nu(x)}\left(\mathbb{E}\left[h(G_\nu)\right] - h(t)\right)dt\right|$$

$$\leq \int_{x}^{\infty} \frac{\hat{p}_\nu(t)}{(x+\nu)\hat{p}_\nu(x)}\,dt = \frac{1}{x+\nu}\int_{x}^{\infty}\left(\frac{t+\nu}{x+\nu}\right)^{\frac{\nu}{2}-1}\frac{e^{-t/2}}{e^{-x/2}}\,dt.$$

Now if $\nu \leq 2$, then $\left(\frac{t+\nu}{x+\nu}\right)^{\nu/2-1} \leq 1$ and thus

$$\left|S(h)(x)\right| \leq \frac{1}{x+\nu}\int_{x}^{\infty}\frac{e^{-t/2}}{e^{-x/2}}\,dt = \frac{2}{x+\nu}\xrightarrow{x\to\infty} 0.$$

When $\nu > 2$, set $r := \lceil \nu/2 - 1\rceil$. Then for every $x \geq -1$, we have

$$\left|S(h)(x)\right| \leq \frac{e^{x/2}}{(x+\nu)^{\nu/2}}\int_{x}^{\infty}(t+\nu)^{\frac{\nu}{2}-1}e^{-t/2}\,dt \leq \frac{e^{x/2}}{(x+\nu)^{\nu/2}}\int_{x}^{\infty}(t+\nu)^{r}e^{-t/2}\,dt$$

$$= e^{\nu/2}\frac{e^{x/2}}{(x+\nu)^{\nu/2}}\int_{x+\nu}^{\infty}t^{r}e^{-t/2}\,dt$$

$$= e^{\nu/2}\frac{e^{x/2}}{(x+\nu)^{\nu/2}}e^{-\frac{x}{2}-\frac{\nu}{2}}\sum_{i=0}^{r}(-1)^{r-i+1}\frac{r!}{i!(-\frac{1}{2})^{r-i+1}}(x+\nu)^{i}$$

$$=: \frac{P(x)}{(x+\nu)^{\nu/2}},$$

where P is a polynomial of degree r. Since we always have $r < \nu/2$, it follows that $\lim_{x \to \infty} |S(h)(x)| = 0$. When $x < -\nu$, again using (3.5) of the explicit representation of the solution function $S(h)$, we get

$$|S(h)(x)| = \left| \int_{-\nu}^{x} \frac{\hat{q}_\nu(t)}{2(x+\nu)\hat{q}_\nu(x)} \Big(h(t) - \mathbb{E}\big[h(G_\nu)\big] \Big) \, dt \right|$$

$$\leq \int_{x}^{-\nu} \frac{\hat{q}_\nu(t)}{(-x-\nu)\hat{q}_\nu(x)} \, dt = \frac{1}{-x-\nu} \int_{x}^{-\nu} \left(\frac{-t-\nu}{-x-\nu} \right)^{\frac{\nu}{2}-1} \frac{e^{-t/2}}{e^{-x/2}} \, dt.$$

Hence, the case $x \to -\infty$ can now be discussed similarly. Note that the upper bounds for $|S(h)(x)|$ that we found do not depend on the choice of the test function h. Therefore, we have shown that, in addition to $S(U_{\mathcal{B}}) \subseteq C_0(\mathbb{R})$, the collection $S(U_{\mathcal{B}})$ is a family of functions that are equivanishing at infinity.

Step (3): Next we show that as $n \to \infty$,

$$\left\| S(h_n) - S(h) \right\|_{\mathcal{B}} = \left\| S(h_n) - S(h) \right\|_\infty + \left\| S(h_n)' - S(h)' \right\|_\infty \to 0. \qquad (3.13)$$

By Step (2), for a given $\epsilon > 0$, there exists a compact interval $K \subset \mathbb{R}$ such that

$$\sup_{n \geq 1} \sup_{x \notin K} \max \left\{ \left| S(h_n)(x) \right|, \left| S(h)(x) \right| \right\} < \epsilon. \qquad (3.14)$$

On the other hand, the family $(S(h_n) : n \geq 1)$ consists of 1-Lipschitz-continuous functions (see part (a), Theorem 3.1), and by step (1) converges pointwise to $S(h)$ on the compact interval K. Hence, Lemma 3.6 yields that

$$S(h_n) \to S(h) \quad \text{uniformly on } K. \qquad (3.15)$$

Finally relations (3.14) and (3.15) readily imply that $S(h_n) \to S(h)$ uniformly on the real line. Now, we are left to show that $\|S(h_n)' - S(h)'\|_\infty \to 0$. To this end, first note that for every $h \in U_{\mathcal{B}}$, and every $x \neq y \in \mathbb{R}$ it holds that $|S(h)'(x) - S(h)'(y)| \leq c_\nu \|h'\|_\infty |x-y| \leq c_\nu |x-y|$. Hence, the family $\{S(h_n)', S(h)' : n \geq 1\}$ consists of c_ν-Lipschitz continuous functions. On the other hand, Lemma 3.2 yields that the family $\{S(h_n)', S(h)' : n \geq 1\}$ is equivanishing at infinity. The result now follows. $\qquad \square$

Theorem 3.8. *Let $\lambda \in \mathbb{R}$ be a non-zero scalar. Then for every $h \in \mathcal{B}$ there exists a unique solution $g \in \mathcal{B}$ to the functional equation*

$$h = (I + \lambda S)(g) = g + \lambda S(g). \qquad (3.16)$$

Proof: This is a direct application of Propositions 3.5, 3.7, and the classical Fredholm alternative Theorem (Megginson, 1998, 3.4.24, page 329). $\qquad \square$

For $r > 0$, let $U_{\mathcal{B}}(r) := \{h \in \mathcal{B} : \|h\|_{\mathcal{B}} \leq r\}$ denote the ball of radius r.

Proposition 3.9. *Let $r_1 > 0$, and $\lambda \in \mathbb{R}$ be a non-zero scalar. Then there exists a universal constant r_2 (may depend on r_1, λ, and ν) such that for every $h \in U_{\mathcal{B}}(r_1)$ the unique solution g of the functional equation (3.16) satisfies $\|g\|_{\mathcal{B}} \leq r_2$.*

Proof: From Proposition 3.5 and Theorem 3.8, the linear bounded operator $I + \lambda S : \mathcal{B} \to \mathcal{B}$ is a bijective map. Hence the result follows at once using the inverse mapping Theorem (Megginson, 1998, 1.6.6 Corollary). $\qquad \square$

4. Optimal Gamma Approximation

4.1. *A General Stein-Malliavin Upper Bound.* In the following, we present a general Malliavin-Stein upper bound that constitutes the cornerstone to achieve our final optimal goal. We start with the following useful result.

Proposition 4.1. *Let F be a centered random variable admitting a finite chaos expansion with $\mathrm{Var}(F) = 2\nu$. Let $G(\nu) \sim CenteredGamma(\nu)$. Then there exists a constant $C > 0$ (only depending on ν), such that*

$$d_2(F, G(\nu)) \leq C \sup_{h \in \mathcal{B}_{1,1}} \mathbb{E}\left|h(F)\left(\overline{\Gamma}_1(F) - 2F\right)\right|, \tag{4.1}$$

where recall that $\mathcal{B}_{1,1} := \left\{h : \mathbb{R} \to \mathbb{R}, \text{ Lipschitz-continuous} : \|h\| \leq 1, \|h'\|_\infty \leq 1\right\}$.

Proof: Consider the centered Gamma Stein equation (3.2). Let $h \in \mathcal{H}_2$ be an arbitrary test function (note that $\mathbb{E}|h(G(\nu))| < \infty$). Then by using the Malliavin integration by parts formula (2.6), we get

$$
\begin{aligned}
|\mathbb{E}[h(F)] - \mathbb{E}[h(G(\nu))]| &= |\mathbb{E}\left[2(F + \nu)\,S(h)'(F) - FS(h)(F)\right]| \\
&= \left|\mathbb{E}\left[2(F + \nu)\,S(h)'(F) - S(h)'(F)\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\right]\right| \\
&= \left|\mathbb{E}\left[S(h)'(F)\left(\overline{\Gamma}_1(F) - 2F\right)\right]\right|.
\end{aligned}
$$

Now the claim follows at once by a direct application of Theorem 3.1. □

To simplify computations, we continue with the following useful Lemmas.

