Available online at www.sciencedirect.com

ScienceDirect

Journal of Functional Analysis 266 (2014) 814–875

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Convergence of densities of some functionals of Gaussian processes

Yaozhong Hu [a,1], Fei Lu [b,*], David Nualart [a,2]

[a] *Department of Mathematics, University of Kansas, Lawrence, KS, 66045, USA*
[b] *Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA*

Received 1 March 2013; accepted 25 September 2013

Available online 14 October 2013

Communicated by B. Driver

Abstract

The aim of this paper is to establish the uniform convergence of the densities of a sequence of random variables, which are functionals of an underlying Gaussian process, to a normal density. Precise estimates for the uniform distance are derived by using the techniques of Malliavin calculus, combined with Stein's method for normal approximation. We need to assume some non-degeneracy conditions. First, the study is focused on random variables in a fixed Wiener chaos, and later, the results are extended to the uniform convergence of the derivatives of the densities and to the case of random vectors in some fixed chaos, which are uniformly non-degenerate in the sense of Malliavin calculus. Explicit upper bounds for the uniform norm are obtained for random variables in the second Wiener chaos, and an application to the convergence of densities of the least square estimator for the drift parameter in Ornstein–Uhlenbeck processes is discussed.
Published by Elsevier Inc.

Keywords: Multiple Wiener–Itô integrals; Wiener chaos; Malliavin calculus; Integration by parts; Stein's method; Convergence of densities; Ornstein–Uhlenbeck process; Least squares estimator; Small deviation

Contents

* Corresponding author.
E-mail addresses: hu@math.ku.edu (Y. Hu), flu@lbl.gov (F. Lu), nualart@math.ku.edu (D. Nualart).

[1] Y. Hu is partially supported by a grant from the Simons Foundation #209206.
[2] D. Nualart is supported by the NSF grant DMS1208625.

0022-1236/$ – see front matter Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.jfa.2013.09.024

1. Introduction

There has been a recent interest in studying normal approximations for sequences of multiple stochastic integrals. Consider a sequence of multiple stochastic integrals of order $q \geqslant 2$, $F_n = I_q(f_n)$, with variance $\sigma^2 > 0$, with respect to an isonormal Gaussian process $X = \{X(h), h \in \mathfrak{H}\}$ associated with a Hilbert space $\mathfrak{H}$. It was proved by Nualart and Peccati [24] and Nualart and Ortiz-Latorre [23] that F_n converges in distribution to the normal law $N(0, \sigma^2)$ as $n \to \infty$ if and only if one of the following three equivalent conditions holds:

(i) $\lim_{n\to\infty} E[F_n^4] = 3\sigma^4$ (convergence of the fourth moments).
(ii) For all $1 \leqslant r \leqslant q-1$, $f_n \otimes_r f_n$ converges to zero, where $\otimes_r$ denotes the contraction of order r (see Eq. (2.5)).
(iii) $\|DF_n\|_{\mathfrak{H}}^2$ (see definition in Section 2) converges to $q\sigma^2$ in $L^2(\Omega)$ as n tends to infinity.

A new methodology to study normal approximations and to derive quantitative results combining Stein's method with Malliavin calculus was introduced by Nourdin and Peccati [15] (see also Nourdin and Peccati [16]). As an illustration of the power of this method, let us mention the following estimate for the total variation distance between the law $\mathcal{L}(F)$ of $F = I_q(f)$ and distribution $\gamma = N(0, \sigma^2)$, where $\sigma^2 = E[F^2]$:

$$d_{TV}\big(\mathcal{L}(F), \gamma\big) \leqslant \frac{2}{q\sigma^2}\sqrt{\operatorname{Var}\big(\|DF\|_{\mathfrak{H}}^2\big)} \leqslant \frac{2\sqrt{q-1}}{\sigma^2\sqrt{3q}}\sqrt{E\big[F^4\big] - 3\sigma^4}. \tag{1.1}$$

This inequality can be used to show the above equivalence (i)–(iii). A recent result of Nourdin and Poly [21] says that the convergence in law for a sequence of multiple stochastic integrals of order $q \geqslant 2$ is equivalent to the convergence in total variation if the limit is not constant. As a consequence, for a sequence F_n of nonzero multiple stochastic integrals of order $q \geqslant 2$, the limit in law to is equivalent to the limit of the densities in $L^1(\mathbb{R})$, provided the limit is not constant. A multivariate extension of this result has been derived in [14].

The aim of this paper is to study the uniform convergence of the densities of a sequence of random vectors F_n to the normal density using the techniques of Malliavin calculus, combined with Stein's method for normal approximation. It is well known that to guarantee that each F_n has a density we need to assume that the norm of the Malliavin derivative of F_n has negative moments. Thus, a natural assumption to obtain uniform convergence of densities is to assume uniform boundedness of the negative moments of the corresponding Malliavin derivatives. Our first result (Theorem 4.1) says that if F is a multiple stochastic integral of order $q \geqslant 2$ such that $E[F^2] = \sigma^2$ and $M := E(\|DF\|_{\mathfrak{H}}^{-6}) < \infty$, we have

$$\sup_{x\in\mathbb{R}} \left| f_F(x) - \phi(x) \right| \leqslant C\sqrt{E\left[F^4\right] - 3\sigma^4}, \tag{1.2}$$

where f_F is the density of F, ϕ is the density of the normal law $N(0, \sigma^2)$ and the constant C depends on q, σ and M. We can also replace the expression in the right-hand side of (1.2) by $\sqrt{\operatorname{Var}(\|DF\|_{\mathfrak{H}}^2)}$. The main idea to prove this result is to express the density of F using Malliavin calculus:

$$f_F(x) = E\left[\mathbf{1}_{\{F>x\}} q \|DF\|_{\mathfrak{H}}^{-2} F\right] - E\left[\mathbf{1}_{\{F>x\}} \left\langle DF, D\left(\|DF\|_{\mathfrak{H}}^{-2}\right)\right\rangle_{\mathfrak{H}}\right].$$

Then, one can find an estimate of the form (1.2) for the terms $E[|\langle DF, D(\|DF\|_{\mathfrak{H}}^{-2})\rangle_{\mathfrak{H}}|]$ and $E[|q\|DF\|_{\mathfrak{H}}^{-2} - \sigma^{-2}|]$. On the other hand, taking into account that

$$\phi(x) = \sigma^{-2} E[\mathbf{1}_{\{N>x\}} N],$$

it suffices to estimate the difference

$$E[\mathbf{1}_{\{F>x\}} F] - E[\mathbf{1}_{\{N>x\}} N],$$

which can be done by Stein's method. The estimate (1.2) leads to the uniform convergence of the densities in the above equivalence of conditions (i) to (iii) if we assume that $\sup_n E(\|DF_n\|_{\mathfrak{H}}^{-6}) < \infty$.

This methodology is extended in the paper in several directions. We consider the uniform approximation of the mth derivative of the density of F by the corresponding densities $\phi^{(m)}$, in the case of random variables in a fixed chaos of order $q \geqslant 2$. In Theorem 4.4 we obtain an inequality similar to (1.2) assuming that $E(\|DF\|_{\mathfrak{H}}^{-\beta}) < \infty$ for some $\beta > 6m + 6(\lfloor \frac{m}{2} \rfloor \vee 1)$. Again the proof is obtained by a combination of Malliavin calculus and the Stein's method. Here we need to consider Stein's equation for functions of the form of $h(x) = \mathbf{1}_{\{x>a\}} p(x)$, where p is a polynomial.

For a d dimensional random vector $F = (F^1, \dots, F^d)$ whose components are multiple stochastic integrals of orders $q_1, \dots, q_d$, $q_i \geqslant 2$, we assume non-degeneracy condition

$E[\det \gamma_F^{-p}] < \infty$ for all $p \geqslant 1$, where $\gamma_F = (\langle DF, , DF \rangle)_{1 \leqslant i,j \leqslant d}$ denotes the Malliavin matrix of F. Then, for any multi-index $\beta = (\beta_1, \dots, \beta_k)$, $1 \leqslant \beta_i \leqslant d$, we obtain the estimate (see Theorem 5.2)

$$\sup_{x \in \mathbb{R}^d} \left| \partial_\beta f_F(x) - \partial_\beta \phi(x) \right| \leqslant C \left(|V - I|^{\frac{1}{2}} + \sum_{j=1}^{d} \sqrt{E\left[F_j^4\right] - 3\left(E\left[F_j^2\right]\right)^2} \right), \tag{1.3}$$

where V is the covariance matrix of F, ϕ is the standard d dimensional normal density, and $\partial_\beta = \frac{\partial^k}{\partial x_{\beta_1} \cdots \partial x_{\beta_k}}$. As a consequence, we derive the uniform convergence of the densities and their derivatives for a sequence of vectors of multiple stochastic integrals, under the assumption $\sup_n E[\det \gamma_{F_n}^{-p}] < \infty$ for all $p \geqslant 1$. A multivariate extension of Stein's method is required for noncontinuous functions with polynomial growth (see Proposition 5.10). While univariate Stein's equations with non-smooth test functions have been extensively studied, relatively few results are available for the multivariate case, see [5,4,12,19,26,27], so this result has its own interest.

We also consider the case of random variables F such that $E[F] = 0$ and $E[F^2] = \sigma^2$, belonging to the Sobolev space $\mathbb{D}^{2,s}$ for some $s > 4$. In this case, under a non-degeneracy assumption of the form $E[|\langle DF, -DL^{-1}F \rangle_{\mathfrak{H}}|^{-r}|] < \infty$ for some $r > 2$, we derive an estimate for the uniform distance between the density of F and the density of the normal law $N(0, \sigma^2)$.

In a recent paper [20], Nourdin, Peccati and Swan have obtained an upper bound on the total variation distance between the law of a vector of multiple stochastic integrals and a normal distribution, using a combination of entropy techniques and Malliavin calculus. Their main result can be briefly stated as follows. Let $F = (F^1, \dots, F^d)$ be a d dimensional random vector whose components are multiple stochastic integrals of orders $q_1, \dots, q_d$, $q_i \geqslant 2$, respectively. Suppose the covariance of F is the identity matrix and it admits a density $f_F(x)$. Denote $\phi(x)$ the density of $N \sim N(0, \mathrm{Id})$. Then the relative entropy $D(F \| N)$ of F satisfies $D(F \| N) := E[\log f_F(F) - \log \phi(F)] \leqslant C \Delta |\log \Delta|$, where $C > 0$ is a constant and $\Delta = E[|F|^4 - |N|^4]$. This leads to the bound

$$\| f_F - \phi \|_{L^1(\mathbb{R}^d)} \leqslant \sqrt{2D(F \| N)} \leqslant C \sqrt{\Delta |\log \Delta|}.$$

This result refines some estimates obtained in [14]. In the case $d = 1$, it is finer than our estimate (4.12) in the special case $p = 1$, where by taking α close to $\frac{1}{2}$ we can only get $\Delta^{\frac{1}{4} - \epsilon}$ with $\epsilon > 0$ arbitrarily small. However, the best such L^1 estimate is still given by (1.1). It is worth mentioning that we don't assume the existence of density and our estimate (4.12) covers the L^p norm for all $p \geqslant 1$.

Convergence of densities in uniform distance has also been studied using Fisher information theory via Shimizu's inequality (see for instance, [29,3,2,8])

$$\sup_{x \in \mathbb{R}} \left| f_F(x) - \phi(x) \right| \leqslant C \sqrt{I(F \| N)}, \tag{1.4}$$

where F is a random variable with density $f \in C^1(\mathbb{R})$, ϕ is the density of $N \sim N(0, 1)$, and $I(F \| N) := \int_{\mathbb{R}} (\frac{f_F'(x)}{f_F(x)} - \frac{\phi'(x)}{\phi(x)})^2 f(x)\, dx$ is the relative Fisher information. Recently, Bobkov, Chistyakov and Götze [3] studied the rate of convergence to 0 of $I(F_n \| N)$ for $F_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$, where $\{X_i\}_{i\geqslant a}$ are i.i.d. random variables with mean 0 and variance 1, assuming that $f'_{F_{n_0}} \in L^1(\mathbb{R})$ for some n_0. In general, when studying uniform convergence of densities, one is necessarily led to introduce some stringent assumptions on the regularity of the laws of the underlying random variables. Here we showed that these assumptions can be reduced to requirements about the finiteness of the negative moments of Malliavin matrices.

The paper is organized as follows. Section 2 introduces some preliminary results of Gaussian analysis, Malliavin calculus and Stein's method for normal approximations. Section 3 is devoted to density formulae with elementary estimates using Malliavin calculus. The density formulae themselves are well-known results, but we present explicit formulae with useful estimates, such as the Hölder continuity and boundedness estimates in Theorems 3.1 and 3.3. The boundedness estimates enable us to prove the L^p convergence of the densities (see (4.12)). The Hölder continuity estimates can be used to provide a short proof for the convergence of densities based on a compactness argument, assuming convergence in law (see Theorem 6.5). Section 4 proves the convergence of densities of random variables in a fixed Wiener chaos, and Section 5 discusses convergence of densities for random vectors. In Section 6, the convergence of densities for sequences of general centered square integrable random variables are studied.

The main difficulty in the application of the above results is to verify the existence of negative moments for the determinant of the Malliavin matrix. We provide explicit sufficient conditions for this condition for random variables in the second Wiener chaos in Section 7. As an application we derive the uniform convergence of the densities and their derivatives to the normal distribution, as time goes to infinity, for the least squares estimator of the parameter θ in the Ornstein–Uhlenbeck process: $dX_t = -\theta X_t\,dt + \gamma\,dB_t$, where $B = \{B_t,\ t\geqslant 0\}$ is a standard Brownian motion. Some technical results and proofs are included in Appendix A.

Along this paper, we denote by C (maybe with subindexes) a generic constant that might depend on quantities such as the order of multiple stochastic integrals q, the order of the derivatives m, the variance σ^2 or the negative moments of the Malliavin derivative. We denote by $\|\cdot\|_p$ the norm in the space $L^p(\Omega)$.

2. Preliminaries

In the first two subsections, we introduce some basic elements of Gaussian analysis and Malliavin calculus, for which we refer to [22,16] for further details. In the last subsection, we shall introduce some basic estimates from the univariate Stein's method.

2.1. Isonormal Gaussian process and multiple integrals

Let $\mathfrak{H}$ be a real separable Hilbert space (with its inner product and norm denoted by $\langle\cdot,\cdot\rangle_{\mathfrak{H}}$ and $\|\cdot\|_{\mathfrak{H}}$, respectively). For any integer $q\geqslant 1$, let $\mathfrak{H}^{\otimes q}$ ($\mathfrak{H}^{\odot q}$) be the qth tensor product (symmetric tensor product) of $\mathfrak{H}$. Let $X = \{X(h),\ h\in\mathfrak{H}\}$ be an isonormal Gaussian process associated with the Hilbert space $\mathfrak{H}$, defined on a complete probability space $(\Omega, \mathcal{F}, P)$. That is, X is a centered Gaussian family of random variables such that $E[X(h)X(g)] = \langle h, g\rangle_{\mathfrak{H}}$ for all $h, g\in\mathfrak{H}$. We assume that the σ-field $\mathcal{F}$ is generated by X.

For every integer $q\geqslant 0$, the *qth Wiener chaos* (denoted by $\mathcal{H}_q$) of X is the closed linear subspace of $L^2(\Omega)$ generated by the random variables $\{H_q(X(h)):\ h\in\mathfrak{H},\ \|h\|_{\mathfrak{H}} = 1\}$, where H_q is the qth *Hermite polynomial* recursively defined by $H_0(x) = 1$, $H_1(x) = x$ and

$$H_{q+1}(x) = xH_q(x) - qH_{q-1}(x), \quad q\geqslant 1. \tag{2.1}$$

For every integer $q \geqslant 1$, the mapping $I_q(h^{\otimes q}) = H_q(X(h))$, where $\|h\|_{\mathfrak{H}} = 1$, can be extended to a linear isometry between $\mathfrak{H}^{\odot q}$ (equipped with norm $\sqrt{q!}\|\cdot\|_{\mathfrak{H}^{\otimes q}}$) and $\mathcal{H}_q$ (equipped with $L^2(\Omega)$ norm). For $q = 0$, $\mathcal{H}_0 = \mathbb{R}$, and I_0 is the identity map. The mapping I_q is called the multiple stochastic integral of order q.

It is well known (Wiener chaos expansion) that $L^2(\Omega)$ can be decomposed into the infinite orthogonal sum of the spaces $\mathcal{H}_q$. That is, any random variable $F \in L^2(\Omega)$ has the following chaos expansion:

$$F = \sum_{q=0}^{\infty} I_q(f_q), \tag{2.2}$$

where $f_0 = E[F]$, and $f_q \in \mathfrak{H}^{\odot q}, q \geqslant 1$, are uniquely determined by F. For every $q \geqslant 0$ we denote by J_q the orthogonal projection on the qth Wiener chaos $\mathcal{H}_q$, so $I_q(f_q) = J_q F$.

Let $\{e_n, n \geqslant 1\}$ be a complete orthonormal basis of $\mathfrak{H}$. Given $f \in \mathfrak{H}^{\odot q}$ and $g \in \mathfrak{H}^{\odot p}$, for $r = 0, \dots, p \wedge q$ the rth contraction of f and g is the element of $\mathfrak{H}^{\otimes(p+q-2r)}$ defined by

$$f \otimes_r g = \sum_{i_1,\dots,i_r=1}^{\infty} \langle f, e_{i_1} \otimes \dots \otimes e_{i_r}\rangle_{\mathfrak{H}^{\otimes r}} \otimes \langle g, e_{i_1} \otimes \dots \otimes e_{i_r}\rangle_{\mathfrak{H}^{\otimes r}}. \tag{2.3}$$

Notice that $f \otimes_r g$ is not necessarily symmetric. We denote by $f \widetilde{\otimes}_r g$ its symmetrization. Moreover, $f \otimes_0 g = f \otimes g$, and for $p = q$, $f \otimes_q g = \langle f, g\rangle_{\mathfrak{H}^{\otimes q}}$. For the product of two multiple stochastic integrals we have the multiplication formula

$$I_p(f)I_q(g) = \sum_{r=0}^{p\wedge q} r! \binom{p}{r}\binom{q}{r} I_{p+q-2r}(f \otimes_r g). \tag{2.4}$$

In the particular case $\mathfrak{H} = L^2(A, \mathcal{A}, \mu)$, where $(A, \mathcal{A})$ is a measurable space and μ is a σ-finite and nonatomic measure, one has that $\mathfrak{H}^{\otimes q} = L^2(A^q, \mathcal{A}^{\otimes q}, \mu^{\otimes q})$ and $\mathfrak{H}^{\odot q}$ is the space of symmetric and square-integrable functions on A^q. Moreover, for every $f \in \mathfrak{H}^{\odot q}$, $I_q(f)$ coincides with the qth *multiple Wiener–Itô integral* of f with respect to X, and (2.3) can be written as

$$f \otimes_r g(t_1, \dots, t_{p+q-2r}) = \int_{A^r} f(t_1, \dots, t_{q-r}, s_1, \dots, s_r) \times g(t_{1+q-r}, \dots, t_{p+q-r}, s_1, \dots, s_r)\, d\mu(s_1) \dots d\mu(s_r). \tag{2.5}$$

2.2. *Malliavin operators*

We introduce some basic facts on Malliavin calculus with respect to the Gaussian process X. Let $\mathcal{S}$ denote the class of smooth random variables of the form $F = f(X(h_1), \dots, X(h_n))$, where $h_1, \dots, h_n$ are in $\mathfrak{H}$, $n \geqslant 1$, and $f \in C_p^\infty(\mathbb{R}^n)$, the set of smooth functions f such that f itself and all its partial derivatives have at most polynomial growth. Given $F = f(X(h_1), \dots, X(h_n))$ in $\mathcal{S}$, its Malliavin derivative DF is the $\mathfrak{H}$-valued random variable given by

$$DF=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\big(X(h_1),\dots,X(h_n)\big)h_i.$$

The derivative operator D is a closable and unbounded operator on $L^2(\Omega)$ taking values in $L^2(\Omega;\mathfrak{H})$. By iteration one can define higher order derivatives $D^kF\in L^2(\Omega;\mathfrak{H}^{\odot k})$. For any integer $k\geqslant 0$ and any $p\geqslant 1$ and we denote by $\mathbb{D}^{k,p}$ the closure of $\mathcal{S}$ with respect to the norm $\|\cdot\|_{k,p}$ given by:

$$\|F\|_{k,p}^p=\sum_{i=0}^{k}E\big(\big\|D^iF\big\|_{\mathfrak{H}^{\otimes i}}^p\big).$$

For $k=0$ we simply write $\|F\|_{0,p}=\|F\|_p$. For any $p\geqslant 1$ and $k\geqslant 0$, we set $\mathbb{D}^{\infty,p}=\bigcap_{k\geqslant 0}\mathbb{D}^{k,p}$, $\mathbb{D}^{k,\infty}=\bigcap_{p\geqslant 1}\mathbb{D}^{k,p}$ and $\mathbb{D}^{\infty}=\bigcap_{k\geqslant 0}\mathbb{D}^{k,\infty}$.