Lemma 4.2. *Let $g : \mathbb{R} \to \mathbb{R}$ be a Lipschitz-continuous function, where g and g' are bounded by a constant only depending on $\nu > 0$. Consider the solution $S(g)$ of the Gamma Stein equation (3.2) associated to the test functions g. Assume that F is a centered random variable belonging to a finite sum of Wiener chaoses with variance $\mathbb{E}[F^2] = 2\nu$. Then for any $r \in \mathbb{N}$:*

$$
\begin{aligned}
\mathbb{E}&\left[g(F)\left(\overline{\Gamma}_r(F) - 2\overline{\Gamma}_{r-1}(F)\right)\right] \\
&= -\mathbb{E}\left[S(g)'(F)\left(\overline{\Gamma}_r(F) - 2\overline{\Gamma}_{r-1}(F)\right)\left(\overline{\Gamma}_1(F) - 2F\right)\right] \\
&\quad - \mathbb{E}\left[S(g)(F)\left(\Gamma_{r+1}(F) - 2\Gamma_r(F)\right)\right].
\end{aligned}
$$

Proof: First note that $2\nu = \mathbb{E}[\Gamma_1(F)]$. Thus

$$
\begin{aligned}
\mathbb{E}&\left[g(F)\left(\overline{\Gamma}_r(F) - 2\overline{\Gamma}_{r-1}(F)\right)\right] = \mathbb{E}\left[\left(g(F) - \mathbb{E}[g(G(\nu))]\right)\left(\overline{\Gamma}_r(F) - 2\overline{\Gamma}_{r-1}(F)\right)\right] \\
&= \mathbb{E}\left[\left(2(F + \nu)S(g)'(F) - FS(g)(F)\right)\left(\overline{\Gamma}_r(F) - 2\overline{\Gamma}_{r-1}(F)\right)\right] \\
&= 2\,\mathbb{E}\left[FS(g)'(F)\,\overline{\Gamma}_r(F)\right] + \mathbb{E}[\Gamma_1(F)]\mathbb{E}\left[S(g)'(F)\,\overline{\Gamma}_r(F)\right] - \mathbb{E}\left[FS(g)(F)\,\overline{\Gamma}_r(F)\right] \\
&\quad - 4\,\mathbb{E}\left[FS(g)'(F)\,\overline{\Gamma}_{r-1}(F)\right] - 2\,\mathbb{E}[\Gamma_1(F)]\mathbb{E}\left[S(g)'(F)\,\overline{\Gamma}_{r-1}(F)\right] \\
&\quad + 2\,\mathbb{E}\left[FS(g)(F)\,\overline{\Gamma}_{r-1}(F)\right] \\
&=: \sum_{i=1}^{6} T_i.
\end{aligned}
$$

Now, we use the integration-by-parts formula (2.6) in combination with the chain rule (2.5) to obtain

$$T_3 + T_2 = -\mathbb{E}\left[FS(g)(F)\,\overline{\Gamma}_r(F)\right] + \mathbb{E}[\Gamma_1(F)]\mathbb{E}\left[S(g)'(F)\,\overline{\Gamma}_r(F)\right]$$

$$= -\mathbb{E}\big[\Gamma_1(F)\,\overline{\Gamma}_r(F)S(g)'(F)\big] - \mathbb{E}\big[S(g)(F)\Gamma_{r+1}(F)\big]$$
$$+ \mathbb{E}\big[\Gamma_1(F)\big]\mathbb{E}\big[S(g)'(F)\,\overline{\Gamma}_r(F)\big]$$
$$= -\mathbb{E}\big[\overline{\Gamma}_1(F)\,\overline{\Gamma}_r(F)S(g)'(F)\big] - \mathbb{E}\big[S(g)(F)\Gamma_{r+1}(F)\big],$$

and similarly

$$T_6 + T_5 = 2\,\mathbb{E}\big[\overline{\Gamma}_1(F)\,\overline{\Gamma}_{r-1}(F)S(g)'(F)\big] + 2\,\mathbb{E}\big[S(g)(F)\Gamma_r(F)\big].$$

Hence, putting everything together, the result follows. $\qquad\square$

Lemma 4.3. *Let $g : \mathbb{R} \to \mathbb{R}$ be a Lipschitz-continuous function, where g and g' are bounded by a constant only depending on $\nu > 0$. Assume that $S(g)$ and $S(S(g))$ stand for the solutions of the Gamma Stein equation* (3.2) *associated to the test functions g and $S(g)$ respectively. Let F be a centered random variable belonging to a finite sum of Wiener chaoses with variance $\mathbb{E}[F^2] = 2\nu$. Then the following identities take place.*

(a)

$$\mathbb{E}\Big[g(F)\Big(2F - \overline{\Gamma}_1(F)\Big)\Big] = \mathbb{E}\Big[S(g)'(F)\Big(\overline{\Gamma}_1(F) - 2F\Big)^2\Big]$$
$$+ \mathbb{E}\Big[S(g)(F)\Big(\overline{\Gamma}_2(F) - 2\overline{\Gamma}_1(F)\Big)\Big] - \Big(\mathbb{E}[S(g)(F)]\Big)\Big(\frac{1}{2}\kappa_3(F) - 2\kappa_2(F)\Big)$$

(b)

$$\mathbb{E}\Big[g(F)\Big(2F - \overline{\Gamma}_1(F)\Big)\Big]$$
$$= \mathbb{E}\Big[S(g)'(F)\Big(\overline{\Gamma}_1(F) - 2F\Big)^2\Big]$$
$$- \mathbb{E}\Big[S(S(g))'(F)\Big(\overline{\Gamma}_2(F) - 2\overline{\Gamma}_1(F)\Big)\Big(\overline{\Gamma}_1(F) - 2F\Big)\Big]$$
$$- \mathbb{E}\Big[S(S(g))(F)\Big(\overline{\Gamma}_3(F) - 2\overline{\Gamma}_2(F)\Big)\Big] - \Big(\mathbb{E}[S(g)(F)]\Big)\Big(\frac{1}{2}\kappa_3(F) - 2\kappa_2(F)\Big)$$
$$- \mathbb{E}\big[S(S(g))(F)\big]\,\mathrm{Var}\,\big(\Gamma_1(F) - 2F\big)$$
$$+ \mathbb{E}\big[S(S(g))(F)\big]\Big(\frac{1}{3}\kappa_4(F) - 3\kappa_3(F) + 4\kappa_2(F)\Big).$$

Proof: We apply Lemma 4.2 twice to obtain

$$\mathbb{E}\Big[g(F)\Big(2F - \overline{\Gamma}_1(F)\Big)\Big]$$
$$= \mathbb{E}\Big[S(g)'(F)\Big(\overline{\Gamma}_1(F) - 2F\Big)^2\Big] + \mathbb{E}\Big[S(g)(F)\Big(\Gamma_2(F) - 2\Gamma_1(F)\Big)\Big]$$
$$= \mathbb{E}\Big[S(g)'(F)\Big(\overline{\Gamma}_1(F) - 2F\Big)^2\Big] - \Big(\mathbb{E}[S(g)(F)]\Big)\Big(\mathbb{E}\,[\Gamma_2(F)] - 2\kappa_2(F)\Big)$$
$$+ \mathbb{E}\Big[S(g)(F)\Big(\overline{\Gamma}_2(F) - 2\overline{\Gamma}_1(F)\Big)\Big] \qquad \Big(\text{this completes the proof of part (a)}\Big)$$
$$= \mathbb{E}\Big[S(g)'(F)\Big(\overline{\Gamma}_1(F) - 2F\Big)^2\Big] - \Big(\mathbb{E}[S(g)(F)]\Big)\Big(\mathbb{E}\,[\Gamma_2(F)] - 2\kappa_2(F)\Big)$$
$$- \mathbb{E}\Big[S(S(g))'(F)\Big(\overline{\Gamma}_2(F) - 2\overline{\Gamma}_1(F)\Big)\Big(\overline{\Gamma}_1(F) - 2F\Big)\Big]$$
$$- \mathbb{E}\Big[S(S(g))(F)\Big(\Gamma_3(F) - 2\Gamma_2(F)\Big)\Big]$$
$$= \mathbb{E}\Big[S(g)'(F)\Big(\overline{\Gamma}_1(F) - 2F\Big)^2\Big]$$

$$- \mathbb{E}\Big[S\,(S(g))'\,(F)\Big(\overline{\Gamma}_2(F) - 2\,\overline{\Gamma}_1(F)\Big)\Big(\overline{\Gamma}_1(F) - 2F\Big)\Big]$$

$$- \mathbb{E}\Big[S\,(S(g))\,(F)\Big(\overline{\Gamma}_3(F) - 2\,\overline{\Gamma}_2(F)\Big)\Big] - \Big(\mathbb{E}[S(g)(F)]\Big)\Big(\frac{1}{2}\kappa_3(F) - 2\kappa_2(F)\Big)$$

$$+ \Big(\mathbb{E}\big[S\,(S(g))\,(F)\big]\Big)\Big(\mathbb{E}\big[\Gamma_3(F)\big] - \kappa_3(F)\Big).$$

Note that we cannot translate $\mathbb{E}[\Gamma_3(F)]$ directly into the fourth cumulant, but instead by Proposition 2.1 part *(c)*, we have $\mathbb{E}[\Gamma_3(F)] = \frac{1}{3}\kappa_4(F) - \mathrm{Var}(\Gamma_1(F))$. The variance term can be written as

$$\begin{aligned}
\mathrm{Var}\left(\Gamma_1(F)\right) &= \mathrm{Var}\left(\Gamma_1(F) - 2F\right) - 4\kappa_2(F) + 4\mathbb{E}\big[F\Gamma_1(F)\big] \\
&= \mathrm{Var}\left(\Gamma_1(F) - 2F\right) - 4\kappa_2(F) + 4\mathbb{E}\big[\Gamma_2(F)\big] \\
&= \mathrm{Var}\left(\Gamma_1(F) - 2F\right) - 4\kappa_2(F) + 2\kappa_3(F).
\end{aligned}$$

Putting everything together, the claim follows.   □

Remark 4.4. We point out that for both linear cumulant combinations appearing in the right hand sides of parts (a) and (b) in Lemma 4.3 it holds that

$$\frac{1}{2}\kappa_3(G(\nu)) - 2\kappa_2(G(\nu)) = 0, \quad \text{and} \quad \frac{1}{3}\kappa_4(G(\nu)) - 3\kappa_3(G(\nu)) + 4\kappa_2(G(\nu)) = 0.$$

Now, we are ready to state the main result of this section.