We denote by δ (the *divergence* operator) the adjoint operator of D, which is an unbounded operator from a domain in $L^2(\Omega;\mathfrak{H})$ to $L^2(\Omega)$. An element $u\in L^2(\Omega;\mathfrak{H})$ belongs to the domain of δ if and only if it verifies

$$\big|E\big[\langle DF,u\rangle_{\mathfrak{H}}\big]\big|\leqslant c_u\sqrt{E\big[F^2\big]}$$

for any $F\in\mathbb{D}^{1,2}$, where c_u is a constant depending only on u. In particular, if $u\in\operatorname{Dom}\delta$, then $\delta(u)$ is characterized by the following duality relationship

$$E\big(\delta(u)F\big)=E\big(\langle DF,u\rangle_{\mathfrak{H}}\big) \tag{2.6}$$

for any $F\in\mathbb{D}^{1,2}$. This formula extends to the multiple integral δ^q, that is, for $u\in\operatorname{Dom}\delta^q$ and $F\in\mathbb{D}^{q,2}$ we have

$$E\big(\delta^q(u)F\big)=E\big(\big\langle D^qF,u\big\rangle_{\mathfrak{H}^{\otimes q}}\big).$$

We can factor out a scalar random variable in the divergence in the following sense. Let $F\in\mathbb{D}^{1,2}$ and $u\in\operatorname{Dom}\delta$ such that $Fu\in L^2(\Omega;\mathfrak{H})$. Then $Fu\in\operatorname{Dom}\delta$ and

$$\delta(Fu)=F\delta(u)-\langle DF,u\rangle_{\mathfrak{H}}, \tag{2.7}$$

provided the right-hand side is square integrable. The operators δ and D have the following commutation relationship

$$D\delta(u)=u+\delta(Du) \tag{2.8}$$

for any $u\in\mathbb{D}^{2,2}(\mathfrak{H})$ (see [22, page 37]).

The following version of Meyer's inequality (see [22, Proposition 1.5.7]) will be used frequently in this paper. Let V be a real separable Hilbert space. We can also introduce Sobolev spaces $\mathbb{D}^{k,p}(V)$ of V-valued random variables for $p \geqslant 1$ and integer $k \geqslant 1$. Then, for any $p > 1$ and $k \geqslant 1$, the operator δ is continuous from $\mathbb{D}^{k,p}(V \otimes \mathfrak{H})$ into $\mathbb{D}^{k-1,p}(V)$. That is,

$$\|\delta(u)\|_{k-1,p} \leqslant C_{k,p}\|u\|_{k,p}. \tag{2.9}$$

The operator L defined on the Wiener chaos expansion as $L = \sum_{q=0}^{\infty}(-q)J_q$ is the infinitesimal generator of the Ornstein–Uhlenbeck semigroup $T_t = \sum_{q=0}^{\infty} e^{-qt} J_q$. Its domain in $L^2(\Omega)$ is

$$\operatorname{Dom} L = \left\{F \in L^2(\Omega)\colon \sum_{q=1}^{\infty} q^2\|J_q F\|_2^2 < \infty\right\} = \mathbb{D}^{2,2}.$$

The relation between the operators D, δ and L is explained in the following formula (see [22, Proposition 1.4.3]). For $F \in L^2(\Omega)$, $F \in \operatorname{Dom} L$ if and only if $F \in \operatorname{Dom}(\delta D)$ (i.e., $F \in \mathbb{D}^{1,2}$ and $DF \in \operatorname{Dom}\delta$), and in this case

$$\delta DF = -LF. \tag{2.10}$$

For any $F \in L^2(\Omega)$, we define $L^{-1}F = -\sum_{q=1}^{\infty} q^{-1}J_q(F)$. The operator L^{-1} is called the pseudo-inverse of L. Indeed, for any $F \in L^2(\Omega)$, we have that $L^{-1}F \in \operatorname{Dom} L$, and

$$LL^{-1}F = L^{-1}LF = F - E[F].$$

We list here some properties of multiple integrals which will be used in Section 4. Fix $q \geqslant 1$ and let $f \in \mathfrak{H}^{\odot q}$. We have $I_q(f) = \delta^q(f)$ and $DI_q(f) = qI_{q-1}(f)$, and hence $I_q(f) \in \mathbb{D}^{\infty,2}$. The multiple stochastic integral $I_q(f)$ satisfies *hypercontractivity* property:

$$\|I_q(f)\|_p \leqslant C_{q,p}\|I_q(f)\|_2 \quad \text{for any } p \in [2,\infty). \tag{2.11}$$

This easily implies that $I_q(f) \in \mathbb{D}^{\infty}$ and for any $1 \leqslant k \leqslant q$ and $p \geqslant 2$,

$$\|I_q(f)\|_{k,p} \leqslant C_{q,k,p}\|I_q(f)\|_2.$$

As a consequence, for any $F \in \oplus_{l=1}^{q}\mathcal{H}_l$, we have

$$\|F\|_{k,p} \leqslant C_{q,k,p}\|F\|_2. \tag{2.12}$$

For any random variable F in the chaos of order $q \geqslant 2$, we have (see [16], Eq. (5.2.7))

$$\frac{1}{q^2}\operatorname{Var}\left(\|DF\|_{\mathfrak{H}}^2\right) \leqslant \frac{q-1}{3q}\left(E\left[F^4\right] - \left(E\left[F^2\right]\right)^2\right) \leqslant (q-1)\operatorname{Var}\left(\|DF\|_{\mathfrak{H}}^2\right). \tag{2.13}$$

In the case where $\mathfrak{H}$ is $L^2(A, \mathcal{A}, \mu)$, for an integrable random variable $F = \sum_{q=0}^{\infty} I_q(f_q) \in \mathbb{D}^{1,2}$, its derivative can be represented as an element in of $L^2(A \times \Omega)$ given by

$$D_t F = \sum_{q=1}^{\infty} q I_q\big(f_q(\cdot, t)\big), \quad t \in A.$$

2.3. Stein's method of normal approximation

We shall now give a brief account of Stein's method of univariate normal approximation and its connection with Malliavin calculus. For a more detailed exposition we refer to [5,16,30].

Let F be an arbitrary random variable and let N be an $N(0, \sigma^2)$ distributed random variable, where $\sigma^2 > 0$. Consider the distance between the law of F and the law of N given by

$$d_{\mathcal{H}}(F, N) = \sup_{h \in \mathcal{H}} \big| E\big[h(F) - h(N)\big]\big| \tag{2.14}$$

for a class of functions $\mathcal{H}$ such that $E[h(F)]$ and $E[h(N)]$ are well-defined for $h \in \mathcal{H}$. Notice first the following fact (which is usually referred as *Stein's lemma*): a random variable N is $N(0, \sigma^2)$ distributed if and only if $E[\sigma^2 f'(N) - N f(N)] = 0$ for all absolutely continuous functions f such that $E[|f'(N)|] < \infty$. This suggests that the distance of $E[\sigma^2 f'(F) - F f(F)]$ from zero may quantify the distance between the law of F and the law of N. To see this, for each function h such that $E[|h(N)|] < \infty$, Stein [30] introduced the *Stein's equation*:

$$f'(x) - \frac{x}{\sigma^2} f(x) = h(x) - E\big[h(N)\big] \tag{2.15}$$

for all $x \in \mathbb{R}$. For a random variable F such that $E[|h(F)|] < \infty$, any solution f_h to Eq. (2.15) verifies

$$\frac{1}{\sigma^2} E\big[\sigma^2 f_h'(F) - F f_h(F)\big] = E\big[h(F) - h(N)\big], \tag{2.16}$$

and the distance defined in (2.14) can be written as

$$d_{\mathcal{H}}(F, N) = \frac{1}{\sigma^2} \sup_{h \in \mathcal{H}} \big| E\big[\sigma^2 f_h'(F) - F f_h(F)\big]\big|. \tag{2.17}$$

The unique solution to (2.15) verifying $\lim_{x \to \pm\infty} e^{-x^2/(2\sigma^2)} f(x) = 0$ is

$$f_h(x) = e^{x^2/(2\sigma^2)} \int_{-\infty}^{x} \big\{h(y) - E\big[h(N)\big]\big\} e^{-y^2/(2\sigma^2)}\, dy. \tag{2.18}$$

From (2.17) and (2.18), one can get bounds for probability distances like the total variation distance, where we let $\mathcal{H}$ consist of all indicator functions of measurable sets, Kolmogorov distance, where we consider all the half-line indicator functions and Wasserstein distance, where we take $\mathcal{H}$ to be the set of all Lipschitz-continuous functions with Lipschitz constant equal to 1.

In the present paper, we shall consider the case when h: $\mathbb{R} \to \mathbb{R}$ is given by $h(x) = \mathbf{1}_{\{x>z\}} H_k(x)$ for any integer $k \geqslant 1$ and $z \in \mathbb{R}$, where $H_k(x)$ is the kth Hermite polynomial. More generally, we have the following lemma whose proof can be found in Appendix A. It should be pointed out that the univariate Stein's equations have been extensively studied. For example, we refer to [5, Section 2.2] and [18, Lemma 8.2] when the test functions have sub-polynomial growth.

Lemma 2.1. *Suppose $|h(x)| \leqslant a|x|^k + b$ for some integer $k \geqslant 0$ and some nonnegative numbers a, b. Then, the solution f_h to the Stein's equation (2.15) given by (2.18) satisfies*

$$\left|f_h'(x)\right| \leqslant aC_k \sum_{i=0}^{k} \sigma^{k-i} |x|^i + 4b$$

for all $x \in \mathbb{R}$, where C_k is a constant depending only on k.

Nourdin and Peccati [15,16] combined Stein's method with Malliavin calculus to estimate the distance between the distributions of regular functionals of an isonormal Gaussian process and the normal distribution $N(0, \sigma^2)$. The basic ingredient is the following integration by parts formula. For $F \in \mathbb{D}^{1,2}$ with $E[F] = 0$ and any function $f \in C^1$ such that $E[|f'(F)|] < \infty$, using (2.10) and (2.6) we have

$$\begin{aligned} E\big[Ff(F)\big] &= E\big[LL^{-1}Ff(F)\big] = E\big[-\delta DL^{-1}Ff(F)\big] \\ &= E\big[\big\langle -DL^{-1}F, Df(F)\big\rangle\big] = E\big[f'(F)\big\langle -DL^{-1}F, DF\big\rangle_{\mathfrak{H}}\big]. \end{aligned}$$

Then, it follows that

$$E\big[\sigma^2 f'(F) - Ff(F)\big] = E\big[f'(F)\big(\sigma^2 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\big)\big]. \tag{2.19}$$

Combining Eq. (2.19) with (2.16) and Lemma 2.1 we obtain the following result.

Lemma 2.2. *Suppose $h : \mathbb{R} \to \mathbb{R}$ verifies $|h(x)| \leqslant a|x|^k + b$ for some $a, b \geqslant 0$ and some integer $k \geqslant 0$. Let $N \sim N(0, \sigma^2)$ and let $F \in \mathbb{D}^{1,2k}$ with $\|F\|_{2k} \leqslant c\sigma$ for some $c > 0$. Then there exists a constant $C_{k,c}$ depending only on k and c such that*

$$\left|E\big[h(F) - h(N)\big]\right| \leqslant \sigma^{-2}\big[aC_{k,c}\sigma^k + 4b\big]\big\|\sigma^2 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\big\|_2.$$

Proof. From (2.16), (2.19) and Lemma 2.1, it suffices to notice that $\|\sum_{i=0}^{k} \sigma^{k-i}|F|^i\|_2 \leqslant \sum_{i=0}^{k} \|F\|_{2k}^i \sigma^{k-i} \leqslant C_{k,c}\sigma^k$. $\square$

3. Density formulae

In this section, we present explicit formulae for the density of a random variable and its derivatives, using the techniques of Malliavin calculus.

3.1. Density formulae

We shall present two explicit formulae for the density of a random variable, with estimates of its uniform and Hölder norms.

Theorem 3.1. *Let $F \in \mathbb{D}^{2,s}$ such that $E[|F|^{2p}] < \infty$ and $E[\|DF\|_{\mathfrak{H}}^{-2r}] < \infty$ for $p, r, s > 1$ satisfying $\frac{1}{p} + \frac{1}{r} + \frac{1}{s} = 1$. Denote*

$$w = \|DF\|_{\mathfrak{H}}^2, \qquad u = w^{-1} DF.$$

Then $u \in \mathbb{D}^{1,p'}$ with $p' = \frac{p}{p-1}$ and F has a density given by

$$f_F(x) = E\big[\mathbf{1}_{\{F>x\}} \delta(u)\big]. \tag{3.1}$$

Furthermore, $f_F(x)$ is bounded and Hölder continuous of order $\frac{1}{p}$, that is

$$f_F(x) \leqslant C_p \big\|w^{-1}\big\|_r \|F\|_{2,s} \big(1 \wedge \big(|x|^{-2} \|F\|_{2p}^2\big)\big), \tag{3.2}$$

$$\big|f_F(x) - f_F(y)\big| \leqslant C_p \big\|w^{-1}\big\|_r^{1+\frac{1}{p}} \|F\|_{2,s}^{1+\frac{1}{p}} |x-y|^{\frac{1}{p}} \tag{3.3}$$

for any $x, y \in \mathbb{R}$, where C_p is a constant depending only on p.

Proof. Note that

$$Du = w^{-1} D^2 F - 2w^{-2} \big(D^2 F \otimes_1 DF\big) \otimes DF.$$

Applying Meyer's inequality (2.9) and Hölder's inequality we have

$$\begin{aligned} \big\|\delta(u)\big\|_{p'} &\leqslant C_p \|u\|_{1,p'} \leqslant C_p \big(\|u\|_{p'} + \|Du\|_{p'}\big) \\ &\leqslant C_p \big(\big\|w^{-1} \|DF\|_{\mathfrak{H}}\big\|_{p'} + 3\big\|w^{-1} \big\|D^2 F\big\|_{\mathfrak{H}\otimes\mathfrak{H}}\big\|_{p'}\big) \\ &\leqslant 3C_p \big\|w^{-1}\big\|_r \big(\|DF\|_s + \big\|D^2 F\big\|_s\big). \end{aligned} \tag{3.4}$$

Then $u \in \mathbb{D}^{1,p'}$ and the density formula (3.1) holds (see, for instance, Nualart [22, Proposition 2.1.1]). From $E[\delta(u)] = 0$ and Hölder's inequality it follows that

$$\big|E\big[\mathbf{1}_{\{F>x\}} \delta(u)\big]\big| \leqslant P\big(|F| > |x|\big)^{\frac{1}{p}} \big\|\delta(u)\big\|_{p'} \leqslant \big(1 \wedge \big(|x|^{-2p} \|F\|_{2p}^{2p}\big)\big)^{\frac{1}{p}} \big\|\delta(u)\big\|_{p'}. \tag{3.5}$$

Then (3.2) follows from (3.5) and (3.4).

Finally, for $x < y \in \mathbb{R}$, noticing that $\mathbf{1}_{\{F>x\}} - \mathbf{1}_{\{F>y\}} = \mathbf{1}_{\{x<F\leqslant y\}}$, we have

$$\big|f_F(x) - f_F(y)\big| \leqslant \big(E[\mathbf{1}_{\{x<F\leqslant y\}}]\big)^{\frac{1}{p}} \big\|\delta(u)\big\|_{p'}.$$

Applying (3.2) and (3.4) with the fact that $E[\mathbf{1}_{\{x<F\leqslant y\}}] = \int_x^y f_F(z)\,dz$ one gets (3.3). □

With the exact proof of [22, Propositions 2.1.1], one can prove the following slightly more general result.

Proposition 3.2. *Let $F \in \mathbb{D}^{1,p}$ and $h : \Omega \to \mathfrak{H}$, and suppose that $\langle DF, h\rangle_{\mathfrak{H}} \neq 0$ a.s. and $\frac{h}{\langle DF,h\rangle_{\mathfrak{H}}} \in \mathbb{D}^{1,q}(\mathfrak{H})$ for some $p, q > 1$. Then the law of F has a density given by*

$$f_F(x) = E\left[\mathbf{1}_{\{F>x\}}\delta\left(\frac{h}{\langle DF, h\rangle_{\mathfrak{H}}}\right)\right]. \tag{3.6}$$

Our next goal is to take h to be $-DL^{-1}F$ in formula (3.6) and get a result similar to Theorem 3.1. First, to get a sufficient condition for $\frac{-DL^{-1}F}{\langle DF,-DL^{-1}F\rangle_{\mathfrak{H}}} \in \mathbb{D}^{1,p'}$ for some $p' > 1$, we need some technical estimates on $DL^{-1}F$ and $D^2L^{-1}F$. Estimates of this type have been obtained by Nourdin, Peccati and Reinert [17] (see also Nourdin and Peccati's book [16, Lemma 5.3.8]), when proving an infinite-dimensional Poincaré inequality. More precisely, by using Mehler's formula, they proved that for any $p \geqslant 1$, if $F \in \mathbb{D}^{2,p}$, then

$$E\big[\big\|DL^{-1}F\big\|_{\mathfrak{H}}^p\big] \leqslant E\big[\|DF\|_{\mathfrak{H}}^p\big]. \tag{3.7}$$

$$E\big[\big\|D^2L^{-1}F\big\|_{op}^p\big] \leqslant 2^{-p}E\big[\big\|D^2F\big\|_{op}^p\big], \tag{3.8}$$

where $\|D^2F\|_{op}$ denotes the operator norm of the Hilbert–Schmidt operator from $\mathfrak{H}$ to $\mathfrak{H}$: $f \mapsto f \otimes_1 D^2F$. Furthermore, the operator norm $\|D^2F\|_{op}$ satisfies the following "*random contraction inequality*"

$$\big\|D^2F\big\|_{op}^4 \leqslant \big\|D^2F \otimes_1 D^2F\big\|_{\mathfrak{H}^{\otimes 2}}^2 \leqslant \big\|D^2F\big\|_{\mathfrak{H}^{\otimes 2}}^4. \tag{3.9}$$

Sometimes in application, the use of L^{-1} in the integration by parts formula is useful. The next proposition gives a density formula with estimates similar to Theorem 3.1 with the use of L^{-1}. Let

$$\bar{w} = \big\langle DF, -DL^{-1}F\big\rangle_{\mathfrak{H}}, \qquad \bar{u} = -\bar{w}^{-1}DL^{-1}F.$$

Proposition 3.3. *Let $F \in \mathbb{D}^{2,s}$, $E[|F|^{2p}] < \infty$ and suppose that $E[|\bar{w}|^{-r}] < \infty$, where $p > 1$, $r > 2$, $s > 3$ satisfy $\frac{1}{p} + \frac{2}{r} + \frac{3}{s} = 1$. Then $\bar{u} \in \mathbb{D}^{1,p'}$ with $p' = \frac{p}{p-1}$ and the law of F has a density given by*

$$f_F(x) = E\big[\mathbf{1}_{\{F>x\}}\delta(\bar{u})\big]. \tag{3.10}$$

Furthermore, $f_F(x)$ is bounded and Hölder continuous of order $\frac{1}{p}$, that is

$$f_F(x) \leqslant K_0\big(1 \wedge \big(|x|^{-2}\|F\|_{2p}^2\big)\big), \tag{3.11}$$

$$\big|f_F(x) - f_F(y)\big| \leqslant K_0^{1+\frac{1}{p}}|x-y|^{\frac{1}{p}} \tag{3.12}$$

for any $x, y \in \mathbb{R}$, where $K_0 = C_p\|\bar{w}^{-1}\|_r\|F\|_{2,s}(\|\bar{w}^{-1}\|_r\|DF\|_s^2 + 1)$, and C_p depends only on p.

Proof. Note that $D\bar{w} = -D^2F \otimes_1 DL^{-1}F - DF \otimes_1 D^2L^{-1}F$. Then, applying (3.7) and (3.8) we obtain

$$\|D\bar{w}\|_{\frac{s}{2}} \leqslant \left(1+2^{-s}\right)\left\| \left\|D^2F\right\|_{op}\right\|_s \|DF\|_s. \tag{3.13}$$

From $\bar{u} = -\bar{w}^{-1}DL^{-1}F$ we get $D\bar{u} = -\bar{w}^{-1}D^2L^{-1}F + w^{-2}D\bar{w} \otimes DL^{-1}F$. Then, using (3.7)–(3.9) we have for $t > 0$ satisfying $\frac{1}{p'} = \frac{1}{r} + \frac{1}{t}$,

$$\|\bar{u}\|_{p'} \leqslant \left\|\bar{w}^{-1}\left\|DL^{-1}F\right\|_{\mathfrak{H}}\right\|_{p'} \leqslant \left\|\bar{w}^{-1}\right\|_r \|DF\|_t,$$

and

$$\begin{aligned}\|D\bar{u}\|_{p'} &\leqslant \left\|\bar{w}^{-1}\left\|D^2L^{-1}F\right\|_{\mathfrak{H}\otimes\mathfrak{H}}\right\|_{p'} + \left\|\bar{w}^{-2}\|D\bar{w}\|_{\mathfrak{H}}\left\|DL^{-1}F\right\|_{\mathfrak{H}}\right\|_{p'} \\ &\leqslant \left\|\bar{w}^{-1}\right\|_r\left\|D^2F\right\|_t + \left\|\bar{w}^{-2}\right\|_r \|D\bar{w}\|_{\frac{s}{2}}\|DF\|_s.\end{aligned}$$

Noticing that $\|D^2F\|_t \leqslant \|D^2F\|_s$ because $t < s$, and applying Meyer's inequality (2.9) with (3.13) and (3.9) we obtain

$$\left\|\delta(\bar{u})\right\|_{p'} \leqslant C_p\|\bar{u}\|_{1,p'} \leqslant K_0. \tag{3.14}$$

Then $u \in \mathbb{D}^{1,p'}$ and the density formula (3.10) holds. As in the proof of Theorem 3.1, (3.11) and (3.12) follow from (3.14) and

$$\left|E\left[\mathbf{1}_{\{F>x\}}\delta(\bar{u})\right]\right| \leqslant P\left(|F|>|x|\right)^{\frac{1}{p}}\left\|\delta(\bar{u})\right\|_{p'} \leqslant \left(1 \wedge \left(|x|^{-2}\|F\|_{2p}^2\right)\right)\left\|\delta(\bar{u})\right\|_{p'},$$

$$\left|f_F(x) - f_F(y)\right| \leqslant \left(E[\mathbf{1}_{\{x<F\leqslant y\}}]\right)^{\frac{1}{p}}\left\|\delta(u)\right\|_{p'}. \qquad \square$$

3.2. Derivatives of the density

Next we present a formula for the derivatives of the density function, under additional conditions. A sequence of recursively defined random variables given by $G_0 = 1$ and $G_{k+1} = \delta(G_k u)$ where u is an $\mathfrak{H}$-valued process, plays an essential role in the formula. The following technical lemma gives an explicit formula for the sequence G_k, relating it to Hermite polynomials. To simplify the notation, for an $\mathfrak{H}$-valued random variable u, we denote

$$\delta_u = \delta(u), \qquad D_uG = \langle DG, u\rangle_{\mathfrak{H}}, \qquad D_u^kG = \left\langle D\left(D_u^{k-1}G\right), u\right\rangle_{\mathfrak{H}}. \tag{3.15}$$

Recall $H_k(x)$ denotes the kth Hermite polynomial. For $\lambda > 0$ and $x \in \mathbb{R}$, we define the generalized kth Hermite polynomial as

$$H_k(\lambda, x) = \lambda^{\frac{k}{2}} H_k\left(\frac{x}{\sqrt{\lambda}}\right). \tag{3.16}$$

From the property $H_k'(x) = kH_{k-1}(x)$ it follows by induction that the kth Hermite polynomials has the form $H_k(x) = \sum_{0\leqslant i\leqslant \lfloor k/2\rfloor} c_{k,i}x^{k-2i}$, where we denote by $\lfloor k/2\rfloor$ the largest integer less than or equal to $k/2$. Then (3.16) implies

$$H_k(\lambda, x) = \sum_{0\leqslant i\leqslant \lfloor k/2\rfloor} c_{k,i}x^{k-2i}\lambda^i. \tag{3.17}$$

Lemma 3.4. *Fix an integer $m \geqslant 1$ and a number $p > m$. Suppose $u \in \mathbb{D}^{m,p}(\mathfrak{H})$. We define recursively a sequence $\{G_k\}_{k=0}^m$ by $G_0 = 1$ and $G_{k+1} = \delta(G_k u)$. Then, these variables are well-defined and for $k = 1, 2, \dots, m$, $G_k \in \mathbb{D}^{m-k,\frac{p}{k}}$ and*

$$G_k = H_k(D_u\delta_u, \delta_u) + T_k, \tag{3.18}$$

where we denote by T_k the higher order derivative terms which can be defined recursively as follows: $T_1 = T_2 = 0$ and for $k \geqslant 2$,

$$T_{k+1} = \delta_u T_k - D_u T_k - \partial_\lambda H_k(D_u\delta_u, \delta_u)D_u^2\delta_u. \tag{3.19}$$

The following lemma is proved in Appendix A.

Lemma 3.5. *From (3.19) we can deduce that for $k \geqslant 3$*

$$T_k = \sum_{(i_0,\dots,i_{k-1})\in J_k} a_{i_0,i_1,\dots,i_{k-1}}\delta_u^{i_0}(D_u\delta_u)^{i_1}\left(D_u^2\delta_u\right)^{i_2}\cdots\left(D_u^{k-1}\delta_u\right)^{i_{k-1}}, \tag{3.20}$$

where the coefficients $a_{i_0,i_1,\dots,i_{k-1}}$ are real numbers and J_k is the set of multi-indices $(i_0, i_1, \dots, i_{k-1}) \in \mathbb{N}^k$ satisfying the following three conditions

$$\text{(a)}\quad i_0 + \sum_{j=1}^{k-1} j i_j \leqslant k-1; \qquad \text{(b)}\quad i_2 + \dots + i_{k-1} \geqslant 1; \qquad \text{(c)}\quad \sum_{j=1}^{k-1} i_j \leqslant \left\lfloor \frac{k-1}{2} \right\rfloor.$$

From (b) *we see that every term in T_k contains at least one factor of the form $D_u^j\delta_u$ with some $j \geqslant 2$. We shall show this type of factors will converge to zero. For this reason we call these terms high order terms.*

Proof of Lemma 3.4. First, we prove by induction on k that the above sequence G_k is well-defined and $G_k \in \mathbb{D}^{m-k,\frac{p}{k}}$. Suppose first that $k = 1$. Then, Meyer's inequality implies that $G_1 = \delta_u \in \mathbb{D}^{m-1,p}$. Assume now that for $k \leqslant m-1$, $G_k \in \mathbb{D}^{m-k,\frac{p}{k}}$. Then it follows from Meyer's and Hölder's inequalities (see [22, Proposition 1.5.6]) that

$$\|G_{k+1}\|_{m-k-1,\frac{p}{k+1}} \leqslant C_{m,p}\|G_k u\|_{m-k,\frac{p}{k+1}} \leqslant C'_{m,p}\|G_k\|_{m-k,\frac{p}{k}}\|u\|_{m-k,p} < \infty.$$

Let us now show, by induction, the decomposition (3.18). When $k = 1$ (3.18) is true because $G_1 = \delta_u$ and $T_1 = 0$. Assume now (3.18) holds for $k \leqslant m-1$. Noticing that $\partial_x H_k(\lambda, x) = kH_{k-1}(\lambda, x)$ (since $H_k'(x) = kH_{k-1}(x)$), we get

$$D_u H_k(D_u\delta_u, \delta_u) = kH_{k-1}(D_u\delta_u, \delta_u)D_u\delta_u + \partial_\lambda H_k(D_u\delta_u, \delta_u)D_u^2\delta_u.$$

Hence, applying the operator D_u to both sides of (3.18),

$$D_u G_k = kH_{k-1}(D_u\delta_u, \delta_u)D_u\delta_u + \widetilde{T}_{k+1},$$

where

$$\widetilde{T}_{k+1} = D_u T_k + \partial_\lambda H_k(D_u\delta_u, \delta_u)D_u^2\delta_u. \tag{3.21}$$

From the definition of G_{k+1} and using (2.7) we obtain

$$\begin{aligned} G_{k+1} &= \delta(uG_k) = G_k\delta_u - D_u G_k \\ &= \delta_u H_k(D_u\delta_u, \delta_u) + \delta_u T_k - kH_{k-1}(D_u\delta_u, \delta_u)D_u\delta_u - \widetilde{T}_{k+1}. \end{aligned}$$

Note that $H_{k+1}(x) = xH_k(x) - kH_{k-1}(x)$ implies $xH_k(\lambda, x) - k\lambda H_{k-1}(\lambda, x) = H_{k+1}(\lambda, x)$. Hence,

$$G_{k+1} = H_{k+1}(D_u\delta_u, \delta_u) + \delta_u T_k - \widetilde{T}_{k+1}.$$

The term $T_{k+1} = \delta_u T_k - \widetilde{T}_{k+1}$ has the form given in (3.19). This completes the proof. □

Now we are ready to present some formulae for the derivatives of the density function under certain sufficient conditions on the random variable F. For a random variable F in $\mathbb{D}^{1,2}$ and for any $\beta \geqslant 1$ we are going to use the notation

$$M_\beta(F) = \left(E\left[\|DF\|_{\mathfrak{H}}^{-\beta}\right]\right)^{\frac{1}{\beta}}. \tag{3.22}$$

Proposition 3.6. *Fix an integer $m \geqslant 1$. Let F be a random variable in $\mathbb{D}^{m+2,\infty}$ such that $M_\beta(F) < \infty$ for some $\beta > 3m + 3(\lfloor\frac{m}{2}\rfloor \vee 1)$. Denote $w = \|DF\|_{\mathfrak{H}}^2$ and $u = \frac{DF}{w}$. Then, $u \in \mathbb{D}^{m+1,p}(\mathfrak{H})$ for some $p > 1$, and the random variables $\{G_k\}_{k=0}^{m+1}$ introduced in Lemma 3.4 are well-defined. Under these assumptions, F has a density f of class C^m with derivatives given by*

$$f_F^{(k)}(x) = (-1)^k E[\mathbf{1}_{\{F>x\}}G_{k+1}] \tag{3.23}$$

for $k = 1, \ldots, m$.

Proof. It is enough to show that $\{G_k\}_{k=0}^{m+1}$ are well-defined, since it follows from [22, Exercise 2.1.4] that the kth derivative of the density of F is given by (3.23). To do this we will show that G_k defined in (3.18) are in $L^1(\Omega)$ for all $k = 1, \dots, m+1$. From (3.18) we can write

$$E\big[|G_k|\big] \leqslant E\big[\big|H_k(D_u\delta_u, \delta_u)\big|\big] + E\big[|T_k|\big].$$

Recall the explicit expression of $H_k(\lambda, x)$ in (3.17). Since $\beta > 3(m+1)$, we can choose $r_0 < \frac{\beta}{3}$, $r_1 < \frac{\beta}{6}$ such that

$$1 \geqslant \frac{k-2i}{r_0} + \frac{i}{r_1} > \frac{3(k-2i)}{\beta} + \frac{6i}{\beta} = \frac{3k}{\beta},$$

for any $0 \leqslant i \leqslant \lfloor k/2 \rfloor$ and $1 \leqslant k \leqslant m+1$. Then, applying Hölder's inequality with (3.17), (A.11) and (A.12) we have

$$E\big[\big|H_k(D_u\delta_u, \delta_u)\big|\big] \leqslant C_k \sum_{0 \leqslant i \leqslant \lfloor k/2 \rfloor} \|\delta_u\|_{r_0}^{k-2i} \|D_u\delta_u\|_{r_1}^{i} < \infty.$$

To prove that $E[|T_k|] < \infty$, applying Hölder's inequality to the expression (3.20) and choosing $r_j > 0$ for $0 \leqslant j \leqslant k-1$ such that

$$1 \geqslant \frac{i_0}{r_0} + \sum_{j=1}^{k-1} \frac{i_j}{r_j} > \frac{3i_0}{\beta} + \sum_{j=1}^{k-1} \frac{(3j+3)i_j}{\beta},$$

we obtain that, (assuming $k \geqslant 3$, otherwise $T_k = 0$)

$$E\big[|T_k|\big] \leqslant C \sum_{(i_0, \dots, i_k) \in J_k} \|\delta_u\|_{r_0}^{i_0} \prod_{j=1}^{k-1} \big\|D_u^j \delta_u\big\|_{r_j}^{i_j}.$$

Due to (A.11) and (A.12), this expression is finite, provided $r_j < \frac{\beta}{3j+3}$ for $0 \leqslant j \leqslant k-1$. We can choose $(r_j, 0 \leqslant j \leqslant k-1)$ satisfying the above conditions because $\beta > 3(k-1) + 3\lfloor \frac{k-1}{2} \rfloor$ for all $1 \leqslant k \leqslant m+1$, and from properties (a) and (c) of J_k in Lemma 3.5 we have

$$\frac{3i_0}{\beta} + \sum_{j=1}^{k-1} \frac{(3j+3)i_j}{\beta} \leqslant \frac{3(k-1) + 3\lfloor \frac{k-1}{2} \rfloor}{\beta}.$$

This completes the proof. □

Example 3.7. Consider a random variable in the first Wiener chaos $N = I_1(h)$, where $\in \mathfrak{H}$ with $\|h\|_{\mathfrak{H}} = \sigma$. Then N has the normal distribution $N \sim N(0, \sigma^2)$ with density denoted by $\phi(x)$. Clearly $\|DN\|_{\mathfrak{H}} = \sigma$, $u = \frac{h}{\sigma^2}$, $\delta_u = \frac{N}{\sigma^2}$ and $D_u\delta_u = \frac{h}{\sigma^2}$. Then $G_k = H_k(\frac{1}{\sigma^2}, \frac{N}{\sigma^2})$ and from (3.23) we obtain the formula

$$\phi^{(k)}(x) = (-1)^k E\bigg[\mathbf{1}_{\{N>x\}} H_{k+1}\bigg(\frac{1}{\sigma^2}, \frac{N}{\sigma^2}\bigg)\bigg], \tag{3.24}$$

which can also be obtained by analytic arguments.

Remark 3.8. Let g be a function differentiable of order m, and denote $u_j(g,x) = \sup_{|t|\leqslant x}|g^{(j)}(t) - g^{(j)}(0)|$. Let F be a random variable with $E(F) = 0$, $E(F^2) = 1$ and $E\{|F|^{m+1}u_m(g,F)\} < \infty$. It is proved in [1] that the following expansion holds:

$$E\big(Fg(F)\big) - E\big(g'(F)\big) = \sum_{j=2}^{m} \frac{\gamma_{j+1}}{j!} Eg^{(j)}(F) + R,$$

where γ_j is the jth cumulant of X and $|R| \leqslant CE\{|F|^{m+1}u_m(g,F)\}$ for some constant $C > 0$. For any function h, let f be the solution of the Stein's equation (2.15) given by (2.18). Then

$$\begin{aligned} E\big[h(F)\big] - E\big[h(N)\big] &= E\big[f'(F)\big] - E\big[Ff(F)\big] \\ &= -\sum_{j=2}^{m} \frac{\gamma_{j+1}}{j!} E\big[f^{(j)}(F)\big] - R. \end{aligned}$$

This is the so-called Edgeworth expansion (see also [28] and references therein). Eq. (3.23) can also be used to compute $E[f^{(j)}(F)]$. We have easily

$$E\big[h(F)\big] - E\big[h(N)\big] = -\sum_{j=2}^{m} \frac{\gamma_{j+1}}{j!} E\big[f(F)G_k\big] - R, \tag{3.25}$$

where G_k is given in Lemma 3.4. Thus, it is possible to use Malliavin calculus to obtain the full Edgeworth expansion without assuming the differentiability of f. However, we shall not pursue this aspect in the present work.

Remark 3.9. The recursive algorithms used in Lemma 3.4 have some similarities with the recursive formula developed by Privault in [25] to compute $E(F[\delta(u)]^n)$.

4. Random variables in the qth Wiener chaos

In this section we establish our main results on uniform estimates and uniform convergence of densities and their derivatives. We shall deal first with the convergence of densities and later we consider their derivatives.

4.1. Uniform estimates of densities

Let $F = I_q(f)$ for some $f \in \mathfrak{H}^{\odot q}$ and $q \geqslant 2$. To simplify the notation, along this section we denote

$$w = \|DF\|^2_{\mathfrak{H}}, \qquad u = w^{-1}DF.$$

Note that $LF = -qF$ and using (2.7) and (2.10) we can write

$$\delta_u = \delta(u) = qFw^{-1} - \big\langle Dw^{-1}, DF\big\rangle_{\mathfrak{H}}. \tag{4.1}$$

Theorem 4.1. *Let $F = I_q(f)$, $q \geqslant 2$, for some $f \in \mathfrak{H}^{\odot q}$ be a random variable in the qth Wiener chaos with $E[F^2] = \sigma^2$. Assume that $M_6(F) < \infty$, where $M_6(F)$ is defined in (3.22). Let $\phi(x)$ be the density of $N \sim N(0, \sigma^2)$. Then F has a density $f_F(x)$ given by (3.1). Furthermore,*

$$\sup_{x\in\mathbb{R}} \left| f_F(x) - \phi(x) \right| \leqslant C\sqrt{E\left[F^4\right] - 3\sigma^4}, \tag{4.2}$$

where the constant C has the form $C = C_q(\sigma^{-1}M_6(F)^2 + M_6(F)^3 + \sigma^{-3})$ and C_q depends only on q.

We begin with a lemma giving an estimate for the contraction $D^k F \otimes_1 D^l F$ with $k + l \geqslant 3$.