Theorem 4.5. *Let F be a centered random variable admitting a finite chaos expansion with $\mathrm{Var}(F) = 2\nu$. Let $G(\nu) \sim CenteredGamma(\nu)$. Then there exists a general constant $C > 0$ (only depending on ν), such that*

$$\begin{aligned}
d_2(F, G(\nu)) \leq C \Bigg\{ &\mathrm{Var}\left(\Gamma_1(F) - 2F\right) \\
&+ \sqrt{\mathrm{Var}\left(\Gamma_2(F) - 2\Gamma_1(F)\right)} \times \sqrt{\mathrm{Var}\left(\Gamma_1(F) - 2F\right)} \\
&+ \sqrt{\mathrm{Var}\left(\big(\Gamma_3(F) - 2\Gamma_2(F)\big) - 2\big(\Gamma_2(F) - 2\Gamma_1(F)\big)\right)} \\
&+ \Big|\kappa_3(F) - \kappa_3(G(\nu))\Big| + \Big|\kappa_4(F) - \kappa_4(G(\nu))\Big| \Bigg\}.
\end{aligned} \tag{4.2}$$

In particular, when $F = I_q(f)$ belongs to qth Wiener chaos with q an even integer, then for some general constant $C > 0$ (may depend on ν and q but, independent of the kernel f) it holds that

$$\begin{aligned}
d_2(F, G(\nu)) \leq C \Bigg\{ &\sqrt{\mathrm{Var}\left(\big(\Gamma_3(F) - 2\Gamma_2(F)\big) - 2\big(\Gamma_2(F) - 2\Gamma_1(F)\big)\right)} \\
&+ \Big|\kappa_3(F) - \kappa_3(G(\nu))\Big| + \Big|\kappa_4(F) - \kappa_4(G(\nu))\Big| \Bigg\}.
\end{aligned} \tag{4.3}$$

Proof: Using Proposition 4.1, Theorem 3.8 with $\lambda = 2$, and Proposition 3.9 we obtain that

$$\begin{aligned}
d_2(F, G(\nu)) &\leq C \sup_{h \in \mathcal{B}_{1,1}} \mathbb{E}\left|h(F)\big(\overline{\Gamma}_1(F) - 2F\big)\right| \\
&\leq C \sup_{h \in \mathcal{B}_{1,1}} \mathbb{E}\left|\big(h(F) + 2S(h)(F)\big)\big(\overline{\Gamma}_1(F) - 2F\big)\right|
\end{aligned}$$

where C stands for a general constant depending only on the parameter ν. Now, we apply Lemma 4.3 item (b) on $\mathbb{E}\left[h(F)\big(\overline{\Gamma}_1(F) - 2F\big)\right]$, and item (a) on term $\mathbb{E}\left[S(h)(F)\big(\overline{\Gamma}_1(F) - 2F\big)\right]$. Then putting everything together the estimate (4.2) follows by applying Cauchy-Schwarz inequality, Theorem 3.1, as well as using the fact that $\kappa_2(G(\nu)) = \kappa_2(F) = 2\nu$, $\kappa_3(G(\nu)) = 8\nu$ and $\kappa_4(G(\nu)) = 48\nu$, see (2.11). For the estimate (4.3), we first verify that for a general constant $C > 0$,

$$\mathrm{Var}\left(\Gamma_2(F) - 2\Gamma_1(F)\right) \leq C\,\mathrm{Var}\left(\Gamma_1(F) - 2F\right). \tag{4.4}$$

Using Lemma 5.1 we can write

$$\overline{\Gamma}_1(F) - 2F = I_q\left(c_q(q/2)f\,\widetilde{\otimes}_{q/2}\,f - 2f\right) + \sum_{\substack{r=1 \\ r \neq q/2}}^{q-1} c_q(r)I_{2q-2r}\left(f\,\widetilde{\otimes}_r\,f\right).$$

Hence,

$$\begin{aligned}
\mathrm{Var}\left(\Gamma_1(F) - 2F\right) &= q!\left\|c_q(q/2)f\,\widetilde{\otimes}_{q/2}\,f - 2f\right\|_{\mathfrak{H}}^2 \\
&\quad + \sum_{\substack{r=1 \\ r \neq q/2}}^{q-1} c_q(r)^2(2q-2r)!\|f\,\widetilde{\otimes}_r\,f\|_{\mathfrak{H}}^2.
\end{aligned} \tag{4.5}$$

Therefore, the relation (4.5) implies that for a general constant C it holds that

$$\mathrm{Var}\left(\Gamma_1(F) - 2F\right) \geq C\max\left\{\left\|c_q(q/2)f\,\widetilde{\otimes}_{q/2}\,f - 2f\right\|_{\mathfrak{H}}^2, \max_{\substack{1 \leq r \leq q-1 \\ r \neq q/2}} \|f\,\widetilde{\otimes}_r\,f\|_{\mathfrak{H}}^2\right\}$$

$$\tag{4.6}$$

On the other hand, again using Lemma 5.1, we have

$$\begin{aligned}
\overline{\Gamma}_2(F) - 2\overline{\Gamma}_1(F) &= \sum_{\substack{r=1 \\ (r,s) \neq (q/2,q) \\ r+s \leq q/2}}^{q-1} \sum_{s=1}^{[2q-2r]\wedge q} c_q(r,s)I_{3q-2(r+s)}\left((f\,\widetilde{\otimes}_r\,f)\,\widetilde{\otimes}_s\,f\right) \\
&\quad + \left\{\sum_{\substack{r=1 \\ (r,s) \neq (q/2,q) \\ q/2 < r+s < 3q/2}}^{q-1} \sum_{s=1}^{[2q-2r]\wedge q} c_q(r,s)I_{3q-2(r+s)}\left((f\,\widetilde{\otimes}_r\,f)\,\widetilde{\otimes}_s\,f\right) - 2\sum_{r'=1}^{q-1} c_q(r')I_{2q-2r'}\left(f\,\widetilde{\otimes}_{r'}\,f\right)\right\} \\
&:= S_1 + \{S_2 - 2S_3\},
\end{aligned}$$

where the coefficients c_q are defined according to relation (5.2).

Estimating S_1: for every indices (r, s) in the sum S_1 (noticing that in particular $1 \leq r \leq q - 1$, and restriction $r + s \leq q/2$ yields that $\mathbf{r} \neq \mathbf{q/2}$ and $s \neq q$) we have

$$\left\|(f\,\widetilde{\otimes}_r\,f)\,\widetilde{\otimes}_s\,f\right\|_{\mathfrak{H}}^2 \leq \|f\|_{\mathfrak{H}}^2\|f\,\widetilde{\otimes}_r\,f\|_{\mathfrak{H}}^2 = \frac{2\nu}{q!}\|f\,\widetilde{\otimes}_r\,f\|_{\mathfrak{H}}^2 \leq C\max_{\substack{1 \leq r \leq q-1 \\ r \neq q/2}} \|f\,\widetilde{\otimes}_r\,f\|_{\mathfrak{H}}^2.$$

$$\tag{4.7}$$

Furthermore, if two indices (r,s) and (r',s') in the sum S_1 produce multiple integrals of the same order then we use estimate (4.7) with the elementary inequality

$$\left\| c_q(r,s)(f \widetilde{\otimes}_r f) \widetilde{\otimes}_s f + c_q(r',s')(f \widetilde{\otimes}_{r'} f) \widetilde{\otimes}_{s'} f \right\|_{\mathfrak{H}}^2$$
$$\leq \frac{4\nu}{q!} \max\{c_q^2(r,s), c_q^2(r',s')\} \left\{ \|f \widetilde{\otimes}_r f\|_{\mathfrak{H}}^2 + \|f \widetilde{\otimes}_{r'} f\|_{\mathfrak{H}}^2 \right\}$$
$$\leq C \max_{\substack{1 \leq r \leq q-1 \\ r \neq q/2}} \|f \widetilde{\otimes}_r f\|_{\mathfrak{H}}^2. \quad (4.8)$$

Estimating S_2 and S_3: next, for every $1 \leq r' \leq q-1$ in the sum S_3, we split indices (r,s) with the property $3q - 2(r+s) = 2q - 2r'$ in the sum S_2 in two seperate parts: (i) $(r,s) = (q/2, r')$ and (ii) otherwise. Moreover, it holds that $c_q(q/2, r) = c_q(q/2)c_q(r)$. Hence, for every $1 \leq r' \leq q-1$ we get

$$\left\| c_q(q/2, r')(f \widetilde{\otimes}_{q/2} f) \widetilde{\otimes}_{r'} f - 2c_q(r')f \widetilde{\otimes}_{r'} f \right.$$
$$+ \sum_{\substack{r=1 \\ (r,s) \neq (q/2,q) \\ q/2 < r+s < 3q/2 \\ (r,s) \neq (q/2,1),\dots,(q/2,q-1)}}^{q-1} \sum_{s=1}^{[2q-2r]\wedge q} \left. c_q(r,s)(f \widetilde{\otimes}_r f) \widetilde{\otimes}_s f \right\|_{\mathfrak{H}}^2$$