Lemma 4.2. *Let $F = I_q(f)$ be a random variable in the qth Wiener chaos with $E[F^2] = \sigma^2$. Then for any integers $k \geqslant l \geqslant 1$ satisfying $k + l \geqslant 3$, there exists a constant $C_{k,l,q}$ depending only on k, l, q such that*

$$\left\| D^k F \otimes_1 D^l F \right\|_2 \leqslant C_{k,l,q} \left\| q\sigma^2 - \|DF\|_{\mathfrak{H}}^2 \right\|_2. \tag{4.3}$$

Proof. Note that $D^k F = q(q-1)\cdots(q-k+1)I_{q-k}(f)$. Applying (2.4), we get

$$D^k F \otimes_1 D^l F = q^2(q-1)^2 \cdot\cdots\cdot (q-l+1)^2(q-l)\cdots(q-k+1) \times \sum_{r=0}^{q-k} r! \binom{q-k}{r}\binom{q-l}{r} I_{2q-k-l-2r}(f \widetilde{\otimes}_{r+1} f).$$

Taking into account the orthogonality of multiple integrals of different orders, we obtain

$$\begin{aligned} &E\left[\left\| D^k F \otimes_1 D^l F \right\|^2_{\mathfrak{H}^{\otimes(k+l-2)}}\right] \\ &\quad= \frac{(q!)^4}{(q-l)!^2(q-k)!^2} \\ &\qquad\times \sum_{r=0}^{q-k} r!^2 \binom{q-k}{r}^2 \binom{q-l}{r}^2 (2q-k-l-2r)! \| f \widetilde{\otimes}_{r+1} f \|^2_{\mathfrak{H}^{\otimes 2q-2-2r}}. \end{aligned} \tag{4.4}$$

Applying (4.4) with $k = l = 1$, we obtain

$$\begin{aligned} E\left[\|DF\|_{\mathfrak{H}}^4\right] = E\left[|DF \otimes_1 DF|^2\right] &= q^4 \sum_{r=0}^{q-1} r!^2 \binom{q-1}{r}^4 (2q-2-2r)! \| f \widetilde{\otimes}_{r+1} f \|^2_{\mathfrak{H}^{\otimes 2q-2-2r}} \\ &= q^4 \sum_{r=0}^{q-2} r!^2 \binom{q-1}{r}^4 (2q-2-2r)! \| f \widetilde{\otimes}_{r+1} f \|^2_{\mathfrak{H}^{\otimes 2q-2-2r}} \\ &\quad + q^2 q!^2 \| f \|^4_{\mathfrak{H}^{\otimes q}}. \end{aligned} \tag{4.5}$$

Taking into account that $\sigma^2 = E[F^2] = q!\|f\|^2_{\mathfrak{H}^{\otimes q}}$, we obtain that for any $k + l \geqslant 3$, there exists a constant $C_{k,l,q}$ such that

$$E\big[\big\|D^k F \otimes_1 D^l F\big\|^2_{\mathfrak{H}^{\otimes(k+l-2)}}\big] \leqslant C^2_{k,l,q} E\big[\|DF\|^4_{\mathfrak{H}} - q^2\sigma^4\big].$$

Meanwhile, it follows from $E[\|DF\|^2_{\mathfrak{H}}] = q\|f\|^2_{\mathfrak{H}^{\otimes q}} = q\sigma^2$ that

$$\begin{aligned} E\big[\|DF\|^4_{\mathfrak{H}} - q^2\sigma^4\big] &= E\big[\|DF\|^4_{\mathfrak{H}} - 2q\sigma^2\|DF\|^2_{\mathfrak{H}} + q^2\sigma^4\big] \\ &= E\big[\big(\|DF\|^2_{\mathfrak{H}} - q\sigma^2\big)^2\big]. \end{aligned} \tag{4.6}$$

Combining (4.4), (4.5) and (4.6) we have

$$E\big[\big\|D^k F \otimes_1 D^l F\big\|^2_{\mathfrak{H}^{\otimes(k+l-2)}}\big] \leqslant C^2_{k,l,q} E\big[\big(\|DF\|^2_{\mathfrak{H}} - q\sigma^2\big)^2\big],$$

which completes the proof. □

Proof of Theorem 4.1. It follows from Theorem 3.1 that F admits a density $f_F(x) = E[\mathbf{1}_{\{F>x\}}\delta(u)]$. By (3.24) with $k = 1$ we can write $\phi(x) = \frac{1}{\sigma^2} E[\mathbf{1}_{\{N>x\}}N]$. Then, using (4.1), for all $x \in \mathbb{R}$ we obtain

$$\begin{aligned} f_F(x) - \phi(x) &= E\big[\mathbf{1}_{\{F>x\}}\delta(u)\big] - \sigma^{-2} E[\mathbf{1}_{\{N>x\}}N] \\ &= E\bigg[\mathbf{1}_{\{F>x\}}\bigg(F\bigg(\frac{q}{w} - \sigma^{-2}\bigg) - \big\langle Dw^{-1}, DF\big\rangle_{\mathfrak{H}}\bigg)\bigg] \\ &\quad + \sigma^{-2} E[F\mathbf{1}_{\{F>x\}} - N\mathbf{1}_{\{N>x\}}] \\ &= A_1 + A_2. \end{aligned} \tag{4.7}$$

For the first term A_1, Hölder's inequality implies

$$\begin{aligned} |A_1| &= \bigg|E\bigg[\mathbf{1}_{\{F>x\}}\bigg(F\bigg(\frac{q}{w} - \sigma^{-2}\bigg) - \big\langle Dw^{-1}, DF\big\rangle_{\mathfrak{H}}\bigg)\bigg]\bigg| \\ &\leqslant \sigma^{-2} E\big[\big|Fw^{-1}\big(w - q\sigma^2\big)\big|\big] + 2E\big[w^{-\frac{3}{2}}\big\|D^2F \otimes_1 DF\big\|_{\mathfrak{H}}\big] \\ &\leqslant \sigma^{-2}\big\|w^{-1}\big\|_3 \|F\|_3 \big\|w - q\sigma^2\big\|_3 + 2\big\|w^{-\frac{3}{2}}\big\|_2 \big\|\big\|D^2F \otimes_1 DF\big\|_{\mathfrak{H}}\big\|_2. \end{aligned}$$

Note that (2.12) implies

$$\big\|w - q\sigma^2\big\|_3 \leqslant C\big\|w - q\sigma^2\big\|_2$$

and $\|F\|_3 \leqslant C\|F\|_2 = C\sigma$. Combining these estimates with (4.3) we obtain

$$|A_1| \leqslant C\big(\sigma^{-1}\big\|w^{-1}\big\|_3 + \big\|w^{-1}\big\|^{\frac{3}{2}}_3\big)\big\|w - q\sigma^2\big\|_2. \tag{4.8}$$

For the second term A_2, applying Lemma 2.2 to the function $h(z) = z\mathbf{1}_{\{z>x\}}$, which satisfies $|h(z)| \leqslant |z|$, we have

$$|A_2| = \sigma^{-2}\big|E[F\mathbf{1}_{\{F>x\}} - N\mathbf{1}_{\{N>x\}}]\big|$$
$$\leqslant C\sigma^{-3}\big\|\sigma^2 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\big\|_2 \leqslant C\sigma^{-3}\big\|q\sigma^2 - w\big\|_2. \tag{4.9}$$

Combining (4.7) with (4.8)–(4.9) we obtain

$$\sup_{x\in\mathbb{R}}\big|f_F(x) - \phi(x)\big| \leqslant C\big(\sigma^{-1}\big\|w^{-1}\big\|_3 + \big\|w^{-1}\big\|_3^{\frac{3}{2}} + \sigma^{-3}\big)\big\|w - q\sigma^2\big\|_2.$$

Then (4.2) follows from (2.13). This completes the proof. □

Using the estimates shown in Theorem 4.1 we can deduce the following uniform convergence and convergence in L^p of densities for a sequence of random variables in a fixed qth Wiener chaos.

Corollary 4.3. *Let $\{F_n\}_{n\in\mathbb{N}}$ be a sequence of random variables in the qth Wiener chaos with $q \geqslant 2$. Set $\sigma_n^2 = E[F_n^2]$ and assume that $\lim_{n\to\infty}\sigma_n^2 = \sigma^2$, $0 < \delta \leqslant \sigma_n^2 \leqslant K$ for all n, $\lim_{n\to\infty} E[F_n^4] = 3\sigma^4$ and*

$$M := \sup_n\big(\mathbb{E}\big[\|DF_n\|_{\mathfrak{H}}^{-6}\big]\big)^{1/6} < \infty. \tag{4.10}$$

Let $\phi(x)$ be the density of the law $N(0,\sigma^2)$. Then, each F_n admits a density $f_{F_n} \in C(\mathbb{R})$ and there exists a constant C depending only on q, σ, δ and M such that

$$\sup_{x\in\mathbb{R}}\big|f_{F_n}(x) - \phi(x)\big| \leqslant C\big(\big|E\big[F_n^4\big] - 3\sigma_n^4\big|^{\frac{1}{2}} + |\sigma_n - \sigma|\big). \tag{4.11}$$

Furthermore, for any $p \geqslant 1$ and $\alpha \in (\frac{1}{2}, p)$,

$$\|f_{F_n} - \phi\|_{L^p(\mathbb{R})} \leqslant C\big(\big|E\big[F_n^4\big] - 3\sigma_n^4\big|^{\frac{1}{2}} + |\sigma_n - \sigma|\big)^{\frac{p-\alpha}{p}}, \tag{4.12}$$

where C is a constant depending on q, σ, M, p, α and K.

Proof. Let $\phi_n(x)$ be the density of $N(0,\sigma_n^2)$. Then Theorem 4.1 implies that

$$\sup_{x\in\mathbb{R}}\big|f_{F_n}(x) - \phi_n(x)\big| \leqslant C\big|E\big[F_n^4\big] - 3\sigma_n^4\big|^{\frac{1}{2}}.$$

On the other hand, if $N_n \sim N(0,\sigma_n^2)$, it is easy to see that

$$\sup_{x\in\mathbb{R}}\big|\phi_n(x) - \phi(x)\big| \leqslant C|\sigma_n - \sigma|.$$

Then (4.11) follows from triangle inequality. To show (4.12), first notice that (3.2) implies

$$f_{F_n}(x) \leqslant C\big(1 \wedge |x|^{-2}\big).$$

Therefore, if $\alpha > \frac{1}{2}$ the function $(f_{F_n}(x) + \phi(x))^\alpha$ is integrable. Then, (4.12) follows from (4.11) and the inequality

$$\left|f_{F_n}(x) - \phi(x)\right|^p \leqslant \left|f_{F_n}(x) - \phi(x)\right|^{p-\alpha}\left(f_{F_n}(x) + \phi(x)\right)^\alpha. \qquad \square$$

4.2. Uniform estimates of derivatives of densities

In this subsection, we establish the uniform convergence for derivatives of densities of random variables to a normal distribution. We begin with the following theorem which estimates the uniform distance between the derivatives of the densities of a random variable F in the qth Wiener chaos and the normal law $N(0, E[F^2])$.

Theorem 4.4. *Let $m \geqslant 1$ be an integer. Let F be a random variable in the qth Wiener chaos, $q \geqslant 2$, with $E[F^2] = \sigma^2$ and $M_\beta := M_\beta(F) < \infty$ for some $\beta > 6m + 6(\lfloor \frac{m}{2} \rfloor \vee 1)$ (recall the definition of $M_\beta(F)$ in (3.22)). Let $\phi(x)$ be the density of $N \sim N(0, \sigma^2)$. Then F has a density $f_F(x) \in C^m(\mathbb{R})$ with derivatives given by (3.23). Moreover, for any $k = 1, \dots, m$*

$$\sup_{x\in\mathbb{R}}\left|f_F^{(k)}(x) - \phi^{(k)}(x)\right| \leqslant \sigma^{-k-3} C\sqrt{E\left[F^4\right] - 3\sigma^2},$$

where the constant C depends on q, β, m, σ and M_β with polynomial growth in σ and M_β.

To prove Theorem 4.4, we need some technical results. Recall the notation we introduced in (3.15), where we denote $\delta_u = \delta(u)$, $D_u\delta_u = \langle D\delta_u, u\rangle_{\mathfrak{H}}$.

Lemma 4.5. *Let F be a random variable in the qth Wiener chaos with $E[F^2] = \sigma^2$. Let $w = \|DF\|_{\mathfrak{H}}^2$ and $u = w^{-1}DF$.*

(i) *If $M_\beta(F) < \infty$ for some $\beta > 6$, then for any $1 < r \leqslant \frac{2\beta}{\beta+6}$*

$$\left\|\delta_u - \sigma^{-2}F\right\|_r \leqslant C\sigma^{-1}\left(M_\beta^3 \vee 1\right)\left\|q\sigma^2 - w\right\|_2. \tag{4.13}$$

(ii) *If $M_\beta(F) < \infty$ for some $\beta > 12$, then for any $1 < r < \frac{2\beta}{\beta+12}$*

$$\left\|D_u\delta_u - \sigma^{-2}\right\|_r \leqslant C\sigma^{-2}\left(M_\beta^6 \vee 1\right)\left\|q\sigma^2 - w\right\|_2, \tag{4.14}$$

where the constant C depends on σ.

Proof. Recall that $\delta_u = qFw^{-1} - D_{DF}w^{-1}$. Using Hölder's inequality and (A.3) we can write

$$\begin{aligned}\left\|\delta_u - \sigma^{-2}F\right\|_r &\leqslant \left\|\sigma^{-2}Fw^{-1}\left(q\sigma^2 - w\right)\right\|_r + \left\|D_{DF}w^{-1}\right\|_r \\ &\leqslant C\left(\sigma^{-2}\left\|Fw^{-1}\right\|_s + \left(M_\beta^3 \vee 1\right)\right)\left\|q\sigma^2 - w\right\|_2,\end{aligned}$$

provided $\frac{1}{r} = \frac{1}{s} + \frac{1}{2}$. By the hypercontractivity property (2.11) $\|F\|_\gamma \leqslant C_{q,\gamma}\|F\|_2$ for any $\gamma \geqslant 2$. Thus, by Hölder's inequality, if $\frac{1}{s} = \frac{1}{\gamma} + \frac{1}{p}$

$$\|Fw^{-1}\|_s \leqslant \|F\|_\gamma \|w^{-1}\|_p \leqslant C_{q,\gamma}\sigma M_{2p}^2.$$

Choosing p such that $2p < \beta$ we get (4.13).

We can compute $D_u\delta_u$ as

$$D_u\delta_u = qw^{-1} + qFw^{-1}D_{DF}w^{-1} - w^{-1}D_{DF}^2 w^{-1} - w^{-1}\langle D^2F, DF \otimes Dw^{-1}\rangle_{\mathfrak{H}}.$$

Applying Hölder's inequality we obtain

$$\begin{aligned}\|D_u\delta_u - \sigma^{-2}\|_r &\leqslant \|w^{-1}[\sigma^{-2}(q\sigma^2 - w) + qFD_{DF}w^{-1} - D_{DF}^2 w^{-1}]\|_r \\ &\leqslant \sigma^{-2}\|w^{-1}\|_{\frac{2r}{2-r}}\|q\sigma^2 - w\|_2 + C_\sigma\|w^{-1}\|_p(\|D_{DF}w^{-1}\|_s + \|D_{DF}^2 w^{-1}\|_s),\end{aligned}$$

if $\frac{1}{r} > \frac{1}{p} + \frac{1}{s}$. Then, using (A.3) and (A.4) with $k = 2$ and assuming that $s < \frac{2\beta}{\beta+8}$ and that $2p < \beta$ we obtain (4.14). □

Proof of Theorem 4.4. Proposition 3.6 implies that $f_F(x) \in C^{m-1}(\mathbb{R})$ and for $k = 0, 1, \dots, m-1$,

$$f_F^{(k)}(x) = (-1)^k E[\mathbf{1}_{\{F>x\}}G_{k+1}],$$

where $G_0 = 1$ and $G_{k+1} = \delta(G_k u) = G_k\delta(u) - \langle DG_k, u\rangle_{\mathfrak{H}}$. From (3.24),

$$\phi^{(k)}(x) = (-1)^k E[\mathbf{1}_{\{N>x\}}H_{k+1}(\sigma^{-2}, \sigma^{-2}N)].$$

Then, the identity $G_{k+1} = H_{k+1}(D_u\delta_u, \delta_u) + T_{k+1}$ (see formula (3.18)), suggests the following triangle inequality

$$\begin{aligned}|f_F^{(k)}(x) - \phi^{(k)}(x)| &= |E[\mathbf{1}_{\{F>x\}}G_{k+1} - \mathbf{1}_{\{N>x\}}H_{k+1}(\sigma^{-2}, \sigma^{-2}N)]| \\ &\leqslant |E[\mathbf{1}_{\{F>x\}}G_{k+1} - \mathbf{1}_{\{F>x\}}H_{k+1}(\sigma^{-2}, \sigma^{-2}F)]| \\ &\quad + |E[\mathbf{1}_{\{F>x\}}H_{k+1}(\sigma^{-2}, \sigma^{-2}F) - \mathbf{1}_{\{N>x\}}H_{k+1}(\sigma^{-2}, \sigma^{-2}N)]| \\ &= A_1 + A_2.\end{aligned}$$

We first estimate the term A_2. Note that $\|F\|_{2k+2} \leqslant C_{q,k}\|F\|_2 = C_{q,k}\sigma$ by the hypercontractivity property (2.11). Applying Lemma 2.2 with $h(z) = \mathbf{1}_{\{z>x\}}H_{k+1}(\sigma^{-2}, \sigma^{-2}z)$, which satisfies $|h(z)| \leqslant C_k(|z|^{k+1} + \sigma^{-k-1})$, we obtain

$$\begin{aligned}A_2 &= |E[h(F) - h(N)]| \\ &\leqslant C_{q,k}\sigma^{-2}|\sigma^k + 4\sigma^{-k-1}|\|\sigma^2 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\|_2 \\ &\leqslant C_{q,k,\sigma}\sigma^{-k-3}\|q\sigma^2 - w\|_2,\end{aligned} \tag{4.15}$$

where in the second inequality we used the fact that $\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}} = \frac{w}{q}$.