$$\text{(note that when } s = q \text{ forces that } r \neq q/2)$$

$$\leq 2 \left\| c_q(q/2, r')(f \widetilde{\otimes}_{q/2} f) \widetilde{\otimes}_{r'} f - 2c_q(r')f \widetilde{\otimes}_{r'} f \right\|_{\mathfrak{H}}^2$$
$$+ 2 \sum_{\substack{r=1 \\ (r,s) \neq (q/2,q) \\ q/2 < r+s < 3q/2 \\ (r,s) \neq (q/2,r')}}^{q-1} \sum_{s=1}^{[2q-2r]\wedge q} c_q^2(r,s) \left\| (f \widetilde{\otimes}_r f) \widetilde{\otimes}_s f \right\|_{\mathfrak{H}}^2 \quad (4.9)$$

$$\leq \frac{4\nu}{q!} c_q^2(r) \left\| c_q(q/2)f \widetilde{\otimes}_{q/2} f - 2f \right\|_{\mathfrak{H}}^2$$
$$+ \frac{4\nu}{q!} \sum_{\substack{r=1 \\ (r,s) \neq (q/2,q) \\ q/2 < r+s < 3q/2 \\ r \neq q/2}}^{q-1} \sum_{s=1}^{[2q-2r]\wedge q} c_q^2(r,s) \left\| f \widetilde{\otimes}_r f \right\|_{\mathfrak{H}}^2$$

$$\leq C \left\{ \left\| c_q(q/2)f \widetilde{\otimes}_{q/2} f - 2f \right\|_{\mathfrak{H}}^2 + \max_{\substack{1 \leq r \leq q-1 \\ r \neq q/2}} \|f \widetilde{\otimes}_r f\|_{\mathfrak{H}}^2 \right\}$$

$$\leq C \max \left\{ \left\| c_q(q/2)f \widetilde{\otimes}_{q/2} f - 2f \right\|^2, \max_{\substack{1 \leq r \leq q-1 \\ r \neq q/2}} \|f \widetilde{\otimes}_r f\|_{\mathfrak{H}}^2 \right\}.$$

Now, the the variance estimate (4.4) follows directly from the relations (4.7), (4.8), (4.9), and (4.6). Finally, estimate (4.3) follows from Proposition 2.1, item (a) and,

the classical estimate (see Nourdin et al. (2010, Theorem 3.6))

$$\mathrm{Var}\left(\Gamma_1(F) - 2F\right) \leq \frac{q-1}{3q}\left\{\left|\kappa_4(F) - \kappa_4(G(\nu)) - 12\big(\kappa_3(F) - \kappa_3(G(\nu))\big)\right|\right\}.$$

$\square$

Remark 4.6. **(a):** The splitting technique implemented in the proof of Theorem 4.5 by using operator theory is vital to obtain an optimal upper bound. In fact, not doing it, instead of estimate (4.3), the best estimate one can achieve (under the assumption in Theorem 4.5) is a similar bound as (4.3) with the quantity $\sqrt{\mathrm{Var}\left(\Gamma_3(F) - 2\Gamma_2(F)\right)}$ instead of

$$\sqrt{\mathrm{Var}\left(\big(\Gamma_3(F) - 2\Gamma_2(F)\big) - 2\big(\Gamma_2(F) - 2\Gamma_1(F)\big)\right)}.$$

On the other hand, it is not difficult to see that for a sequence $\{F_n = \sum_{1 \leq i \leq \nu} c_{i,n}(N_i^2 - 1) : n \geq 1\}$ in the second Wiener chaos with a finite number of non-zero spectral coefficients such that for every $i = 1, \ldots, \nu$, $c_{i,n} \to 1$ as $n \to \infty$ it holds that

$$\mathrm{Var}\left(\Gamma_3(F_n) - 2\Gamma_2(F_n)\right) \approx_C \mathrm{Var}\left(\Gamma_2(F_n) - 2\Gamma_1(F_n)\right) \approx_C \mathrm{Var}\left(\Gamma_1(F_n) - 2F_n\right),$$

resulting in a suboptimal rate. See also illustrating Example 4.8 for further clarifications. Furthermore, only the special splitting with $\lambda = 2$ in Theorem 3.8 provides the optimal upper bound. This is the message of Proposition 4.7 estimate (4.10).

(b): Using Lemma 5.1 and orthogonality of Wiener chaoses (2.4), one can also further translate the right hand sides of estimates (4.2), and (4.3) in the language of norm of contraction operators. However, in order to keep our presentation transparent we avoid it due to presence of complicated contractions.

(c): When $q = 2$, i.e. the random variable F belongs to the second Wiener chaos, the estimate (4.4) can readily be deduced from representation (2.12) together with the classical estimate (4.4) in Biermé et al. (2012, Lemma 4.2).

4.2. *Proof of Theorem 1.3: Upper Bound.* In order to handle the involved variance quantity of the iterated Gamma operators appearing in the right hand side of estimate (4.3) in terms of cumulants, we consider the case of second Wiener chaos. In this setting, the connection is apparent thanks to Lemma 2.4.

Proposition 4.7. *Let $\nu > 0$, and $F = I_2(f)$ in the second Wiener chaos such that $\mathbb{E}[F^2] = 2\nu$. Assume $r \geq 1$. Then there exists a general constant C (possibly depending on the parameters ν and r but, independent of the kernel f) such that*

$$\mathrm{Var}\Big((\Gamma_{2r+1}(F) - 2\Gamma_{2r}(F)) - 2\left(\Gamma_{2r}(F) - 2\Gamma_{2r-1}(F)\right)\Big)$$
$$\leq 2\,\mathrm{Var}^2\left(\Gamma_r(F) - 2\Gamma_{r-1}(F)\right)$$
$$\leq_C \mathrm{Var}\left(\Gamma_{r-1}(F) - 2\Gamma_{r-2}(F)\right) \times \mathrm{Var}\left(\Gamma_{r+1}(F) - 2\Gamma_r(F)\right).$$

In particular, by choosing $r = 1$, we obtain the crucial estimate

$$\mathrm{Var}\Big((\Gamma_3(F) - 2\Gamma_2(F)) - 2\left(\Gamma_2(F) - 2\Gamma_1(F)\right)\Big) \leq 2\,\mathrm{Var}^2\left(\Gamma_1(F) - 2F\right). \quad (4.10)$$

Proof: For the first estimate, using representation (2.12) we can write

$$\mathrm{Var}\left(\left(\Gamma_{2r+1}(F) - 2\Gamma_{2r}(F)\right) - 2\left(\Gamma_{2r}(F) - 2\Gamma_{2r-1}(F)\right)\right)$$

$$= 2^{4r+3}\left\|f \otimes_1^{(2r+2)} f - 2f \otimes_1^{(2r+1)} f + f \otimes_1^{(2r)} f\right\|_{\mathfrak{H}^{\otimes 2}}^2$$

$$= 2^{4r+3}\left\|\left(f \otimes_1^{(r+1)} f - f \otimes_1^{(r)} f\right) \otimes_1 \left(f \otimes_1^{(r+1)} f - f \otimes_1^{(r)} f\right)\right\|_{\mathfrak{H}^{\otimes 2}}^2$$

$$\leq 2^{4r+3}\left\|f \otimes_1^{(r+1)} f - f \otimes_1^{(r)} f\right\|_{\mathfrak{H}^{\otimes 2}}^4$$

$$= 2\,\mathrm{Var}^2\left(\Gamma_r(F) - 2\Gamma_{r-1}(F)\right),$$

where we have used the classical estimate (4.4) in Biermé et al. (2012, Lemma 4.2). The second estimate is a direct application of Lemma 5.2 with $P = (A_f^{r+1} - A_f^r)^2$, $C = A_f^2$ combined with $\mathrm{Var}\left(\Gamma_{r+1}(F) - 2\Gamma_r(F)\right) = 2^{2r+3}\,\mathrm{Tr}\left((A_f^{r+2} - A_f^{r+1})^2\right)$ for every $r \geq 0$, see the proof of Lemma 2.4. $\qquad\square$

Proof for the upper bound of Theorem 1.3: combine Theorem 4.5 estimate (4.3) with Proposition 4.7 estimate (4.10) as well as Lemma 2.4 with $r = 1$.

4.3. *Proof of Theorem 1.3: Lower Bound.* Proof for the lower bound of Theorem 1.3: Fix a real number $\rho > 0$ whose range of values will be determined later on. Taking into account the second moment assumption, it is a classical result (see Lukacs, 1970, Chapter 7) that the characteristic functions ϕ_F and $\phi_{G(\nu)}$ are analytic inside the strip $\Delta_\nu := \{z \in \mathbb{C} : |\mathrm{Im}\, z| < \frac{1}{2\sqrt{\nu}}\}$. Moreover, in the strip of regularity Δ_ν, they follow the integral representations

$$\phi_F(z) = \int_{\mathbb{R}} e^{izx} \mu(dx) \quad \text{and} \quad \phi_{G(\nu)}(z) = \int_{\mathbb{R}} e^{izx} \mu_\nu(dx),$$

where μ and μ_ν stand for the probability measures of F and $G(\nu)$ respectively. Recall that all elements in the second Wiener chaos have exponential moments, see Nourdin and Peccati (2012, Proposition 2.7.13, item (iii)). Denote by $\Omega_{\rho,\nu}$ the domain

$$\Omega_{\rho,\nu} := \left\{z = t + iy \in \mathbb{C} : |\mathrm{Re}\, z| < \rho, |\mathrm{Im}\, z| < \min\{(2\sqrt{\nu})^{-1}, e^{-1}\}\right\}.$$