For the term A_1, Lemma 3.4 implies

$$A_1 \leqslant E\big[\big|H_{k+1}(D_u\delta_u,\delta_u) - H_{k+1}\big(\sigma^{-2},\sigma^{-2}F\big)\big|\big] + E\big[|T_{k+1}|\big]. \tag{4.16}$$

To proceed with the first term above, applying (3.17) we have

$$\begin{aligned}
&\big|H_{k+1}(D_u\delta_u,\delta_u) - H_{k+1}\big(\sigma^{-2},\sigma^{-2}F\big)\big| \\
&\quad\leqslant \sum_{0\leqslant i\leqslant \lfloor (k+1)/2\rfloor} |c_{k,i}|\big|\delta_u^{k+1-2i}(D_u\delta_u)^i - \big(\sigma^{-2}F\big)^{k+1-2i}\sigma^{-2i}\big| \\
&\quad\leqslant \sum_{0\leqslant i\leqslant \lfloor (k+1)/2\rfloor} |c_{k,i}| \\
&\qquad\times\big[\big|\delta_u^{k+1-2i} - \big(\sigma^{-2}F\big)^{k+1-2i}\big||D_u\delta_u|^i + \big|\sigma^{-2}F\big|^{k+1-2i}\big|(D_u\delta_u)^i - \sigma^{-2i}\big|\big].
\end{aligned} \tag{4.17}$$

Using the fact that $|x^k - y^k| \leqslant C_k|x-y|\sum_{0\leqslant j\leqslant k-1}|x|^{k-1-j}|y|^j$ and applying Hölder's inequality and the hypercontractivity property (2.11) we obtain

$$\begin{aligned}
&E\big[\big|\delta_u^{k+1-2i} - \big(\sigma^{-2}F\big)^{k+1-2i}\big||D_u\delta_u|^i\big] \\
&\quad\leqslant C_k E\Big[\big|\delta_u - \sigma^{-2}F\big||D_u\delta_u|^i \sum_{0\leqslant j\leqslant k-2i}|\delta_u|^{k-2i-j}\big|\sigma^{-2}F\big|^j\Big] \\
&\quad\leqslant C_{q,k,\sigma}\big\|\delta_u - \sigma^{-2}F\big\|_r\|D_u\delta_u\|_s^i \sum_{0\leqslant j\leqslant k-2i}\|\delta_u\|_p^{k-2i-j}\sigma^{-j},
\end{aligned} \tag{4.18}$$

provided $1 \geqslant \frac{1}{r} + \frac{i}{s} + \frac{k-2i-j}{p}$, which is implied by $1 \geqslant \frac{1}{r} + \frac{i}{s} + \frac{k-2i}{p}$. In order to apply the estimates (4.13), (A.12) (with $k=1$) and (A.11) we need $\frac{1}{r} > \frac{3}{\beta} + \frac{1}{2}$, $\frac{1}{s} > \frac{6}{\beta}$ and $\frac{1}{p} > \frac{3}{\beta}$, respectively. These are possible because $\beta > 6k+6$. Then we obtain an estimate of the form

$$E\big[\big|\delta_u^{k+1-2i} - \big(\sigma^{-2}F\big)^{k+1-2i}\big||D_u\delta_u|^i\big] \leqslant C_{q,k,\sigma}\sigma^{-k}\big(M_\beta^{3k+3}\vee 1\big)\big\|q\sigma^2 - w\big\|_2. \tag{4.19}$$

Similarly,

$$\begin{aligned}
&E\big[\big|\sigma^{-2}F\big|^{k+1-2i}\big|(D_u\delta_u)^i - \sigma^{-2i}\big|\big] \\
&\quad\leqslant C_{q,k,\sigma}E\Big[\big|\sigma^{-2}F\big|^{k+1-2i}\big|D_u\delta_u - \sigma^{-2}\big|\sum_{0\leqslant j\leqslant i-1}|D_u\delta_u|^j\sigma^{-2(i-1-j)}\Big] \\
&\quad\leqslant C_{q,k,\sigma}\sigma^{-(k-1)}\big\|D_u\delta_u - \sigma^{-2}\big\|_r\sum_{0\leqslant j\leqslant i-1}\|D_u\delta_u\|_s^j,
\end{aligned} \tag{4.20}$$

provided $1 > \frac{1}{r} + \frac{j}{s}$. In order to apply the estimates (4.14) and (A.12) (with $k=1$) we need $\frac{1}{r} > \frac{6}{\beta} + \frac{1}{2}$ and $\frac{1}{s} > \frac{6}{\beta}$, respectively. This implies

$$\frac{1}{r} + \frac{j}{s} > \frac{6+6j}{\beta} + \frac{1}{2}.$$

Notice that $6+6j \leqslant 6i \leqslant 3k+3$. So, we need $1 > \frac{1}{2} + \frac{3k+3}{\beta}$. The above r, s and p exist because $\beta > 6k+6$. Thus, we obtain an estimate of the form

$$E\big[\big|\sigma^{-2}F\big|^{k+1-2i}\big|(D_u\delta_u)^i - \sigma^{-2i}\big|\big] \leqslant C_{q,k,\sigma,\beta}\sigma^{-(k-1)}\big(M_\beta^{3k+3} \vee 1\big)\big\|q\sigma^2 - w\big\|_2. \tag{4.21}$$

Combining (4.19) and (4.21) we have

$$\begin{aligned} &E\big[\big|H_{k+1}(D_u\delta_u, \delta_u) - H_{k+1}\big(\sigma^{-2}, \sigma^{-2}F\big)\big|\big] \\ &\quad \leqslant C_{q,k,\sigma,\beta}\sigma^{-k}\big(M_\beta^{3k+3} \vee 1\big)\big\|q\sigma^2 - w\big\|_2. \end{aligned} \tag{4.22}$$

Applying Hölder's inequality to the expression (3.20) we obtain (assuming $k \geqslant 2$, otherwise $T_{k+1} = 0$)

$$E\big[|T_{k+1}|\big] \leqslant C_{q,k,\sigma,\beta} \sum_{(i_0,\dots,i_k)\in J_{k+1}} \|\delta_u\|_{r_0}^{i_0} \prod_{j=1}^{k} \big\|D_u^j\delta_u\big\|_{r_j}^{i_j},$$

where $1 = \frac{i_0}{r_0} + \sum_{j=1}^k \frac{i_j}{r_j}$. From property (b) in Lemma 3.5 there is at least one factor of the form $\|D_u^j\delta_u\|_{s_j}$ with $j \geqslant 2$. We apply the estimate (A.13) to one of these factors, and the estimate (A.12) to all the remaining factors. We also use the estimate (A.11) to control $\|\delta_u\|_{r_0}$. Notice that

$$1 = \frac{i_0}{r_0} + \sum_{j=1}^{k} \frac{i_j}{r_j} > \frac{3i_0}{\beta} + \sum_{j=1}^{k} \frac{i_j(3j+3)}{\beta} + \frac{1}{2},$$

and, on the other hand, using properties (a) and (c) in Lemma 3.5

$$\frac{3i_0}{\beta} + \sum_{j=1}^{k} \frac{i_j(3j+3)}{\beta} + \frac{1}{2} \leqslant \frac{3k + 3\lfloor\frac{k}{2}\rfloor}{\beta} + \frac{1}{2}.$$

We can choose the r_j's satisfying the above properties because $\beta > 6k + 6\lfloor\frac{k}{2}\rfloor$, and we obtain

$$E|T_{k+1}| \leqslant C_{q,k,\sigma,\beta}\big(M_\beta^{3k+3\lfloor\frac{k}{2}\rfloor} \vee 1\big)\big\|q\sigma^2 - w\big\|_2. \tag{4.23}$$

Combining (4.22) and (4.23) we complete the proof. □

Corollary 4.6. *Fix an integer $m \geqslant 1$. Let $\{F_n\}_{n\in\mathbb{N}}$ be a sequence of random variables in the qth Wiener chaos with $q \geqslant 2$ and $E[F_n^2] = \sigma_n^2$. Assume $\lim_{n\to\infty}\sigma_n = \sigma$, $0 < \delta \leqslant \sigma_n^2 \leqslant K$ for all n, $\lim_{n\to\infty} E[F_n^4] = 3\sigma^4$ and*

$$M := \sup_n \big(\mathbb{E}\big[\|DF_n\|_{\mathfrak{H}}^{-\beta}\big]\big)^{\frac{1}{\beta}} < \infty \tag{4.24}$$

for some $\beta > 6(m) + 6(\lfloor \frac{m}{2} \rfloor \vee 1)$. *Let* $\phi(x)$ *be the density of* $N(0, \sigma^2)$. *Then, each* F_n *admits a probability density function* $f_{F_n} \in C^m(\mathbb{R})$ *with derivatives given by* (3.23) *and for any* $k = 1, \dots, m$,

$$\sup_{x\in\mathbb{R}} \left| f_{F_n}^{(k)}(x) - \phi^{(k)}(x) \right| \leqslant C\left(\sqrt{E\left[F_n^4\right] - 3\sigma_n^4} + |\sigma_n - \sigma| \right),$$

where the constant C *depends only on* $q, m, \beta, M, \sigma, \delta$ *and* K.

Proof. Let $\phi_n(x)$ be the density of $N(0, \sigma_n^2)$. Then Theorem 4.4 implies that

$$\sup_{x\in\mathbb{R}} \left| f_{F_n}^{(k)}(x) - \phi_n^{(k)}(x) \right| \leqslant C_{q,m,\beta,M,\sigma} \sqrt{E\left[F_n^4\right] - 3\sigma_n^4}.$$

On the other hand, by the mean value theorem we can write

$$\left| \phi_n^{(k)}(x) - \phi^{(k)}(x) \right| \leqslant |\sigma_n - \sigma| \sup_{\gamma\in[\frac{\sigma}{2}, 2\sigma]} \left| \partial_\gamma \phi_\gamma^{(k)}(x) \right| = \frac{1}{2} |\sigma_n - \sigma| \sup_{\gamma\in[\frac{\sigma}{2}, 2\sigma]} \gamma \left| \phi_\gamma^{(k+2)}(x) \right|,$$

where $\phi_\gamma(x)$ is the density of the law $N(0, \gamma^2)$. Then, using the expression

$$\phi_\gamma^{(k+2)}(x) = E\left[\mathbf{1}_{N>x} H_{k+3}\left(\gamma^{-2}, \gamma^{-2} Z\right)\right],$$

where $Z \sim N(0, \gamma^2)$ and the explicit form of $H_{k+3}(\lambda, x)$, we obtain

$$\sup_{\gamma\in[\frac{\sigma}{2}, 2\sigma]} \gamma \left| \phi_\gamma^{(k+2)}(x) \right| \leqslant C_{k,\sigma}.$$

Therefore,

$$\sup_{x\in\mathbb{R}} \left| \phi_n^{(k)}(x) - \phi^{(k)}(x) \right| \leqslant C_{k,\sigma} |\sigma_n - \sigma|.$$

This completes the proof. □

5. Random vectors in Wiener chaos

5.1. Main result

In this section, we study the multidimensional counterpart of Theorem 4.6. We begin with a density formula for a smooth random vector.

A random vector $F = (F_1, \dots, F_d)$ in $\mathbb{D}^\infty$ is called *non-degenerate* if its *Malliavin matrix* $\gamma_F = (\langle DF_i, DF_j \rangle_{\mathfrak{H}})_{1\leqslant i,j\leqslant d}$ is invertible a.s. and $(\det \gamma_F)^{-1} \in \bigcap_{p\geqslant 1} L^p(\Omega)$. For any multi-index

$$\beta = (\beta_1, \beta_2, \dots, \beta_k) \in \{1, 2, \dots, d\}^k$$

of length $k \geqslant 1$, the symbol ∂_β stands for the partial derivative $\frac{\partial^k}{\partial x_{\beta_1}\cdots\partial x_{\beta_k}}$. For β of length 0 we make the convention that $\partial_\beta f = f$. We denote by $\mathcal{S}(\mathbb{R}^d)$ the Schwartz space of rapidly decreasing smooth functions, that is, the space of all infinitely differentiable functions $f : \mathbb{R}^d \to \mathbb{R}$ such that $\sup_{x\in\mathbb{R}^d} |x|^m |\partial_\beta f(x)| < \infty$ for any nonnegative integer m and for all multi-index β. The following lemma (see Nualart [22, Proposition 2.1.5]) gives an explicit formula for the density of F.

Lemma 5.1. *Let $F = (F_1, \dots, F_d)$ be a non-degenerate random vector. Then, F has a density $f_F \in \mathcal{S}(\mathbb{R}^d)$, and f_F and its partial derivative $\partial_\beta f_F$, for any multi-index $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ of length $k \geqslant 0$, are given by*

$$f_F(x) = E\big[\mathbf{1}_{\{F>x\}} H_{(1,2,\dots,d)}(F)\big], \tag{5.1}$$

$$\partial_\beta f_F(x) = (-1)^k E\big[\mathbf{1}_{\{F>x\}} H_{(1,2,\dots,d,\beta_1,\beta_2,\dots,\beta_k)}(F)\big], \tag{5.2}$$

where $\mathbf{1}_{\{F>x\}} = \prod_{i=1}^d \mathbf{1}_{\{F_i>x_i\}}$ and the elements $H_\beta(F)$ are recursively defined by

$$\begin{cases} H_\beta(F) = 1, & \text{if } k = 0; \\ H_{(\beta_1,\beta_2,\dots,\beta_k)}(F) = \displaystyle\sum_{j=1}^d \delta\big(H_{(\beta_1,\beta_2,\dots,\beta_{k-1})}(F)\big(\gamma_F^{-1}\big)^{\beta_1 j} DF_j\big), & \text{if } k \geqslant 1. \end{cases} \tag{5.3}$$

Fix d natural numbers $1 \leqslant q_1 \leqslant \dots \leqslant q_d$. We will consider a random vector of multiple stochastic integrals: $F = (F_1, \dots, F_d) = (I_{q_1}(f_1), \dots, I_{q_d}(f_d))$, where $f_i \in \mathfrak{H}^{\odot q_i}$. Denote

$$V = \big(E[F_i F_j]\big)_{1\leqslant i,j\leqslant d}, \qquad Q = \operatorname{diag}(q_1, \dots, q_d)$$
$$\text{(diagonal matrix of elements } q_1, \dots, q_d). \tag{5.4}$$

Along this section, we denote by $N = (N_1, \dots, N_d)$ a standard normal vector given by $N_i = I_1(h_i)$, where $h_i \in \mathfrak{H}$ are orthonormal. We denote by I the d dimensional identity matrix, and by $|\cdot|$ the Hilbert–Schmidt norm of a matrix. The following is the main theorem of this section.

Theorem 5.2. *Let $F = (F_1, \dots, F_d) = (I_{q_1}(f_1), \dots, I_{q_d}(f_d))$ be non-degenerate and let ϕ be the density of N. Then for any multi-index β of length $k \geqslant 0$, the density f_F of F satisfies*

$$\sup_{x\in\mathbb{R}^d} \big|\partial_\beta f_F(x) - \partial_\beta \phi(x)\big| \leqslant C\left(|V - I| + \sum_{1\leqslant j\leqslant d} \sqrt{E\big[F_j^4\big] - 3\big(E\big[F_j^2\big]\big)^2}\right), \tag{5.5}$$

where the constant C depends on d, V, Q, k and $\|(\det \gamma_F)^{-1}\|_{(k+4)2^{k+3}}$.

Proof. Note that $\partial_\beta \phi(x) = (-1)^k E[\mathbf{1}_{\{N>x\}} H_{(1,2,\dots,d,\beta_1,\beta_2,\dots,\beta_k)}(N)]$. Then, in order to estimate the difference between $\partial_\beta f_{F_n}$ and $\partial_\beta \phi$, it suffices to estimate

$$E\big[\mathbf{1}_{\{F>x\}} H_\beta(F)\big] - E\big[\mathbf{1}_{\{N>x\}} H_\beta(N)\big]$$

for all multi-index β of length k for all $k \geqslant d$.

Fix a multi-index β of length k for some $k \geqslant d$. For the above standard normal random vector N, we have $\gamma_N = I$ and $\delta(DN_i) = N_i$. We can deduce from the expression (5.3) that $H_\beta(N) = g_\beta(N)$, where $g_\beta(x)$ is a polynomial on $\mathbb{R}^d$ (see Remark 5.4). Then,

$$\begin{aligned}
&\big|E\big[\mathbf{1}_{\{F>x\}}H_\beta(F)\big] - E\big[\mathbf{1}_{\{N>x\}}H_\beta(N)\big]\big| \\
&\quad \leqslant \big|E\big[\mathbf{1}_{\{F>x\}}g_\beta(F)\big] - E\big[\mathbf{1}_{\{N>x\}}g_\beta(N)\big]\big| + E\big[\big|H_\beta(F) - g_\beta(F)\big|\big] \\
&\quad = A_1 + A_2.
\end{aligned} \tag{5.6}$$

The term $A_1 = |E[\mathbf{1}_{\{F>x\}}g_\beta(F) - \mathbf{1}_{\{N>x\}}g_\beta(N)]|$ will be studied in Subsection 5.3 by using the multivariate Stein's method. Proposition 5.10 will imply that A_1 is bounded by the right-hand side of (5.5).

Consider the term $A_2 = E[|H_\beta(F) - g_\beta(F)|]$. We introduce an auxiliary term $K_\beta(F)$, which is defined similar to $H_\beta(F)$ with γ_F^{-1} replaced by $(VQ)^{-1}$. That is, for any multi-index $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ of length $k \geqslant 0$, we define

$$\begin{cases}
K_\beta(F) = 1 & \text{if } k = 0; \\
K_\beta(F) = \delta\big(K_{(\beta_1,\beta_2,\dots,\beta_{k-1})}(F)\big((VQ)^{-1}DF\big)_{\beta_k}\big) & \text{if } k \geqslant 1.
\end{cases} \tag{5.7}$$

We have

$$A_2 \leqslant E\big[\big|H_\beta(F) - K_\beta(F)\big|\big] + E\big[\big|K_\beta(F) - g_\beta(F)\big|\big] =: A_3 + A_4. \tag{5.8}$$

Lemma 5.11 below shows that the term $A_3 = E[|H_\beta(F) - K_\beta(F)|]$ is bounded by the right-hand of (5.5).

It remains to estimate A_4. For this we need the following lemma which provides an explicit expression for the term $K_\beta(F)$. Before stating this lemma we need to introduce some notation. For any multi-index $\beta = (\beta_1, \beta_2, \dots, \beta_k)$, $k \geqslant 1$, denote by $\widehat{\beta}_{i_1\dots i_m}$ the multi-index obtained from β after taking away the elements $\beta_{i_1}, \beta_{i_2}, \dots, \beta_{i_m}$. For example, $\widehat{\beta}_{14} = (\beta_2, \beta_3, \beta_5, \dots, \beta_k)$. For any d dimensional vector G we denote by G_β the product $G_{\beta_1}G_{\beta_2}\cdots G_{\beta_k}$ and set $G_\beta = 1$ if the length of β is 0. Denote by $(S_k^m; 0 \leqslant m \leqslant \lfloor\frac{k}{2}\rfloor)$ the following sets

$$\begin{cases}
S_k^{-1} = S_k^0 = \varnothing, \\
S_k^m = \left\{ \begin{array}{l} \{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\} \in \{1,2,\dots,k\}^{2m}\colon \\ i_{2l-1} < i_{2l} \text{ for } 1 \leqslant l \leqslant m \text{ and } i_l \neq i_j \text{ if } l \neq j \end{array} \right\}.
\end{cases} \tag{5.9}$$

For each element $\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\} \in S_k^m$, we emphasize that the m pairs of indices are unordered. In other words, for $m \geqslant 1$, the set S_k^m can be viewed as the set of all partitions of $\{1,2,\dots,k\}$ into m pairs and $k-2m$ singletons.

Denote $M = V^{-1}\gamma_F V^{-1}Q^{-1}$ for V and Q given in (5.4) and denote M_{ij} the (i,j)th entry of M. Denote by D_{β_i} the Malliavin derivative in the direction of $(V^{-1}Q^{-1}DF)_{\beta_i} = V^{-1}Q^{-1}DF_{\beta_i}$, that is,

$$D_{\beta_i}G = \big\langle DG, \big(V^{-1}Q^{-1}DF\big)_{\beta_i}\big\rangle_{\mathfrak{H}} \tag{5.10}$$

for any random variable $G \in \mathbb{D}^{1,2}$.