Then for any $z \in \Omega_{\rho,\nu}$, together with a Fubini's argument, we have that

$$\left|\phi_F(z) - \phi_{G(\nu)}(z)\right| = \left|\int_{\mathbb{R}} e^{itx - yx}(\mu - \mu_\nu)(dx)\right| = \left|\sum_{k\geq 0} \frac{(-y)^k}{k!} \int_{\mathbb{R}} x^k e^{itx}(\mu - \mu_\nu)(dx)\right|$$

$$\leq \sum_{k\geq 0} \frac{e^{-k}}{k!}\left|\phi_F^{(k)}(t) - \phi_{G(\nu)}^{(k)}(t)\right| \leq \sum_{k\geq 0} \frac{e^{-k}}{k!} \rho^{k+1} d_2(F, G(\nu))$$

$$= \rho\, e^{\rho e^{-1}} d_2(F, G(\nu)).$$

Hence $\left|\phi_F(z) - \phi_{G(\nu)}(z)\right| \leq_{C_\rho} d_2(F, G(\nu))$ for every $z \in \Omega_{\rho,\nu}$. Let $R > 0$ such that the disk $D_R \subset \mathbb{C}$ with the origin as center and radius R is contained in the domain $\Omega_{\rho,\nu}$ (note that R depends only on ν, since ρ is a free parameter. For example, one

can choose $\min\{(2\sqrt{\nu})^{-1}, e^{-1}\} < \rho < 2\min\{(2\sqrt{\nu})^{-1}, e^{-1}\}$). Now for any $z \in D_R$, and using the fact that

$$\frac{1}{\phi^2_{G(\nu)}(z)} = \left(e^{2iz}(1 - 2iz)\right)^\nu,$$

one can readily conclude that the function $\phi_{G(\nu)}(z)$ is bounded away from 0 on the disk D_R. Also, for any $r \geq 2$,

$$\left|\kappa_r(F)\right| \leq 2^{r-1}(r-1)! \sum_{i\geq 1}|c_i|^r \leq 2^{r-1}(r-1)! \max_i|c_i|^{r-2} \sum_{i\geq 1}|c_i|^2 \tag{4.11}$$

$$\leq 2^{r-2}(r-1)!\sqrt{\nu}^{r-2}\,\mathbb{E}(F^2) = 2^{r-2}(r-1)!\sqrt{\nu}^r.$$

Therefore, for any $z \in D_R$,

$$\left|\frac{1}{\phi_F(z)}\right| \leq \exp\left\{\sum_{r\geq 2}\frac{|\kappa_r(F)|}{r!}|z|^r\right\} \leq \exp\left\{\sum_{r\geq 2}\frac{2^{r-2}(r-1)!\sqrt{\nu}^r}{r!}|z|^r\right\}$$

$$\leq \exp\left\{\sum_{r\geq 2}\frac{2^{r-2}(r-1)!\sqrt{\nu}^r}{r!}R^r\right\} := C_{R,\nu} < \infty.$$

Hence the function $\phi_F(z)$ is also bounded away from 0 on the disk D_R. Also, relation (4.11) implies that the following power series (complex variable) converge to some analytic function as soon as $|z| < R$;

$$\sum_{r\geq 1}\frac{\kappa_r(F)}{r!}(iz)^r, \quad \sum_{r\geq 1}\frac{\kappa_r(G(\nu))}{r!}(iz)^r. \tag{4.12}$$

Thus we come to the conclusion that the functions $\phi_{G(\nu)}(z)$ and $\phi_F(z)$ are analytic on the disk D_R. Moreover, there exists a constant $c > 0$ such that $|\phi_{G(\nu)}(z)|, |\phi_F(z)| \geq c > 0$ for every $z \in D_R$. This implies that on the disk D_R there exist two analytic functions g and g_ν such that

$$\phi_F(z) = e^{g(z)}, \quad \phi_{G(\nu)}(z) = e^{g_\nu(z)},$$

i.e. $g(z) = \log(\phi_F(z))$ and $g_\nu(z) = \log(\phi_{G(\nu)}(z))$, for $z \in D_R$. In fact, the functions g and g_ν are given by the power series (4.12). Since the derivative of the analytic branch of the complex logarithm is $(\log z)' = \frac{1}{z}$ (see Conway, 1995, Corollary 2.21), one can infer that for some constant C whose value may differ from line to line and for every $z \in D_R$, we have

$$\left|\sum_{r\geq 2}\frac{\kappa_r(F) - \kappa_r(G(\nu))}{r!}(iz)^r\right| = \left|\log(\phi_F(z)) - \log(\phi_{G(\nu)}(z))\right|$$

$$\leq_C \left|\phi_F(z) - \phi_{G(\nu)}(z)\right| \leq_C d_2(F, G(\nu)).$$

Now, using Cauchy's estimate for the coefficients of analytic functions, for any $r \geq 3$, we obtain that

$$\left|\kappa_r(F) - \kappa_r(G(\nu))\right| \leq r!R^r \sup_{|z|\leq R}\left|\log\phi_F(z) - \log\phi_{G(\nu)}(z)\right|.$$

Therefore, $\max\left\{\left|\kappa_3(F) - \kappa_3(G(\nu))\right|, \left|\kappa_4(F) - \kappa_4(G(\nu))\right|\right\} \leq_C d_2(F, G(\nu)).$

4.4. *Examples.* We start with the following naive example that illustrates the essential role of our operator theory technique to achieve the optimal rate. It is worth mentioning that all the rates achieved in the forthcoming examples are better (by a square power) than those that can be obtained by the Malliavin-Stein bound (Nourdin and Peccati, 2009b, Theorem 1.5). In the following, when $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ are two non-negative real number sequences, we write $a_n \approx_C b_n$ if $\lim_{n\to\infty} \frac{a_n}{b_n} = C$, for some constant $C > 0$.

Example 4.8. Let $N_1, N_2 \sim \mathcal{N}(0,1)$ be independent. Consider the sequence

$$F_n = c_{1,n}(N_1^2 - 1) + c_{2,n}(N_2^2 - 1) := \sqrt{1 + \frac{1}{n}}(N_1^2 - 1) + \sqrt{1 - \frac{1}{n}}(N_2^2 - 1)$$

$$\xrightarrow{\mathcal{D}} G(2), \text{ as } n \to \infty.$$

First note that $\mathbb{E}[F_n^2] = 4$ for every $n \in \mathbb{N}$. Also, using Proposition 2.2 item 2, and relation (2.11), simple computations yield that $\kappa_4(F_n) - \kappa_4(G(2)) = 48\frac{2}{n^2} \approx_C \frac{1}{n^2}$. Similarly $\kappa_3(F_n) - \kappa_3(G(2)) = 8\sum_{j=1}^2 (c_{j,n}^3 - 1) \approx_C \frac{1}{n^2}$. Therefore, our main Theorem 1.3 implies

$$d_2(F_n, G(2)) \approx_C \max\left\{|\kappa_3(F_n) - \kappa_3(G(2))|, |\kappa_4(F_n) - \kappa_4(G(2))|\right\} \approx_C \frac{1}{n^2}. \quad (4.13)$$

The following important remarks are in order. (a) This example represents a typical scenario, in which, in order to obtain the optimal upper bound, one needs to join together two Gamma quantities $\Gamma_3(F_n) - 2\Gamma_2(F_n)$ and $\Gamma_2(F_n) - 2\Gamma_1(F_n)$. In fact, it is not difficult, using Lemma 2.4, to see that

$$\text{Var}\left(\Gamma_r(F_n) - 2\Gamma_{r-1}(F_n)\right) \approx_C \text{Var}\left(\Gamma_1(F_n) - 2F_n\right) \approx_C \frac{1}{n^2}, \forall r \geq 1.$$

And now consider Remark 4.6. (b) It is classical that the density function f_n of the random variable F_n admits the following explicit representation in terms of *confluent hypergeometric functions*,

$$f_n(x) = \frac{1}{2\sqrt{c_{1,n}c_{2,n}}} e^{-\frac{x+c_{1,n}+c_{2,n}}{2c_{1,n}}} \times {}_1F_1\left(\frac{1}{2}, 1, -\frac{c_{1,n} - c_{2,n}}{2c_{1,n}c_{2,n}}(c_{1,n} + c_{2,n} + x)\right)$$

$$\times \mathbb{1}_{\{x > -c_{1,n} - c_{2,n}\}}(x).$$

Also recall that the density of the target $G(2)$ is given by $f_\nu(x) = \frac{1}{2}e^{-\frac{x}{2}-1}\mathbb{1}_{\{x > -2\}}(x)$. Using rather long and tedious computations, one can show that the optimal estimate (4.13) continues to hold in the stronger distance of *total variation*, namely that

$$d_{TV}(F_n, G(2)) = \frac{1}{2}\int_{-\infty}^\infty |f_n(x) - f_\nu(x)|\, dx$$

$$\approx_C \max\left\{|\kappa_3(F_n) - \kappa_3(G(2))|, |\kappa_4(F_n) - \kappa_4(G(2))|\right\} \approx_C \frac{1}{n^2}.$$

In the next example we consider the important problem of the asymptotic behavior of the *least squares* estimators in the *autoregressive* models in the *nearly non-stationary* regime, where the target distribution $G(\nu)$ shows up. For more details on this fascinating subject, we refer the reader to Chan and Wei (1987, 1988); White (1958); Rao (1978); Buchmann and Chan (2013); Liu et al. (2011) and references therein when the noise is a martingale difference, and Buchmann and Chan (2007) when the innovation process exhibits *long-range* dependence. We also refer

to Götze and Tikhomirov (2005, Proposition 2) for a study of optimal rates in a general context of quadratic forms.

Example 4.9. (Least square estimator in nearly non stationary $AR(1)$ model) Let $n \in \mathbb{N}$. Let $\beta_n := 1 - \frac{\beta}{n}$. We consider the first order autoregressive process $X_t(n) = \beta_n X_{t-1}(n) + Z_t$, where $t = 1, \ldots, n$, $X_0(n) = 0$ for all n and (Z_i) is a white noise, i.e. a sequence of i.i.d. $\mathcal{N}(0, 1)$ random variables. It is classical that the least squares estimator of the unknown parameter β_n, based on discrete observations $X_1(n), \ldots, X_n(n)$, is given by

$$\widehat{\beta}_n = \frac{\sum_{t=1}^n X_{t-1}(n) X_t(n)}{\sum_{t=1}^n X_{t-1}^2(n)}.$$

Define

$$W_n^\beta := \frac{2}{\sqrt{n(n-1)}} \left(\sum_{t=1}^n X_{t-1}^2(n) \right) (\widehat{\beta}_n - \beta_n) = \frac{2}{\sqrt{n(n-1)}} \sum_{i=1}^n \sum_{j=1}^{i-1} \beta_n^{i-j} Z_i Z_j.$$

Then Chan and Wei (1987, Theorem 1) implies that as $n \to \infty$:

$$W_n^\beta \xrightarrow{\mathcal{D}} W_\infty^\beta := 2 \int_0^1 \left(1 + t(e^{2\beta} - 1) \right)^{-1} B_t dB_t,$$

where $B = (B_t)_{t \in [0,1]}$ is a standard Brownian motion. In particular when $\beta = 0$, we observe that $W_\infty := W_\infty^{\beta=0} = G(1)$ (equality in law), and hence we obtain that $W_n := W_n^{\beta=0} \xrightarrow{\mathcal{D}} G(1)$. Now, let $\{h_i\}_{i \geq 1}$ be an orthonormal basis of $\mathfrak{H}$ and for $i \geq 1$ set $Z_i := I_1(h_i)$. Note that

$$W_n = I_2 \left(\frac{2}{\sqrt{n(n-1)}} \sum_{1 \leq i < j \leq n} h_i \widetilde{\otimes} h_j \right) =: I_2(f_n).$$

We consider the associated Hilbert-Schmidt operator $A_{f_n} g = f_n \otimes_1 g$. Using the fact that $(h_i \otimes h_j) \otimes_1 h_k = \langle h_i, h_k \rangle_{\mathfrak{H}} h_j$ we can explicitly compute the non-zero eigenvalues $c_{1,n}, \ldots, c_{n,n}$ of A_{f_n}. They are

$$c_{1,n} = \sqrt{\frac{n-1}{n}}, \text{ and } c_{2,n} = \ldots = c_{n,n} = \frac{-1}{\sqrt{n(n-1)}}. \tag{4.14}$$

Therefore, as $n \to \infty$, gathering Proposition 2.2 item 2, relation (4.14) and Theorem 1.3 we get that

$$d_2 \big(W_n, G(1) \big) \approx_C \big| \kappa_3(W_n) - \kappa_3(G(1)) \big| \approx_C \big| \kappa_4(W_n) - \kappa_4(G(1)) \big| \approx_C \frac{1}{n}.$$

Example 4.10. (Least square estimator in $AR(2)$ model) In this example, we consider the second order autoregressive $AR(2)$ model:

$$X_n = \beta_1 X_{n-1} + \beta_2 X_{n-2} + Z_n, \tag{4.15}$$

where (Z_k) is a white noise, and $X_0 = X_{-1} = 0$. Further, assume that the roots of the associated characteristic polynomial $1 - \beta_1 z - \beta_2 z^2$ are $e^{i\theta}$ and $e^{-i\theta}$, and lie on the unit disk. Under this condition it is easy to see that $\beta_1 = 2 \cos \theta$ and $\beta_2 = -1$. The least square estimator $\widehat{\boldsymbol{\beta}}_n = (\widehat{\beta}_{1,n}, \widehat{\beta}_{2,n})'$ of the parameter $\boldsymbol{\beta} = (\beta_1, \beta_2)' = (2 \cos \theta, -1)'$ for $n \geq 2$ is given by

$$\widehat{\boldsymbol{\beta}}_n = \left(\sum_{k=0}^{n-1} \mathbf{X_k X_k}' \right)^{-1} \sum_{k=1}^n \mathbf{X_{k-1} X_k}, \text{ where } \mathbf{X_k} = (X_k, X_{k-1})'.$$

In Chan and Wei (1988), the asymptotic behavior of $n(\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) = \mathbf{A}_n^{-1}\boldsymbol{b}_n$ has been derived where

$$\mathbf{A}_n = \frac{1}{n^2}\sum_{k=2}^{n}\begin{bmatrix} X_{k-1}^2 & X_{k-1}X_{k-2} \\ X_{k-1}X_{k-2} & X_{k-2}^2 \end{bmatrix}, \text{ and } \boldsymbol{b}_n = \begin{bmatrix} b_{1,n} \\ b_{2,n} \end{bmatrix} := \frac{1}{n}\sum_{k=2}^{n}\begin{bmatrix} X_{k-1}Z_k \\ X_{k-2}Z_k \end{bmatrix}.$$

Following Chan and Wei (1988, Corollary 3.3.8), as $n \to \infty$, one can deduce that

$$W_n^\theta := 4(\cos\theta\, b_{1,n} - b_{2,n}) = 4\left((\cos\theta)\frac{1}{n}\sum_{k=1}^{n}X_{k-1}Z_k - \frac{1}{n}\sum_{k=1}^{n}X_{k-2}Z_k\right) \xrightarrow{\mathcal{D}} G(2).$$

Note that the sequence $(W_n^\theta : n \geq 1)$ belongs to the second Wiener chaos. An interesting feature of the previous limit theorem is that although the sequence does depend on the parameter θ in the model, the target distribution is independent of θ. On the other hand, relation (4.15) together with the assumption $(\beta_1, \beta_2) = (2\cos\theta, -1)$ yields that

$$X_k = \sum_{j=1}^{k}\frac{\sin(k - j + 1)\theta}{\sin\theta}Z_j.$$

Therefore,

$$W_n^\theta = \frac{4}{n}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{\cos\theta\sin(i - j)\theta - \sin(i - j - 1)\theta}{\sin\theta}Z_i Z_j.$$

By elementary combinatorics, we have for any function $f : \mathbb{N} \to \mathbb{R}$ that $\sum_{i=2}^{n}\sum_{j=1}^{i-1}f(i - j) = \sum_{k=1}^{n-1}(n - k)f(k)$. Using this, and evaluating the sums of sine functions (which are just geometric sums after writing them in terms of complex exponentials), we get

$$\mathbb{E}[(W_n^\theta)^2]$$
$$= \frac{16}{n^2}\left\{\frac{1}{8\sin^4\theta}\left(\cos\left(2\theta(n - 1)\right) - 2\cos(\theta)\cos\left(\theta(2n - 1)\right) + \cos^2(\theta)\cos(2n\theta)\right)\right.$$
$$\left. + \frac{1}{8\sin^2(\theta)}\left(n\cos(2\theta) + 1 - n\right) + \frac{n(n - 1)}{4}\right\}. \tag{4.16}$$

Note that $\left|\kappa_2(W_n^\theta) - 4\right| \approx_C 1/n$ as $n \to \infty$. Now we scale W_n^θ so that it has variance equal to 4 for every $n \in \mathbb{N}$. Set $\sigma_n := \sqrt{\mathrm{Var}(W_n^\theta)}$, and let $\widetilde{W}_n^\theta := \frac{2}{\sigma_n}W_n^\theta$. Using (2.10), and after some tedious computations, we get that

$$\kappa_3(\widetilde{W}_n^\theta) = \frac{512}{\sigma_n^3 n^3}\left(-\frac{3n\cos^2(n\theta)}{4\sin^2(\theta)} + \frac{3\cos^2(n\theta)}{2\sin^2(\theta)} + \frac{5n\cos^4(\theta)}{4\sin^4(\theta)} + \frac{13n}{4\sin^2(\theta)} - \frac{3}{2\sin^2(\theta)}\right.$$
$$\left. - \frac{5n}{4\sin^4(\theta)} + \frac{n^3}{4} - \frac{3n^2}{2} + \frac{3n}{4}\right).$$

Using that $\sigma_n^3 \to 8$ as $n \to \infty$, we see that $\lim_{n \to \infty}\kappa_3(\widetilde{W}_n^\theta) = 16 = 8\nu$ (note that $\nu = 2$), and furthermore,

$$\left|\kappa_3(\widetilde{W}_n^\theta) - \kappa_3(G(2))\right| \approx_C \frac{1}{n}.$$

Similar computations yield that $|\kappa_4(\widetilde{W}_n^\theta) - \kappa_4(G(2))| \approx_C 1/n$. Therefore, Theorem 1.3 can be applied to deduce that $d_2(\widetilde{W}_n^\theta, G(2)) \approx_C 1/n$.