Lemma 5.3. *Let F be a non-degenerate random vector. For a multi-index $\beta = (\beta_1, \dots, \beta_k)$ of length $k \geqslant 0$, $K_\beta(F)$ defined by (5.7) can be computed as follows:*

$$K_\beta(F) = G_\beta(F) + T_\beta(F), \tag{5.11}$$

where

$$G_\beta(F) = \sum_{m=0}^{\lfloor k/2 \rfloor} (-1)^m \sum_{\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\}\in S_k^m} \left(V^{-1}F\right)_{\widehat{\beta}_{i_1\cdots i_{2m}}} M_{\beta_{i_1}\beta_{i_2}} \cdots M_{\beta_{i_{2m-1}}\beta_{2m}}, \tag{5.12}$$

and $T_\beta(F)$ are defined recursively by

$$\begin{aligned} T_\beta(F) &= \left(V^{-1}F\right)_{\beta_k} T_{\widehat{\beta}_k}(F) - D_{\beta_k} T_{\widehat{\beta}_k}(F) \\ &\quad - \sum_{\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\}\in S_{k-1}^m} \left(V^{-1}F\right)_{\widehat{\beta}_{k i_1\cdots i_{2m}}} D_{\beta_k}\left(M_{\beta_{i_1}\beta_{i_2}} \cdots M_{\beta_{i_{2m-1}}\beta_{2m}}\right), \end{aligned} \tag{5.13}$$

for $k \geqslant 2$ and $T_1(F) = T_2(F) = 0$.

Proof. For simplicity, we write K_β, G_β and T_β for $K_\beta(F)$, $G_\beta(F)$ and $T_\beta(F)$, respectively. By using the fact that $\delta(((VQ)^{-1}DF)_{\beta_i}) = (V^{-1}F)_{\beta_i}$ we obtain from (5.7) that

$$K_\beta = \left(V^{-1}F\right)_{\beta_k} K_{\widehat{\beta}_k} - D_{\beta_k} K_{\widehat{\beta}_k}. \tag{5.14}$$

If $k = 1$, namely, $\beta = (\beta_1)$, then

$$K_\beta = \left(V^{-1}F\right)_{\beta_1} = G_\beta.$$

If $k = 2$, namely, $\beta = (\beta_1, \beta_2)$, then

$$K_\beta = \left(V^{-1}F\right)_\beta - M_{\beta_1\beta_2} = G_\beta.$$

Hence, the identity (5.11) is true for $k = 1, 2$. Assume now (5.11) is true for all multi-index of length less than or equal to k. Let $\beta = (\beta_1, \dots, \beta_{k+1})$. Then, (5.14) implies

$$K_\beta = \left(V^{-1}F\right)_{\beta_{k+1}} \left(G_{\widehat{\beta}_{k+1}} + T_{\widehat{\beta}_{k+1}}\right) - D_{\beta_{k+1}}\left(G_{\widehat{\beta}_{k+1}} + T_{\widehat{\beta}_{k+1}}\right). \tag{5.15}$$

Noticing that

$$D_{\beta_{k+1}}\left(V^{-1}F\right)_{\widehat{\beta}_{(k+1)i_1\cdots i_{2m}}} = \sum_{j\in\{1,\dots,k\}\setminus\{i_1,\dots,i_{2m}\}} \left(V^{-1}F\right)_{\widehat{\beta}_{(k+1)ji_1\cdots i_{2m}}} M_{\beta_j\beta_{k+1}},$$

we have

$$D_{\beta_{k+1}} G_{\widehat{\beta}_{k+1}} = B_\beta + \sum_{m=0}^{\lfloor k/2 \rfloor} (-1)^m \sum_{\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\} \in S_k^m} \left(V^{-1}F\right)_{\widehat{\beta}_{(k+1)i_1\cdots i_{2m}}} D_{\beta_{k+1}} (M_{\beta_{i_1}\beta_{i_2}} \cdots M_{\beta_{i_{2m-1}}\beta_{2m}}), \tag{5.16}$$

where we let

$$B_\beta = \sum_{m=0}^{\lfloor k/2 \rfloor} (-1)^m \sum_{\substack{\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\} \in S_k^m, \\ j \in \{1,\dots,k\} \setminus \{i_1,\dots,i_{2m}\}}} \left(V^{-1}F\right)_{\widehat{\beta}_{j(k+1)i_1\cdots i_{2m}}} M_{\beta_j \beta_{k+1}} M_{\beta_{i_1}\beta_{i_2}} \cdots M_{\beta_{i_{2m-1}}\beta_{i_{2m}}}.$$

Substituting the expression (5.16) for $D_{\beta_{k+1}} G_{\widehat{\beta}_{k+1}}$ into (5.15) and using (5.13) we obtain

$$K_\beta = \left(V^{-1}F\right)_{\beta_{k+1}} G_{\widehat{\beta}_{k+1}} - B_\beta + T_\beta.$$

To arrive at (5.11) it remains to verify

$$G_\beta = \left(V^{-1}F\right)_{\beta_{k+1}} G_{\widehat{\beta}_{k+1}} - B_\beta. \tag{5.17}$$

Introduce the following notation

$$C_{k+1}^m = \left\{\left\{(i_1,i_2),\dots,(i_{2m-3},i_{2m-2}),(j,k+1)\right\}: \left\{(i_1,i_2),\dots,(i_{2m-3},i_{2m-2})\right\} \in S_k^{m-1}\right\} \tag{5.18}$$

for $1 \leqslant m \leqslant \lfloor \frac{k}{2} \rfloor$. Then, S_{k+1}^m can be decomposed as follows

$$S_{k+1}^m = S_k^m \cup C_{k+1}^m. \tag{5.19}$$

We consider first the case when k is even. In this case, noticing that for any element in $\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\} \in S_k^{\lfloor \frac{k}{2} \rfloor}$, $\{1,\dots,k\} \setminus \{i_1,\dots,i_{2m}\} = \varnothing$. For any collection of indices $i_1,\dots,i_{2m} \subset \{1,2,\dots,k\}$, we set

$$\Phi_{i_1\dots i_{2m}} = \left(V^{-1}F\right)_{\widehat{\beta}_{i_1\cdots i_{2m}}} M_{\beta_{i_1}\beta_{i_2}} \cdots M_{\beta_{i_{2m-1}}\beta_{i_{2m}}}.$$

Then, we have

$$-B_\beta = \sum_{m=0}^{\lfloor \frac{k}{2} \rfloor - 1} (-1)^{m+1} \sum_{\substack{\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\} \in S_k^m, \\ j \in \{1,\dots,k\} \setminus \{i_1,\dots,i_{2m}\}}} \Phi_{j(k+1)i_1\dots i_{2m}}$$

$$= \sum_{m=1}^{\lfloor \frac{k}{2} \rfloor} (-1)^n \sum_{\substack{\{(i_1,i_2),\dots,(i_{2m-3},i_{2m-2})\}\in S_k^{m-1},\\ j\in\{1,\dots,k\}\setminus\{i_1,\dots,i_{2n-2}\}}} \Phi_{j(k+1)i_1\dots i_{2m-2}}$$

$$= \sum_{m=1}^{\lfloor \frac{k+1}{2} \rfloor} (-1)^m \sum_{\substack{\{(i_1,i_2),\dots,(i_{2m-3},i_{2m-2}),(j,k+1)\}\\ \in C_{k+1}^m}} \Phi_{j(k+1)i_1\dots i_{2m-2}}, \tag{5.20}$$

where in the last equality we used (5.18) and the fact that $\lfloor \frac{k}{2} \rfloor = \lfloor \frac{k+1}{2} \rfloor$ since k is even. Taking into account that $(V^{-1}F)_{\beta_{k+1}}(V^{-1}F)_{\widehat{\beta}_{(k+1)i_1\cdots i_{2m}}} = (V^{-1}F)_{\widehat{\beta}_{i_1\cdots i_{2m}}}$, we obtain from (5.12) that

$$\begin{aligned}
&\left(V^{-1}F\right)_{\beta_{k+1}} G_{\widehat{\beta}_{k+1}} \\
&= \sum_{m=0}^{\lfloor k/2 \rfloor} (-1)^m \sum_{\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\}\in S_k^m} \left(V^{-1}F\right)_{\widehat{\beta}_{i_1\cdots i_{2m}}} M_{\beta_{i_1}\beta_{i_2}} \cdots M_{\beta_{i_{2m-1}}\beta_{i_{2m}}}.
\end{aligned} \tag{5.21}$$

Now combining (5.20) and (5.21) with (5.19) and using again $\lfloor \frac{k}{2} \rfloor = \lfloor \frac{k+1}{2} \rfloor$ we obtain

$$\begin{aligned}
\left(V^{-1}F\right)_{\beta_{k+1}} G_{\widehat{\beta}_{k+1}} - B_\beta &= \sum_{m=0}^{\lfloor (k+1)/2 \rfloor} (-1)^m \sum_{\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\}\in S_k^m} \Phi_{i_1\dots i_{2m}} \\
&\quad + \sum_{m=1}^{\lfloor (k+1)/2 \rfloor} (-1)^m \sum_{\{(i_1,i_2),\dots,(i_{2m-3},i_{2m-2}),(j,k+1)\}\in C_{k+1}^m} \Phi_{i_1\dots i_{2m}} \\
&= \sum_{m=0}^{\lfloor (k+1)/2 \rfloor} (-1)^m \sum_{\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\}\in S_{k+1}^m} \Phi_{i_1\dots i_{2m}} \\
&= G_\beta
\end{aligned}$$

as desired. This verifies (5.17) for the case k is even. The case when k is odd can be verified similarly. Thus, we have proved (5.11) by induction. □

Remark 5.4. For the random vector $N \sim N(0, I)$, we have $\gamma_N = VQ = I$, so $H_\beta(N) = K_\beta(N)$. Then, it follows from Lemma 5.3 that $H_\beta(N) = K_\beta(N) = g_\beta(N)$ with the function $g_\beta(x) : \mathbb{R}^d \to \mathbb{R}$ given by

$$g_\beta(x) = \sum_{m=0}^{\lfloor k/2 \rfloor} (-1)^m \sum_{\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\}\in S_k^m} x_{\widehat{\beta}_{i_1\cdots i_{2m}}} \delta_{\beta_{i_1}\beta_{i_2}} \cdots \delta_{\beta_{i_{2m-1}}\beta_{2m}}, \tag{5.22}$$

where we used δ_{ij} to denote the Kronecker symbol (without confusion with the divergence operator). Notice that

$$g_\beta(x) = \prod_{i=1}^{d} H_{k_i}(x_i),$$

where H_{k_i} is the k_ith Hermite polynomial and for each $i = 1, \dots, d$, k_i is the number of components of β equal to i.

Let us return to the proof of Theorem 5.2 of estimating the term A_4. From (5.11) we can write

$$A_4 = E\big[\big|K_\beta(F) - g_\beta(F)\big|\big] \leqslant E\big[\big|G_\beta(F) - g_\beta(F)\big|\big] + E\big[\big|T_\beta(F)\big|\big]. \tag{5.23}$$

Observe from the expression (5.13) that $T_\beta(F)$ is the sum of terms of the following form

$$\big(V^{-1}F\big)_{\beta_{i_1}\beta_{i_2}\cdots\beta_{i_s}} D_{\beta_{k_1}} D_{\beta_{k_2}} \cdots D_{\beta_{k_t}} \left(\prod_{i}^{r} M_{\beta_{j_i}\beta_{l_i}}\right) \tag{5.24}$$

for some $\{i_1, \dots, i_s, k_1, \dots, k_t, j_1, l_1, \dots j_r, l_r\} \subset \{1, 2, \dots, k\}$ and $t \geqslant 1$. Applying Lemma 5.5 with (2.11) and (2.12) we obtain

$$E\big[\big|T_\beta(F)\big|\big] \leqslant C \sum_{1\leqslant l\leqslant d} \big\| \|DF_l\|_{\mathfrak{H}}^2 - q_l E\big[F_l^2\big]\big\|_2^{\frac{1}{2}}. \tag{5.25}$$

In order to compare $g_\beta(F)$ with $G_\beta(F)$, from (5.22) we can write $g_\beta(F)$ as

$$g_\beta(F) = \sum_{m=0}^{\lfloor k/2\rfloor} (-1)^m \sum_{\{(i_1,i_2),\dots,(i_{2m-1},i_{2m})\}\in S_k^m} F_{\widehat{\beta}_{i_1\cdots i_{2m}}} \delta_{\beta_{i_1}\beta_{i_2}} \cdots \delta_{\beta_{i_{2m-1}}\beta_{2m}}.$$

Then, it follows from hypercontractivity property (2.11) that

$$E\big[\big|G_\beta(F) - g_\beta(F)\big|\big] \leqslant C\big(\big|V^{-1} - I\big| + \|M - I\|_2\big),$$

where the constant C depends on k, V and Q. From $V^{-1} - I = V^{-1}(I - V)$ we have $|V^{-1} - I| \leqslant C|V - I|$, where C depends on V. We also have $M - I = V^{-1}(\gamma_F - VQ)V^{-1}Q^{-1} + V^{-1} - I$. Then, Lemma 5.5 implies that

$$\begin{aligned}\|M - I\|_2 &\leqslant C\big(\|\gamma_F - VQ\|_2 + \big|V^{-1} - I\big|\big)\\ &\leqslant C\bigg(\sum_{1\leqslant l\leqslant d} \big\| \|DF_l\|_{\mathfrak{H}}^2 - q_l E F_l^2\big\|_2 + |V - I|\bigg),\end{aligned}$$

where the constant C depends on k, V and Q. Therefore

$$E\big[\big|G_\beta(F) - g_\beta(F)\big|\big] \leqslant C\bigg(|V - I| + \sum_{1\leqslant l\leqslant d} \big\| \|DF_l\|_{\mathfrak{H}}^2 - q_l E F_l^2\big\|_2^{\frac{1}{2}}\bigg). \tag{5.26}$$

Combining it with (5.25) we obtain from (5.23) that

$$A_4 \leqslant C\Big(|V - I| + \sum_{1\leqslant l\leqslant d} \big\| \|DF_l\|_{\mathfrak{H}}^2 - q_l E F_l^2 \big\|_2^{\frac{1}{2}}\Big),$$

where the constant C depends on d, V, Q. This completes the estimation of the term A_4. □

5.2. Sobolev norms of γ_F^{-1}

In this subsection we estimate the Sobolev norms of γ_F^{-1}, the inverse of the Malliavin matrix γ_F for a random variable F of multiple stochastic integrals. We begin with the following estimate on the variance and Sobolev norms of $(\gamma_F)_{ij} = \langle DF_i, DF_j\rangle_{\mathfrak{H}}$, $1 \leqslant i, j \leqslant d$, following the approach of [13,16,19].

Lemma 5.5. *Let $F = I_p(f)$ and $G = I_q(g)$ with $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$ for $p, q \geqslant 1$. Then for all $k \geqslant 0$ there exists a constant $C_{p,q,k}$ such that*

$$\begin{aligned} &\big\| D^k\big(\langle DF, DG\rangle_{\mathfrak{H}} - \sqrt{pq}\,E[FG]\big)\big\|_2 \\ &\quad \leqslant C_{p,q,k}\big(\|F\|_2^2 + \|G\|_2^2\big)\big(\big\| \|DF\|_{\mathfrak{H}}^2 - pE\big[F^2\big]\big\|_2^{\frac{1}{2}} + \big\| \|DG\|_{\mathfrak{H}}^2 - pE\big[G^2\big]\big\|_2^{\frac{1}{2}}\big). \end{aligned} \tag{5.27}$$

Proof. Without lost of generality, we assume $p \leqslant q$. Applying (2.4) with the fact that $DI_p(f) = pI_{p-1}(f)$ we have

$$\begin{aligned} \langle DF, DG\rangle_{\mathfrak{H}} &= pq\big\langle I_{p-1}(f), I_{q-1}(g)\big\rangle_{\mathfrak{H}} \\ &= pq\sum_{r=0}^{p-1} r!\binom{p-1}{r}\binom{q-1}{r} I_{p+q-2-2r}(f\widetilde{\otimes_{r+1}}g) \\ &= pq\sum_{r=1}^{p}(r-1)!\binom{p-1}{r-1}\binom{q-1}{r-1} I_{p+q-2r}(f\widetilde{\otimes}_r g). \end{aligned} \tag{5.28}$$

Note that $E[FG] = 0$ if $p < q$ and $E[FG] = \langle f, g\rangle_{\mathfrak{H}^{\otimes p}} = f\widetilde{\otimes}_p g$ if $p = q$. Then

$$\langle DF, DG\rangle_{\mathfrak{H}} - \sqrt{pq}\,E(FG) = pq\sum_{r=1}^{p}(1-\delta_{qr})(r-1)!\binom{p-1}{r-1}\binom{q-1}{r-1} I_{p+q-2r}(f\widetilde{\otimes}_r g),$$

where δ_{qr} is again the Kronecker symbol. It follows that

$$\begin{aligned} &E\big[\langle DF, DG\rangle_{\mathfrak{H}} - \sqrt{pq}\,E[FG]\big]^2 \\ &\quad = p^2q^2\sum_{r=1}^{p}(1-\delta_{qr})(r-1)!^2\binom{p-1}{r-1}^2\binom{q-1}{r-1}^2(p+q-2r)!\|f\widetilde{\otimes}_r g\|_{\mathfrak{H}^{\otimes(p+q-2r)}}^2. \end{aligned} \tag{5.29}$$

Note that if $r < p \leqslant q$, then (see also [16, (6.2.7)])

$$\|f\widetilde{\otimes}_r g\|^2_{\mathfrak{H}^{\otimes(p+q-2r)}} \leqslant \|f\otimes_r g\|^2_{\mathfrak{H}^{\otimes(p+q-2r)}} = \langle f\otimes_{p-r} f, g\otimes_{q-r} g\rangle_{\mathfrak{H}^{\otimes 2r}} \leqslant \frac{1}{2}\left(\|f\otimes_{p-r} f\|^2_{\mathfrak{H}^{\otimes 2r}} + \|g\otimes_{q-r} g\|^2_{\mathfrak{H}^{\otimes 2r}}\right), \tag{5.30}$$

and if $r = p < q$,

$$\|f\widetilde{\otimes}_p g\|^2_{\mathfrak{H}^{\otimes(q-p)}} \leqslant \|f\otimes_p g\|^2_{\mathfrak{H}^{\otimes(q-p)}} \leqslant \|f\|^2_{\mathfrak{H}^{\otimes p}}\|g\otimes_{q-p} g\|_{\mathfrak{H}^{\otimes 2p}}. \tag{5.31}$$

From (4.5) and (4.6) it follows that

$$\big\|\|DF\|^2_{\mathfrak{H}} - pE\big[F^2\big]\big\|^2_2 = p^4\sum_{r=1}^{p-1}(r-1)!^2\binom{p-1}{r-1}^2(2p-2r)!\|f\otimes_r f\|^2_{\mathfrak{H}^{\otimes(2p-2r)}}. \tag{5.32}$$

Combining (5.29)–(5.32) we obtain

$$E\big[\langle DF, DG\rangle_{\mathfrak{H}} - \sqrt{pq}\,E[FG]\big]^2 \leqslant C_{p,q}\left(\big\|\|DF\|^2_{\mathfrak{H}} - pE\big[F^2\big]\big\|^2_2 + \|F\|^2_2\big\|\|DG\|^2_{\mathfrak{H}} - pE\big[G^2\big]\big\|_2\right).$$

Then (5.27) with $k=0$ follows from $\|\|DF\|^2_{\mathfrak{H}} - pE[F^2]\|_2 \leqslant C_p\|F\|^2_2$, which is implied by (2.12). From (5.28) we deduce

$$D^k\langle DF, DG\rangle_{\mathfrak{H}} = pq\sum_{r=1}^{p\wedge[(p+q-k)/2]}(r-1)!\binom{p-1}{r-1}\binom{q-1}{r-1}\frac{p+q-2r}{p+q-k-2r}I_{p+q-k-2r}(f\widetilde{\otimes}_r g).$$

Then it follows from (5.30)–(5.32) that

$$E\big\|D^k\langle DF, DG\rangle_{\mathfrak{H}}\big\|^2_{\mathfrak{H}^{\otimes k}} = p^2q^2\sum_{r=1}^{p\wedge[(p+q-k)/2]}(r-1)!^2\binom{p-1}{r-1}^2\binom{q-1}{r-1}^2\frac{(p+q-2r)!^2}{(p+q-k-2r)!}\|f\widetilde{\otimes}_r g\|^2_{\mathfrak{H}^{\otimes(p+q-2r)}} \leqslant C_{p,q}\left(\big\|\|DF\|^2_{\mathfrak{H}} - pE\big[F^2\big]\big\|^2_2 + \|F\|^2_2\big\|\|DG\|^2_{\mathfrak{H}} - pE\big[G^2\big]\big\|_2\right).$$

This completes the proof. □

The following lemma gives estimates on the Sobolev norms of the entries of γ_F^{-1}.