Example 4.11. (Quadratic forms of de Wet and Venter, 1973 and Arras et al., 2019, section 3.2) In this example, we consider a general quadratic form in independent standard normal random variables

$$F_n := \sum_{1 \leq i,j \leq n} c_n(i,j) Z_i Z_j, \quad n \in \mathbb{N},$$

where $C_n = (c_n(i,j))_{1 \leq i,j \leq n}$ is an $n \times n$ symmetric matrix, and (Z_i) is a sequence of i.i.d standard normal random variables. Let $\nu > 0$ be an integer number. Now, we make the following assumptions:

(a) The second moment assumption: $\sum_{1 \leq i,j \leq n} c_n(i,j)^2 = \nu, \quad \forall n \in \mathbb{N}$.
(b) There exists a sequence $\{b_n^m(i) : n, i = 1, 2, \ldots, m = 1, 2, \ldots, \nu\}$ of real numbers such that as $n \to \infty$:

$$\sum_{1 \leq i \leq n} b_n^m(i) b_n^k(i) \to \delta_{km}, \text{ and } \exists b > 0, \ \forall i, m, n, \ \sqrt{n} \mid b_n^m(i) \mid \leq b < +\infty.$$

(c) For every $1 \leq m \leq \nu$, as $n \to \infty$ it holds that: $\sum_{1 \leq i,j \leq n} c_n(i,j) b_n^m(i) b_n^m(j) \to 1$.

Now a direct application of de Wet and Venter (1973, Theorem 2) implies that $W_n := F_n - \mathbb{E}[F_n] \overset{\mathcal{D}}{\to} G(\nu)$. Note that $\mathbb{E}[W_n^2] = 2\nu$ for every $n \in \mathbb{N}$ relying on condition (a). Moreover, one can write $W_n = I_2(\sum_{1 \leq i,j \leq n} c_n(i,j) h_i \widetilde{\otimes} h_j)$, where $\{h_i\}_{i \geq 1}$ stands for an orthonormal basis of $\mathfrak{H}$, and for $i \geq 1$, as before, we set $Z_i := I_1(h_i)$. Therefore our main Theorem 1.3 entails that

$$d_2(W_n, G(\nu)) \approx_C \max\left\{ \left| \kappa_3(W_n) - \kappa_3(G(\nu)) \right|, \left| \kappa_4(W_n) - \kappa_4(G(\nu)) \right| \right\}. \quad (4.17)$$

Depending on the particular choice of the matrix C_n in the original quadratic form F_n, we can provide explicit rates (in terms of suitable powers of n) in the asymptotic relation (4.17). For example, following de Wet and Venter (1973, remark after Theorem 2) and Arras et al. (2019, Corollary 3.2), assume that $\{e_m : m = 1, \ldots, \nu\}$ is a sequence of distinct orthonormal functions in $L^2[0,1]$ such that $e_m \in C^\alpha([0,1])$ for some $\alpha \in (0,1]$. Here $C^\alpha([0,1])$ denotes the space of all Hölder continuous functions with Hölder exponent α. Consider the square integrable kernel K_ν defined as

$$K_\nu(x,y) = \sum_{1 \leq m \leq \nu} e_m(x) e_m(y), \ \forall (x,y) \in (0,1)^2.$$

Finally, for $n \in \mathbb{N}$ and $1 \leq i,j \leq n$ we set

$$d_n(i,j) := \frac{1}{n} K_\nu\left(\frac{i}{n}, \frac{j}{n}\right), \quad \text{and} \quad c_n(i,j) := \sqrt{\frac{\nu}{\sum_{1 \leq i,j \leq n} d_n^2(i,j)}} \times d_n(i,j).$$

Now consider the sequence $W_n = F_n - \mathbb{E}[F_n]$ associated to the symmetric matrix $C_n = (c_n(i,j))$ belonging to the second Wiener chaos. Then, it is straightforward to check that the conditions (a)-(b)-(c) are in order with $b_n^m(i) = \frac{e_m(i/n)}{\sqrt{n}}$. On the other hand, it has been shown Arras et al. (2019, Corollary 3.2) that:

$$\left| \kappa_r(W_n) - \kappa_r(G(\nu)) \right| \approx_C n^{-\alpha}, \ \forall r \geq 2. \quad (4.18)$$

Putting together the asymptotic estimates (4.17) and (4.18), we obtain the optimal rate $d_2(W_n, G(\nu)) \approx_C n^{-\alpha}$. Also, the example presented on page 107 in Nourdin and Peccati (2009b) can be treated in this framework, and resulting in an improved optimal rate of $1/n$.

5. Appendix

The following lemma provides an explicit representation of the new Gamma operators used in this paper in terms of contractions. Recall that these are not the same as e.g. in Nourdin and Peccati (2010), but rather the new ones introduced in (2.7).

Lemma 5.1. *For $q \geq 1$, let $F = I_q(f)$, for some $f \in \mathfrak{H}^{\odot q}$ be an element of the q-th Wiener chaos. Then*

$$
\Gamma_s(F) = \sum_{r_1=1}^{q} \cdots \sum_{r_s=1}^{[sq-2r_1-\cdots-2r_{s-1}]\wedge q} c_q(r_1,\ldots,r_s) \mathbb{1}_{\{r_1 < q\}} \cdots \mathbb{1}_{\{r_1+\cdots+r_{s-1} < \frac{sq}{2}\}}
$$
$$
\times I_{(s+1)q-2r_1-\cdots-2r_s} \left(\left(\left(\ldots(f \widetilde{\otimes}_{r_1} f) \widetilde{\otimes}_{r_2} f\right)\ldots f\right) \widetilde{\otimes}_{r_s} f\right), \tag{5.1}
$$

where the constants $c_q(r_1, \cdots, r_s)$ are recursively defined via $c_q(r) = q\,(r-1)! \binom{q-1}{r-1}^2$, and for $s \geq 2$,

$$
c_q(r_1, \cdots, r_s) =
$$
$$
(sq - 2r_1 - \cdots - 2r_{s-1})\,(r_s - 1)!
$$
$$
\times \binom{sq - 2r_1 - \cdots - 2r_{s-1} - 1}{r_s - 1}\binom{q-1}{r_s - 1} c_q(r_1, \cdots, r_{s-1}). \tag{5.2}
$$

Proof: It follows by induction on s and similar lines of arguments as in Nourdin and Peccati (2010, Proof of Theorem 5.1).

$\square$

Proof of Proposition 2.1: Part *(a)* is clear from the definition. Part *(b)* for $j = 1$ is also trivial. For $j = 2$, we use the fact that $\Gamma_1 = \Gamma_1^{alt}$, as well as the integration by parts formula (2.6), to get

$$
\mathbb{E}[\Gamma_2(F)] = \mathbb{E}[\langle D\Gamma_1(F), -DL^{-1}F\rangle_{\mathfrak{H}}] = \mathbb{E}[\Gamma_1(F)\,F]
$$
$$
= \mathbb{E}[F\,\Gamma_1^{alt}(F)] = \mathbb{E}[\langle DF, -DL^{-1}\Gamma_1^{alt}(F)\rangle_{\mathfrak{H}}] = \mathbb{E}[\Gamma_2^{alt}(F)].
$$

For part *(c)*, consider

$$
\mathbb{E}[\Gamma_3(F)] = \mathbb{E}[\langle D\Gamma_2(F), -DL^{-1}F\rangle_{\mathfrak{H}}] = \mathbb{E}[F\,\Gamma_2(F)]
$$
$$
= \mathbb{E}[F\,\langle D\Gamma_1(F), -DL^{-1}F\rangle_{\mathfrak{H}}]
$$
$$
= \mathbb{E}[\langle D(F\,\Gamma_1(F)), -DL^{-1}F\rangle_{\mathfrak{H}}] - \mathbb{E}[\Gamma_1(F)\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}]
$$
$$
= \mathbb{E}[\langle D(F\,\Gamma_1(F)), -DL^{-1}F\rangle_{\mathfrak{H}}] - \mathbb{E}[\Gamma_1^{alt}(F)^2]
$$
$$
= \mathbb{E}[F^2\,\Gamma_1^{alt}(F)] - \mathbb{E}[\Gamma_1^{alt}(F)^2]
$$
$$
= \mathbb{E}[F^2]\,\mathbb{E}[\Gamma_1^{alt}(F)] + \mathbb{E}[2F\,\langle DF, -DL^{-1}\Gamma_1^{alt}(F)\rangle_{\mathfrak{H}}] - \mathbb{E}[\Gamma_1^{alt}(F)^2]
$$
$$
= \mathbb{E}[\Gamma_1^{alt}(F)]^2 + 2\,\mathbb{E}[F\,\Gamma_2^{alt}(F)] - \mathbb{E}[\Gamma_1^{alt}(F)^2]
$$
$$
= -\operatorname{Var}(\Gamma_1^{alt}(F)) + 2\,\mathbb{E}[\Gamma_3^{alt}(F)].
$$

For part *(d)*, we follow an induction argument on j. For $j = 1$, this is part *(a)*. Now, let $j \geq 1$, then using induction hypothesis, $-L^{-1}(F) = \frac{1}{2}F$ for F being in the second Wiener chaos, and the fact that $\overline{\Gamma}_j^{alt}(F) = \Gamma_j^{alt}(F) - \mathbb{E}\left[\Gamma_j^{alt}(F)\right]$ belongs to the second Wiener chaos (see Azmoodeh et al., 2015 equation (24)), we can write

$$\Gamma_{j+1}(F) = \langle D\Gamma_j(F), -DL^{-1}(F)\rangle_{\mathfrak{H}} = \frac{1}{2}\langle DF, D\Gamma_j^{alt}(F)\rangle_{\mathfrak{H}} = \frac{1}{2}\langle DF, D\overline{\Gamma}_j^{alt}(F)\rangle_{\mathfrak{H}}$$
$$= \langle DF, -DL^{-1}\Gamma_j^{alt}(F)\rangle_{\mathfrak{H}} = \Gamma_{j+1}^{alt}(F).$$

□

Lemma 5.2 (Dragomir, 2016, Corollary 1)**.** *Let H be a Hilbert space. Assume that $C \in \mathcal{B}(H)$ is a bounded, self-adjoint operator, and $P \in \mathcal{B}^+(H)$ is a positive, trace-class operator on the Hilbert space H. Then,*

$$\left|\frac{\mathrm{Tr}(PC)}{\mathrm{Tr}(P)}\right|^2 \leq \frac{\mathrm{Tr}(PC^2)}{\mathrm{Tr}(P)}.$$

Acknowledgements

References

Arras, B., Azmoodeh, E., Poly, G., and Swan, Y. A bound on the Wasserstein-2 distance between linear combinations of independent random variables. *Stochastic Process. Appl.*, **129** (7), 2341–2375 (2019). MR3958435.