Lemma 5.6. *Let $F = (F_1,\dots,F_d) = (I_{q_1}(f_1),\dots,I_{q_d}(f_d))$ be non-degenerate and let $\gamma_F = (\langle DF_i, DF_j\rangle_{\mathfrak{H}})_{1\leqslant i,j\leqslant d}$. Set $V = (E[F_iF_j])_{1\leqslant i,j\leqslant d}$. Then for any real number $p>1$,*

$$\big\|\gamma_F^{-1}\big\|_p \leqslant C\big\|(\det\gamma_F)^{-1}\big\|_{2p}, \tag{5.33}$$

where the constant C depends on $q_1, \dots, q_d, d, p$ and V. Moreover, for any integer $k \geqslant 1$ and any real number $p > 1$

$$\|\gamma_F^{-1}\|_{k,p} \leqslant C \|(\det \gamma_F)^{-1}\|_{(k+2)2p}^{k+1} \sum_{i=1}^d \big\| \|DF_i\|_{\mathfrak{H}}^2 - q_i E[F_i^2]\big\|_2, \tag{5.34}$$

where the constant C depends on $q_1, \dots, q_d, d, p, k$ and V.

Proof. Let γ_F^* be the adjugate matrix of γ_F. Note that Hölder inequality and (2.12) imply

$$\big\|\langle DF_i, DF_j\rangle_{\mathfrak{H}}\big\|_p \leqslant \|DF_i\|_{2p}\|DF_j\|_{2p} \leqslant C_{V,p}$$

for all $1 \leqslant i, j \leqslant d$, $p \geqslant 1$. Applying Hölder's inequality we obtain that the p norm of γ_F^* is also bounded by a constant. A further application of Hölder's inequality to $\gamma_F^{-1} = (\det \gamma_F)^{-1}\gamma_F^*$ yields

$$\|\gamma_F^{-1}\|_p \leqslant \|(\det \gamma_F)^{-1}\|_{2p} \|\gamma_F^*\|_{2p} \leqslant C_{V,p} \|(\det \gamma_F)^{-1}\|_{2p}, \tag{5.35}$$

which implies (5.33).

Since F is non-degenerate, then (see [22, Lemma 2.1.6]) $(\gamma_F^{-1})_{ij}$ belongs to $\mathbb{D}^\infty$ for all i, j and

$$D(\gamma_F^{-1})_{ij} = -\sum_{m,n=1}^d (\gamma_F^{-1})_{im}(\gamma_F^{-1})_{nj} D(\gamma_F)_{mn}. \tag{5.36}$$

Then, applying Hölder's inequality we obtain

$$\begin{aligned} \|D(\gamma_F^{-1})\|_p &\leqslant \|\gamma_F^{-1}\|_{3p}^2 \|D\gamma_F\|_{3p} \\ &\leqslant C_{V,p} \|(\det \gamma_F)^{-1}\|_{6p}^2 \sum_{i=1}^d \big\| \|DF_i\|_{\mathfrak{H}}^2 - q_i E[F_i^2]\big\|_2, \end{aligned}$$

where in the second inequality we used (5.33) and

$$\|D\gamma_F\|_{3p} \leqslant C_{V,p}\|D\gamma_F\|_2 \leqslant C_{V,p} \sum_{i=1}^d \big\| \|DF_i\|_{\mathfrak{H}}^2 - q_i E[F_i^2]\big\|_2$$

for all $p \geqslant 1$, which follows from (2.12) and (5.27). This implies (5.34) with $k = 1$. For higher order derivatives, (5.34) follows from repeating the use of (5.36), (2.12) and (5.27). □

The following lemma estimates the difference $\gamma_F^{-1} - V^{-1}Q^{-1}$.

Lemma 5.7. *Let $F = (F_1, \dots, F_d) = (I_{q_1}(f_1), \dots, I_{q_d}(f_d))$ be a non-degenerate random vector with $1 \leqslant q_1 \leqslant \dots \leqslant q_d$ and $f_i \in \mathfrak{H}^{\odot q_i}$. Let γ_F be the Malliavin matrix of F. Recall the notation of V and Q in (5.4). Then, for every integer $k \geqslant 1$ and any real number $p > 1$ we have*

$$\left\|\gamma_F^{-1} - V^{-1}Q^{-1}\right\|_{k,p} \leqslant C\left\|(\det \gamma_F)^{-1}\right\|_{(k+2)2p}^{k+1} \sum_{1\leqslant l\leqslant d} \left\| \|DF_l\|_{\mathfrak{H}}^2 - q_l E\left[F_l^2\right]\right\|_2^{\frac{1}{2}}, \quad (5.37)$$

where the constant C depends on d, V, Q, p and k.

Proof. In view of Lemma 5.6, we only need to consider the case when $k = 0$ because V and Q are deterministic matrices. Note that

$$\gamma_F^{-1} - V^{-1}Q^{-1} = \gamma_F^{-1}\,(VQ - \gamma_F)V^{-1}Q^{-1}.$$

Then, applying Hölder's inequality we have

$$\left\|\gamma_F^{-1} - V^{-1}Q^{-1}\right\|_p \leqslant C_{V,Q}\left\|\gamma_F^{-1}\right\|_{2p}\|VQ - \gamma_F\|_{2p}.$$

Note that (2.12) and (5.27) with $k = 0$ imply

$$\|VQ - \gamma_F\|_{2p} \leqslant C_{V,Q,p}\|VQ - \gamma_F\|_2 \leqslant C_{V,Q,p}\sum_{i=1}^{d}\left\| \|DF_i\|_{\mathfrak{H}}^2 - q_i E\left[F_i^2\right]\right\|_2^{\frac{1}{2}}.$$

Then, applying (5.35) we obtain

$$\left\|\gamma_F^{-1} - V^{-1}Q^{-1}\right\|_p \leqslant C_{d,V,Q,p}\sum_{i=1}^{d}\left\| \|DF_i\|_{\mathfrak{H}}^2 - q_i E\left[F_i^2\right]\right\|_2^{\frac{1}{2}} \quad (5.38)$$

as desired. □

5.3. Technical estimates

In this subsection, we study the terms $A_1 = |E[h(F)] - E[h(N)]|$ in Eq. (5.6) and $A_3 = E[|H_\beta(F) - K_\beta(F)|]$ in (5.8). For A_1, we shall use the multivariate Stein's method to give an estimate for a large class of non-smooth test functions h.

Lemma 5.8. *Let $h : \mathbb{R}^d \to \mathbb{R}$ be an almost everywhere continuous function such that $|h(x)| \leqslant c(|x|^m + 1)$ for some $m, c > 0$. Let $F = (F_1, \dots, F_d)$ be non-degenerate with $E[F_i] = 0$, $1 \leqslant i \leqslant d$ and denote $N \sim N(0, I)$. Then there exists a constant $C_{m,c}$ depending on m and c such that*

$$\left|E\left[h(F)\right] - E\left[h(N)\right]\right| \leqslant C_{m,c}\left(\|F\|_{2m}^m + 1\right)\sum_{i,j,k=1}^{d}\left\|\delta\left(A_{ij}\left(\gamma_F^{-1}\right)_{jk}DF_k\right)\right\|_2, \quad (5.39)$$

where γ_F^{-1} is the inverse of the Malliavin matrix of F and

$$A_{ij} = \delta_{ij} - \langle DF_j, -DL^{-1}F_i\rangle_{\mathfrak{H}}. \tag{5.40}$$

Proof. For $\varepsilon > 0$, let

$$h_\varepsilon(x) = (\mathbf{1}_{\{|\cdot|<\frac{1}{\varepsilon}\}}h) * \rho_\varepsilon(x) = \int_{\mathbb{R}^d} \mathbf{1}_{|y|<\frac{1}{\varepsilon}} h(y)\rho_\varepsilon(x-y)\,dy,$$

where ρ_ε is the standard mollifier. That is, $\rho_\varepsilon(x) = \frac{1}{\varepsilon^d}\rho(\frac{x}{\varepsilon})$, where $\rho(x) = C\mathbf{1}_{\{|x|<1\}}\exp(\frac{1}{|x|^2-1})$ and the constant C is such that $\int_{\mathbb{R}^d}\rho(x)\,dx = 1$. Then h_ε is Lipschitz continuous. Hence, the solution f_ε to the following *Stein's equation*:

$$\Delta f_\varepsilon(x) - \langle x, \nabla f_\varepsilon(x)\rangle_{\mathbb{R}^d} = h_\varepsilon(x) - E\big[h_\varepsilon(N)\big] \tag{5.41}$$

exists and its derivative has the following expression [16, page 82]

$$\begin{aligned}\partial_i f_\varepsilon(x) &= \frac{\partial}{\partial x_i}\int_0^1 \frac{1}{2t}E\big[h_\varepsilon(\sqrt{t}x + \sqrt{1-t}N)\big]\,dt \\ &= \int_0^1 E\big[h_\varepsilon(\sqrt{t}x + \sqrt{1-t}N)N_i\big]\frac{1}{2\sqrt{t}\sqrt{1-t}}\,dt.\end{aligned} \tag{5.42}$$

It follows directly from the polynomial growth of h that

$$\big|h_\varepsilon(x)\big| \leqslant C_1|x|^m + C_2 \tag{5.43}$$

for all $\varepsilon < 1$, where $C_1, C_2 > 0$ are two constants depending on c and m. Then, from (5.41) we can write

$$\big|\partial_i f_\varepsilon(x)\big| \leqslant C_1|x|^m + C_2,$$

with two possibly different constants C_1, and C_2. Hence,

$$\big\|\partial_i f_\varepsilon(F)\big\|_2 \leqslant C_1\|F\|_{2m}^m + C_2. \tag{5.44}$$

Meanwhile, note that for $1 \leqslant i \leqslant d$,

$$\begin{aligned}E\big[F_i\partial_i f_\varepsilon(F)\big] &= E\big[LL^{-1}F_i\partial_i f_\varepsilon(F)\big] \\ &= E\big[\langle -DL^{-1}F_i, D\partial_i f_\varepsilon(F)\rangle\big] \\ &= \sum_{j=1}^d E\big[\langle -DL^{-1}F_i, \partial_{ij} f_\varepsilon(F)DF_j\rangle\big].\end{aligned}$$

Then, replacing x by F and taking expectation in Eq. (5.41) yields

$$\big|E\big[h_\varepsilon(F)\big]-E\big[h_\varepsilon(N)\big]\big|=\left|\sum_{i,j=1}^{d}E\big[\partial_{ij}^2 f_\varepsilon(F)A_{ij}\big]\right|. \tag{5.45}$$

Notice that

$$\big\langle DF_i, D\partial_i f_\varepsilon(F)\big\rangle_{\mathfrak{H}}=\left\langle DF_i,\sum_{j=1}^{d}\partial_{ij}^2 f_\varepsilon(F)DF_j\right\rangle_{\mathfrak{H}}=\sum_{j=1}^{d}\partial_{ij}^2 f_\varepsilon(F)\langle DF_i, DF_j\rangle_{\mathfrak{H}}$$

for all $1\leqslant i,k\leqslant d$, which implies

$$\partial_{ij} f_\varepsilon(F)=\sum_{k=1}^{d}\big(\gamma_F^{-1}\big)_{jk}\big\langle DF_k, D\partial_i f_\varepsilon(F)\big\rangle_{\mathfrak{H}},$$

and hence

$$\begin{aligned}\sum_{i,j=1}^{d}E\big[\partial_{ij}^2 f_\varepsilon(F)A_{ij}\big]&=\sum_{i,j=1}^{d}E\left[A_{ij}\left\langle\sum_{k=1}^{d}\big(\gamma_F^{-1}\big)_{jk}DF_k, D\partial_i f_\varepsilon(F)\right\rangle_{\mathfrak{H}}\right]\\&=\sum_{i,j=1}^{d}E\left[\partial_i f_\varepsilon(F)\delta\left(A_{ij}\sum_{k=1}^{d}\big(\gamma_F^{-1}\big)_{jk}DF_k\right)\right].\end{aligned}$$

Substituting this expression in (5.45) and using (5.44) we obtain

$$\begin{aligned}\big|E\big[h_\varepsilon(F)\big]-E\big[h_\varepsilon(N)\big]\big|&=\sum_{i,j,k=1}^{d}E\big[\partial_i f_\varepsilon(F)\delta\big(A_{ij}\big(\gamma_F^{-1}\big)_{jk}DF_k\big)\big]\\&\leqslant\sum_{i,j,k=1}^{d}\big\|\partial_i f_\varepsilon(F)\big\|_2\big\|\delta\big(A_{ij}\big(\gamma_F^{-1}\big)_{jk}DF_k\big)\big\|_2\\&\leqslant\big(C_1\|F\|_{2m}^m+C_2\big)\sum_{i,j,k=1}^{d}\big\|\delta\big(A_{ij}\big(\gamma_F^{-1}\big)_{jk}DF_k\big)\big\|_2.\end{aligned}$$

Then, we can conclude the proof by observing that

$$\lim_{\varepsilon\to 0}\big|E\big[h_\varepsilon(F)\big]-E\big[h_\varepsilon(N)\big]\big|=\big|E\big[h(F)\big]-E\big[h(N)\big]\big|,$$

which follows from (5.43) and the fact that $h_\varepsilon\to h$ almost everywhere. □

The next lemma gives an estimate for $\|\delta(A_{ij}(\gamma_F^{-1})_{jk}DF_k)\|_2$ when F is a vector of multiple stochastic integrals.

Lemma 5.9. *Let $F=(F_1,\dots,F_d)=(I_{q_1}(f_1),\dots,I_{q_d}(f_d))$, where $f_i\in\mathfrak{H}^{\odot q_i}$, be non-degenerate and denote $N\sim N(0,I)$. Recall the notation of V and Q in (5.4) and A_{ij} in (5.40). Then, for all $1\leqslant i,j,k\leqslant d$ we have*

$$\begin{aligned}&\big\|\delta\big(A_{ij}\big(\gamma_F^{-1}\big)_{jk}DF_k\big)\big\|_2\\ &\quad\leqslant C\big\|(\det\gamma_F)^{-1}\big\|_{12}^3\Bigg(|V-I|+\sum_{i=1}^d\big\|\|DF_i\|_{\mathfrak{H}}^2-q_iE\big[F_i^2\big]\big\|_2^{\frac12}\Bigg),\end{aligned}\tag{5.46}$$

where the constant C depends on d, V, Q.

Proof. Applying Meyer's inequality (2.9) we have

$$\big\|\delta\big(A_{ij}\big(\gamma_F^{-1}\big)_{jk}DF_k\big)\big\|_2\leqslant\big\|A_{ij}\big(\gamma_F^{-1}\big)_{jk}DF_k\big\|_2+\big\|D\big(A_{ij}\big(\gamma_F^{-1}\big)_{jk}DF_k\big)\big\|_2.$$

Applying Hölder's inequality and (2.12) we have

$$\big\|A_{ij}\big(\gamma_F^{-1}\big)_{jk}DF_k\big\|_2\leqslant\|A_{ij}\|_2\big\|\big(\gamma_F^{-1}\big)_{jk}\big\|_4\|DF_k\|_4\leqslant C_{d,V,Q}\|A_{ij}\|_2\big\|\big(\gamma_F^{-1}\big)_{jk}\big\|_4.$$

Similarly, Hölder's inequality and (2.12) imply

$$\begin{aligned}&\big\|D\big(A_{ij}\big(\gamma_F^{-1}\big)_{jk}DF_k\big)\big\|_2\\ &\quad\leqslant C_{d,V,Q}\big[\|DA_{ij}\|_2\big\|\big(\gamma_F^{-1}\big)_{jk}\big\|_4+\|A_{ij}\|_2\big\|D\big(\gamma_F^{-1}\big)_{jk}\big\|_4+\|A_{ij}\|_2\big\|\big(\gamma_F^{-1}\big)_{jk}\big\|_4\big].\end{aligned}$$

Combining the above inequalities we obtain

$$\big\|\delta\big(A_{ij}\big(\gamma_F^{-1}\big)_{jk}DF_k\big)\big\|_2\leqslant C_{d,V,Q}\|A_{ij}\|_{1,2}\big\|\big(\gamma_F^{-1}\big)_{jk}\big\|_{1,4}.$$

Note that

$$A_{ij}=\delta_{ij}-\big\langle DF_j,-DL^{-1}F_i\big\rangle_{\mathfrak{H}}=\delta_{ij}-V_{ij}+V_{ij}-\frac{1}{q_i}\langle DF_j,-DF_i\rangle_{\mathfrak{H}}.$$

Then, it follows from Lemma 5.5 that

$$\|A_{ij}\|_{1,2}\leqslant C_{d,V,Q}\Bigg(|V-I|+\sum_{i=1}^d\big\|\|DF_i\|_{\mathfrak{H}}^2-q_iE\big[F_i^2\big]\big\|_2^{\frac12}\Bigg).$$

Then, the lemma follows by taking into account of (5.34) with $k=1$. □

As a consequence of the above two lemmas, we have the following result.

Proposition 5.10. *Let $h : \mathbb{R}^d \to \mathbb{R}$ be an almost everywhere continuous function such that $|h(x)| \leqslant c(|x|^m + 1)$ for some $m, c > 0$. Let $F = (F_1, \dots, F_d) = (I_{q_1}(f_1), \dots, I_{q_d}(f_d))$, where $f_i \in \mathfrak{H}^{\odot q_i}$, be non-degenerate and denote $N \sim N(0, I)$. Recall the notation of V and Q in (5.4). Then*

$$\begin{aligned} &\big|E\big[h(F)\big] - E\big[h(N)\big]\big| \\ &\quad \leqslant C\big\|(\det\gamma_F)^{-1}\big\|_{12}^3\left(|V - I| + \sum_{i=1}^d \big\|\|DF_i\|_{\mathfrak{H}}^2 - q_i E\big[F_i^2\big]\big\|_2^{\frac{1}{2}}\right), \end{aligned} \tag{5.47}$$

where the constant C depends on d, V, Q, m, c.

In the following, we estimate the term $A_3 = E[|H_\beta(F) - K_\beta(F)|]$ in (5.8), where $H_\beta(F)$ and $K_\beta(F)$ are defined in (5.3) and (5.7), respectively.