Arras, B., Mijoule, G., Poly, G., and Swan, Y. A new approach to the Stein-Tikhomirov method: with applications to the second Wiener chaos and Dickman convergence. *ArXiv Mathematics e-prints* (2017). arXiv: 1605.06819v2.

Arras, B. and Swan, Y. A stroll along the gamma. *Stochastic Process. Appl.*, **127** (11), 3661–3688 (2017). MR3707241.

Azmoodeh, E., Campese, S., and Poly, G. Fourth Moment Theorems for Markov diffusion generators. *J. Funct. Anal.*, **266** (4), 2341–2359 (2014). MR3150163.

Azmoodeh, E., Eichelsbacher, P., and Knichel, L. On the Rate of Convergence to a Gamma Distribution on Wiener Space. *ArXiv Mathematics e-prints* (2018). arXiv: 1806.03878v2.

Azmoodeh, E., Malicet, D., Mijoule, G., and Poly, G. Generalization of the Nualart-Peccati criterion. *Ann. Probab.*, **44** (2), 924–954 (2016). MR3474463.

Azmoodeh, E., Peccati, G., and Poly, G. Convergence towards linear combinations of chi-squared random variables: a Malliavin-based approach. In *In memoriam Marc Yor—Séminaire de Probabilités XLVII*, volume 2137 of *Lecture Notes in Math.*, pp. 339–367. Springer, Cham (2015). MR3444306.

Biermé, H., Bonami, A., Nourdin, I., and 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.

Buchmann, B. and Chan, N. H. Asymptotic theory of least squares estimators for nearly unstable processes under strong dependence. *Ann. Statist.*, **35** (5), 2001–2017 (2007). MR2363961.

Buchmann, B. and Chan, N. H. Unified asymptotic theory for nearly unstable AR(p) processes. *Stochastic Process. Appl.*, **123** (3), 952–985 (2013). MR3005011.

Chan, N. H. and Wei, C. Z. Asymptotic inference for nearly nonstationary $AR(1)$ processes. *Ann. Statist.*, **15** (3), 1050–1063 (1987). MR902245.

Chan, N. H. and Wei, C. Z. Limiting distributions of least squares estimates of unstable autoregressive processes. *Ann. Statist.*, **16** (1), 367–401 (1988). MR924877.

Chen, L. H. Y., Goldstein, L., and Shao, Q.-M. *Normal approximation by Stein's method.* Probability and its Applications (New York). Springer, Heidelberg (2011). ISBN 978-3-642-15006-7. MR2732624.

Conway, J. B. *Functions of one complex variable. II*, volume 159 of *Graduate Texts in Mathematics.* Springer-Verlag, New York (1995). ISBN 0-387-94460-5. MR1344449.

de Wet, T. and Venter, J. H. Asymptotic distributions for quadratic forms with applications to tests of fit. *Ann. Statist.*, **1**, 380–387 (1973). MR0353543.

Döbler, C. Stein's method of exchangeable pairs for the beta distribution and generalizations. *Electron. J. Probab.*, **20**, no. 109, 34 (2015). MR3418541.

Döbler, C. and Peccati, G. The gamma Stein equation and noncentral de Jong theorems. *Bernoulli*, **24** (4B), 3384–3421 (2018). MR3788176.

Dragomir, S. S. Some Grüss' type inequalities for trace of operators in Hilbert spaces. *Oper. Matrices*, **10** (4), 923–943 (2016). MR3584684.

Eden, R. and Víquez, J. Nourdin-Peccati analysis on Wiener and Wiener-Poisson space for general distributions. *Stochastic Process. Appl.*, **125** (1), 182–216 (2015). MR3274696.

Eichelsbacher, P. and Thäle, C. Malliavin-Stein method for variance-gamma approximation on Wiener space. *Electron. J. Probab.*, **20**, Paper No. 123, 28 (2015). MR3425543.

Gaunt, R. E., Pickett, A. M., and Reinert, G. Chi-square approximation by Stein's method with application to Pearson's statistic. *Ann. Appl. Probab.*, **27** (2), 720–756 (2017). MR3655852.

Godefroy, G. A survey on Lipschitz-free Banach spaces. *Comment. Math.*, **55** (2), 89–118 (2015). MR3518958.

Götze, F. and Tikhomirov, A. N. Asymptotic expansions in non-central limit theorems for quadratic forms. *J. Theoret. Probab.*, **18** (4), 757–811 (2005). MR2289931.

Kusuoka, S. and Tudor, C. A. Stein's method for invariant measures of diffusions via Malliavin calculus. *Stochastic Process. Appl.*, **122** (4), 1627–1651 (2012). MR2914766.

Kusuoka, S. and Tudor, C. A. Characterization of the convergence in total variation and extension of the fourth moment theorem to invariant measures of diffusions. *Bernoulli*, **24** (2), 1463–1496 (2018). MR3706799.

Ledoux, M. Chaos of a Markov operator and the fourth moment condition. *Ann. Probab.*, **40** (6), 2439–2459 (2012). MR3050508.

Liu, W., Ling, S., and Shao, Q.-M. On non-stationary threshold autoregressive models. *Bernoulli*, **17** (3), 969–986 (2011). MR2817613.

Luk, H. M. *Stein's method for the Gamma distribution and related statistical ap-plications.* ProQuest LLC, Ann Arbor, MI (1994). Thesis (Ph.D.)–University of Southern California. MR2693204.

Lukacs, E. *Characteristic functions.* Hafner Publishing Co., New York (1970). MR0346874.

Megginson, R. E. *An introduction to Banach space theory*, volume 183 of *Graduate Texts in Mathematics.* Springer-Verlag, New York (1998). ISBN 0-387-98431-3. MR1650235.

Nourdin, I. and Peccati, G. Noncentral convergence of multiple integrals. *Ann. Probab.*, **37** (4), 1412–1426 (2009a). MR2546749.

Nourdin, I. and Peccati, G. Stein's method on Wiener chaos. *Probab. Theory Related Fields*, **145** (1-2), 75–118 (2009b). MR2520122.

Nourdin, I. and Peccati, G. Cumulants on the Wiener space. *J. Funct. Anal.*, **258** (11), 3775–3791 (2010). MR2606872.

Nourdin, I. and Peccati, G. *Normal approximations with Malliavin calculus. From Stein's method to universality*, volume 192 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge (2012). ISBN 978-1-107-01777-1. MR2962301.

Nourdin, I. and Peccati, G. The optimal fourth moment theorem. *Proc. Amer. Math. Soc.*, **143** (7), 3123–3133 (2015). MR3336636.

Nourdin, I., Peccati, G., and Reinert, G. Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. *Ann. Probab.*, **38** (5), 1947–1985 (2010). MR2722791.

Nourdin, I. and Poly, G. Convergence in law in the second Wiener/Wigner chaos. *Electron. Commun. Probab.*, **17**, no. 36, 12 (2012). MR2970700.

Nourdin, I. and Rosiński, J. Asymptotic independence of multiple Wiener-Itô integrals and the resulting limit laws. *Ann. Probab.*, **42** (2), 497–526 (2014). MR3178465.

Nualart, D. *The Malliavin calculus and related topics.* Probability and its Applications (New York). Springer-Verlag, Berlin, second edition (2006). ISBN 978-3-540-28328-7; 3-540-28328-5. MR2200233.

Nualart, D. and Nualart, E. *Introduction to Malliavin calculus*, volume 9 of *Institute of Mathematical Statistics Textbooks.* Cambridge University Press, Cambridge (2018). ISBN 978-1-107-61198-6; 978-1-107-03912-4. MR3838464.

Nualart, D. and Peccati, G. Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.*, **33** (1), 177–193 (2005). MR2118863.

Pickett, A. *Rates of convergence of Chi-square approximations via Stein's method.* Ph.D. thesis, University of Oxford (2004).

Rao, M. M. Asymptotic distribution of an estimator of the boundary parameter of an unstable process. *Ann. Statist.*, **6** (1), 185–190 (1978). MR0458676.

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 (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pp. 583–602 (1972). MR0402873.

Weaver, N. *Lipschitz algebras.* World Scientific Publishing Co., Inc., River Edge, NJ (1999). ISBN 981-02-3873-8. MR1832645.

Weaver, N. On the unique predual problem for Lipschitz spaces. *Math. Proc. Cambridge Philos. Soc.*, **165** (3), 467–473 (2018). MR3860399.

White, J. S. The limiting distribution of the serial correlation coefficient in the explosive case. *Ann. Math. Statist.*, **29**, 1188–1197 (1958). MR0100952.