Lemma 5.11. *Let $F = (F_1, \dots, F_d) = (I_{q_1}(f_1), \dots, I_{q_d}(f_d))$ be non-degenerate. Let $\beta = (\beta_1, \dots, \beta_k)$ be a multi-index of length $k \geqslant 1$. Let $H_\beta(F)$ and $K_\beta(F)$ be defined by (5.3) and (5.7), respectively. Then there exists a constant C depending on d, V, Q, k such that*

$$\begin{aligned} &E\big[\big|H_\beta(F) - K_\beta(F)\big|\big] \\ &\quad \leqslant C\big\|(\det\gamma_F)^{-1}\big\|_{(k+4)2^{k+3}}^{k(k+2)} \sum_{i=1}^d \big\|\|DF_i\|_{\mathfrak{H}}^2 - q_i E\big[F_i^2\big]\big\|_2^{\frac{1}{2}}. \end{aligned} \tag{5.48}$$

Proof. To simplify notation, we write H_β and K_β for $H_\beta(F)$ and $K_\beta(F)$, respectively. From (5.3) and (5.7) we see that

$$H_\beta - K_\beta = \delta\big(H_{\widehat{\beta}_k}\big(\gamma_F^{-1}DF\big)_{\beta_k} - K_{\widehat{\beta}_k}\big((VQ)^{-1}DF\big)_{\beta_k}\big),$$

where $\widehat{\beta}_k = (\beta_1, \dots, \beta_{k-1})$. For any $s \geqslant 0$, $p > 1$, using Meyer's inequality (2.9) we obtain

$$\begin{aligned} \|H_\beta - K_\beta\|_{s,p} &\leqslant C_{s,p}\big\|H_{\widehat{\beta}_k}\big(\gamma_F^{-1}DF\big)_{\beta_k} - K_{\widehat{\beta}_k}\big((VQ)^{-1}DF\big)_{\beta_k}\big\|_{s+1,p} \\ &\leqslant C_{s,p}\big\|(H_{\widehat{\beta}_k} - K_{\widehat{\beta}_k})\big((VQ)^{-1}DF\big)_{\beta_k}\big\|_{s+1,p} \\ &\quad + C_{s,p}\big\|H_{\widehat{\beta}_k}\big(\big(\gamma_F^{-1} - (VQ)^{-1}\big)DF\big)_{\beta_k}\big\|_{s+1,p}. \end{aligned}$$

Then, Hölder's inequality yields

$$\begin{aligned} \|H_\beta - K_\beta\|_{s,p} \leqslant{}& \|H_{\widehat{\beta}_k} - K_{\widehat{\beta}_k}\|_{s+1,2p}\big\|\big((VQ)^{-1}DF\big)_{\beta_k}\big\|_{s+1,2p} \\ &+ \|H_{\widehat{\beta}_k}\|_{s+1,2p}\big\|\big(\big(\gamma_F^{-1} - (VQ)^{-1}\big)DF\big)_{\beta_k}\big\|_{s+1,2p}. \end{aligned}$$

Note that (2.12) implies $\|((VQ)^{-1}DF)_{\beta_k}\|_{s+1,2p} \leqslant C_{d,V,Q,s,p}$. Also note that (2.12), Hölder's inequality and (5.37) indicate

$$\big\|\big(\big(\gamma_F^{-1} - (VQ)^{-1}\big)DF\big)_{\beta_k}\big\|_{s+1,2p} \leqslant C_{d,V,Q,s,p}\Delta\big\|(\det\gamma_F)^{-1}\big\|_{(s+3)8p}^{s+2},$$

where we denote

$$\Delta := \sum_{1\leqslant l\leqslant d} \big\| \|DF_l\|_{\mathfrak{H}}^2 - q_l E\big[F_l^2\big]\big\|_2^{\frac{1}{2}}$$

to simplify notation. Thus we obtain

$$\|H_\beta - K_\beta\|_{s,p} \leqslant C_{d,V,Q,s,p}\|H_{\widehat{\beta}_k} - K_{\widehat{\beta}_k}\|_{s+1,2p} + C_{d,V,Q,s,p}\Delta\|H_{\widehat{\beta}_k}\|_{s+1,2p}\big\|(\det\gamma_F)^{-1}\big\|_{(s+3)8p}^{s+2}. \tag{5.49}$$

Similarly, from Meyer's inequality (2.9), Hölder's inequality and (2.12) we obtain by iteration

$$\begin{aligned}\|H_\beta\|_{s,p} &\leqslant C_{s,p}\big\|H_{\widehat{\beta}_k}\big(\gamma_F^{-1}DF\big)_{\beta_k}\big\|_{s+1,p}\\ &\leqslant C_{d,V,Q,s,p}\|H_{\widehat{\beta}_k}\|_{s+1,2p}\big\|(\det\gamma_F)^{-1}\big\|_{(s+3)8p}^{s+2}\\ &\quad\dots\\ &\leqslant C_{d,V,Q,s,p,k}\big\|(\det\gamma_F)^{-1}\big\|_{(s+k+1)2^{k+2}p}^{k(s+k)}.\end{aligned} \tag{5.50}$$

Applying (5.50) into (5.49) and by iteration we can obtain

$$\|H_\beta - K_\beta\|_{s,p} \leqslant C_{d,V,Q,s,p,k}\big\|(\det\gamma_F)^{-1}\big\|_{(2s+k+4)2^{k+2}p}^{k(2s+k+2)}\Delta.$$

Now (5.48) follows by taking $s = 0$, $p = 2$ in the above inequality. □

6. Uniform estimates for densities of general random variables

In this section, we study the uniform convergence of densities for general random variables. We first characterize the convergence of densities with quantitative bounds for a sequence of centered random variables, using the density formula (3.10). In the second part of this section, a short proof of the uniform convergence of densities (without quantitative bounds) is given, using a compactness argument based on the assumption that the sequence converges in law.

6.1. Convergence of densities with quantitative bounds

In this subsection, we estimate the rate of uniform convergence for densities of general random variables. The idea is to use the density formula (3.10).

We use the following notations throughout this section:

$$\bar{w} = \big\langle DF, -DL^{-1}F\big\rangle_{\mathfrak{H}}, \qquad \bar{u} = -\bar{w}^{-1}DL^{-1}F.$$

The following technical lemma is useful.

Lemma 6.1. *Let $F \in \mathbb{D}^{2,s}$ with $s \geqslant 4$ such that $E[F] = 0$ and $E[F^2] = \sigma^2$. Let m be the largest even integer less than or equal to $\frac{s}{2}$. Then there is a positive constant C_m such that for any $t \leqslant m$,*

$$\left\|\bar{w} - \sigma^2\right\|_t \leqslant \left\|\bar{w} - \sigma^2\right\|_m \leqslant C_m \|D\bar{w}\|_m \leqslant C_m \|D\bar{w}\|_{s/2}. \tag{6.1}$$

Proof. It suffices to show the above second inequality. From the integration by parts formula in Malliavin calculus it follows

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

Note that from (3.9) and (3.13) we have $\bar{w} \in \mathbb{D}^{1,\frac{s}{2}}$. Then the lemma follows from the following infinite-dimensional Poincaré inequality [16, Lemma 5.3.8]:

$$E\left[\left(G - E[G]\right)^m\right] \leqslant (m-1)^{m/2} E\left[\|DG\|_{\mathfrak{H}}^m\right],$$

for any even integer m and $G \in \mathbb{D}^{1,m}$. □

The next theorem gives a bound for the uniform distance between the density of a random variable F and the normal density.

Theorem 6.2. *Let $F \in \mathbb{D}^{2,s}$ with $s \geqslant 8$ such that $E[F] = 0$, $E[F^2] = \sigma^2$. Suppose $M^r := E[|\bar{w}|^{-r}] < \infty$, where $\bar{w} = \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}$ and $r > 2$. Assume $\frac{2}{r} + \frac{4}{s} = 1$. Then F admits a density $f_F(x)$ and there is a constant $C_{r,s,\sigma,M}$ depending on r, s, σ and M such that*

$$\sup_{x \in \mathbb{R}} \left|f_F(x) - \phi(x)\right| \leqslant C_{r,s,\sigma,M} \|F\|_{1,s}^2 \left\| \left\|D^2 F\right\|_{op}\right\|_{0,s}, \tag{6.2}$$

where $\phi(x)$ is the density of $N \sim N(0, \sigma^2)$ and $\|D^2F\|_{op}$ indicates the operator norm of D^2F introduced in (3.9).

Proof. It follows from Proposition 3.3 that F admits a density given by $f_F(x) = E[\mathbf{1}_{\{F>x\}}\delta(\bar{u})]$. Then

$$\sup_{x \in \mathbb{R}} \left|f_F(x) - \phi(x)\right| = \sup_{x \in \mathbb{R}} \left|E\left[\mathbf{1}_{\{F>x\}}\delta(\bar{u})\right] - \sigma^{-2} E[\mathbf{1}_{\{N>x\}}N]\right|. \tag{6.3}$$

Note that, from (2.7)

$$\delta(\bar{u}) = \delta\left(-DL^{-1}F\bar{w}^{-1}\right) = F\bar{w}^{-1} + \left\langle D\bar{w}^{-1}, DL^{-1}F\right\rangle_{\mathfrak{H}}.$$

Then

$$\begin{aligned} &\left|E\left[\sigma^2 \mathbf{1}_{\{F>x\}}\delta(\bar{u})\right] - E[\mathbf{1}_{\{N>x\}}N]\right| \\ &\quad \leqslant E\left[\left|F\bar{w}^{-1}\left(\sigma^2 - \bar{w}\right)\right|\right] + \sigma^2 E\left[\left|\left\langle D\bar{w}^{-1}, DL^{-1}F\right\rangle_{\mathfrak{H}}\right|\right] \\ &\qquad + \left|E[F\mathbf{1}_{\{F>x\}} - N\mathbf{1}_{\{N>x\}}]\right|. \end{aligned} \tag{6.4}$$

Note that for $t = (\frac{1}{r} + \frac{3}{s})^{-1}$, we have $\frac{s}{2} - t \geqslant 2$, so there exists an even integer $m \in [t, \frac{s}{2}]$. Also, we have $\frac{1}{r} + \frac{1}{s} + \frac{1}{t} = 1$. Then, we can apply Hölder's inequality and (6.1) to obtain

$$E\big[\big|F\bar{w}^{-1}\big(\bar{w} - \sigma^2\big)\big|\big] \leqslant \|F\|_s \big\|\bar{w}^{-1}\big\|_r \big\|\bar{w} - \sigma^2\big\|_t \\ \leqslant C_{r,s}\|F\|_s \big\|\bar{w}^{-1}\big\|_r \|D\bar{w}\|_{s/2}. \tag{6.5}$$

Meanwhile, applying Hölder's inequality and (3.7) we have

$$E\big[\big|\bar{w}^{-2}\langle D\bar{w}, -DL^{-1}F\rangle_{\mathfrak{H}}\big|\big] \leqslant \big\|\bar{w}^{-1}\big\|_r^2 \|D\bar{w}\|_{\frac{s}{2}} \big\|DL^{-1}F\big\|_{\frac{s}{2}} \\ \leqslant \big\|\bar{w}^{-1}\big\|_r^2 \|D\bar{w}\|_{\frac{s}{2}} \|DF\|_s. \tag{6.6}$$

Also, applying Lemma 2.2 for $h(y) = y\mathbf{1}_{\{y>x\}}$ and (6.1) we have

$$\big|E[F\mathbf{1}_{F>x} - N\mathbf{1}_{N>x}]\big| \leqslant C_\sigma \big\|\sigma^2 - \bar{w}\big\|_2 \leqslant C_\sigma \|D\bar{w}\|_{s/2}. \tag{6.7}$$

Applying the estimates (6.5)–(6.7) to (6.4) we have

$$\big|E\big[\sigma^2\mathbf{1}_{F>x}\delta(\bar{u})\big] - E[\mathbf{1}_{N>x}N]\big| \leqslant C_{r,s,\sigma,M}\|F\|_{1,s}\|D\bar{w}\|_{s/2}. \tag{6.8}$$

Combining (6.3), (6.8) and (3.13) one gets

$$\sup_{x\in\mathbb{R}}\big|f_F(x) - \phi(x)\big| \leqslant C_{r,s,\sigma,M}\|F\|_{1,s}^2 \big\|\big\|D^2F\big\|_{op}\big\|_s.$$

This completes the proof. □

Corollary 6.3. *Let $\{F_n\}_{n\in\mathbb{N}} \subset \mathbb{D}^{2,s}$ with $s \geqslant 8$ such that $E[F_n] = 0$ and $\lim_{n\to\infty} E[F_n^2] = \sigma^2$. Assume $E[F_n^2] \geqslant \delta > 0$ for all n. For $r > 2$ such that $\frac{2}{r} + \frac{4}{s} = 1$, assume*

(i) $M_1 = \sup_n \|F_n\|_{1,s} < \infty$.
(ii) $M_2 = \sup_n E|\langle DF_n, -DL^{-1}F_n\rangle_{\mathfrak{H}}|^{-r} < \infty$.
(iii) $E\|D^2F_n\|_{op}^s \to 0$ *as* $n \to \infty$.

Then each F_n admits a density $f_{F_n}(x)$ and

$$\sup_{x\in\mathbb{R}}\big|f_{F_n}(x) - \phi(x)\big| \leqslant C\big(\big\|\big\|D^2F_n\big\|_{op}\big\|_s + \big|E\big[F_n^2\big] - \sigma^2\big|\big), \tag{6.9}$$

where the constant C depends on σ, M_1, M_2 and δ. Moreover, if $M_3 = \sup_n \|F_n\|_{2s} < \infty$, then for any $k \geqslant 1$ and $\alpha \in (\frac{1}{2}, k)$,

$$\|f_{F_n} - \phi\|_{L^k(\mathbb{R})} \leqslant C\big(\big\|\big\|D^2F_n\big\|_{op}\big\|_s + \big|E\big[F_n^2\big] - \sigma^2\big|\big)^{\frac{k-\alpha}{k}},$$

where the constant C depends on $\sigma, M_1, M_2, M_3, \alpha$ and δ.

Remark 6.4. By the "random contraction inequality" (3.9), a sufficient condition for (iii) is $E\|D^2F_n \otimes_1 D^2F_n\|_{\mathfrak{H}^{\otimes 2}}^{s/2} \to 0$ or $E\|D^2F_n\|_{\mathfrak{H}^{\otimes 2}}^{s} \to 0$.

Proof of Corollary 6.3. It follows from Theorem 6.2 and Proposition 3.3 with an argument similar to Corollary 4.3. □

6.2. Compactness argument

In general, convergence in law does not imply convergence of the corresponding densities even if they exist. The following theorem specifies some additional conditions which ensure that convergence in law will imply convergence of densities.

Theorem 6.5. *Let $\{F_n\}_{n\in\mathbb{N}}$ be a sequence of random variables in $\mathbb{D}^{2,s}$ satisfying any one of the following two conditions*:

$$\sup_n \|F_n\|_{2,s} + \sup_n \|F_n\|_{2p} + \sup_n \big\| \|DF_n\|_{\mathfrak{H}}^{-2} \big\|_r < \infty \tag{6.10}$$

for some $p, r, s > 1$ satisfying $\frac{1}{p} + \frac{1}{r} + \frac{1}{s} = 1$, or

$$\sup_n \|F_n\|_{2,s} + \sup_n \big\| \big| \langle DF_n, -DL^{-1}F_n \rangle_{\mathfrak{H}} \big|^{-1} \big\|_r < \infty \tag{6.11}$$

for some $r, s > 1$ satisfying $\frac{2}{r} + \frac{4}{s} = 1$.

Suppose in addition that $F_n \to N \sim N(0, \sigma^2)$ in law. Then each F_n admits a density $f_{F_n} \in C(\mathbb{R})$ given by either (3.1) or (3.10), and

$$\sup_{x\in\mathbb{R}} \big| f_{F_n}(x) - \phi(x) \big| \to 0$$

as $n \to \infty$, where ϕ is the density of N.

Proof. We assume (6.10). The other condition can be treated identically. From Theorem 3.1 it follows that the density formula (3.1) holds for each n and for all $x, y \in \mathbb{R}$,

$$\big| f_{F_n}(x) \big| \leqslant C\big(1 \wedge x^{-2}\big),$$
$$\big| f_{F_n}(x) - f_{F_n}(y) \big| \leqslant C|x-y|^{\frac{1}{p}}.$$

Hence the sequence $\{f_{F_n}\} \subset C(\mathbb{R})$ is uniformly bounded and equi-continuous. Then applying Azelà–Ascoli theorem, we obtain a subsequence $\{f_{F_{n_k}}\}$ which converges uniformly to a continuous function f on $\mathbb{R}$ such that $0 \leqslant f(x) \leqslant C(1 \wedge x^{-2})$. Then $f_{F_{n_k}} \to f$ in $L^1(\mathbb{R})$ as $k \to \infty$ with $\|f\|_{L^1(\mathbb{R})} = \lim_k \|f_{F_{n_k}}\|_{L^1(\mathbb{R})} = 1$. This implies that f is a density function. Then f must be ϕ because F_n converges to N in law. Since the limit is unique for any subsequence, we get the uniform convergence of f_{F_n} to ϕ. □

Corollary 6.6. *Let $\{F_n\}_{n\in\mathbb{N}}$ be a sequence of centered random variables in $\mathbb{D}^{2,4}$ with the following Wiener chaos expansions*: $F_n = \sum_{q=1}^{\infty} J_q F_n$. *Suppose that*

(i) $\lim_{Q\to\infty} \limsup_{n\to\infty} \sum_{q=Q+1}^{\infty} E[|J_q F_n|^2] = 0$.
(ii) *for every* $q \geqslant 1$, $\lim_{n\to\infty} E[(J_q F_n)^2] = \sigma_q^2$.
(iii) $\sum_{q=1}^{\infty} \sigma_q^2 = \sigma^2$.
(iv) *for all* $q \geqslant 1$, $\langle D(J_q F_n), D(J_q F_n)\rangle_{\mathfrak{H}} \longrightarrow q\sigma_q^2$, *in* $L^2(\Omega)$ *as* $n \to \infty$.
(v) $\sup_n \|F_n\|_{2,4} + \sup_n E[\|DF_n\|_{\mathfrak{H}}^{-8}] < \infty$.

Then each F_n *admits a density* $f_{F_n}(x)$ *and*

$$\sup_{x\in\mathbb{R}} \left| f_{F_n}(x) - \phi(x) \right| \to 0$$

as $n \to \infty$, *where* ϕ *is the density of* $N(0, \sigma^2)$.

Proof. It has been proved by Nualart and Ortiz-Latorre in [23, Theorem 8] that under conditions (i)–(iv), F_n converges to $N \sim N(0, \sigma^2)$ in law. The condition (v) implies (6.10) with $s = 4$, $p = 2$, $r = 4$. Then we can conclude from Theorem 6.5. □

7. Applications

The main difficulty in applying Theorem 4.1 or Theorem 5.2 is the verification of the non-degeneracy condition of the Malliavin matrix: $\sup_n E[\|DF_n\|_{\mathfrak{H}}^{-p}] < \infty$ or $\sup_n E[|\det \gamma_{F_n}|^{-p}] < \infty$, respectively. In this section we consider the particular case of random variables in the second Wiener chaos and we find sufficient conditions for $\sup_n E[\|DF_n\|_{\mathfrak{H}}^{-p}] < \infty$. As an application we consider the problem of estimating the drift parameter in an Ornstein–Uhlenbeck process.

A general approach to verify $E[G^{-p}] < \infty$ for some positive random variable and for some $p \geqslant 1$ is to obtain a small ball probability estimate of the form

$$P(G \leqslant \varepsilon) \leqslant C\varepsilon^{\alpha} \quad \text{for some } \alpha > p \text{ and for all } \varepsilon \in (0, \varepsilon_0), \tag{7.1}$$

where $\varepsilon_0 > 0$ and $C > 0$ is a constant that may depend on ε_0 and α. We refer to the paper by Li and Shao [10] for a survey on this topic. However, finding upper bounds of this type is a challenging topic, and the application of small ball probabilities to Malliavin calculus is still an under-explored domain.

7.1. Random variables in the second Wiener chaos

A random variable F in the second Wiener chaos can always be written as $F = I_2(f)$ where $f \in \mathfrak{H}^{\odot 2}$. Without loss of generality we can assume that

$$f = \sum_{i=1}^{\infty} \lambda_i e_i \otimes e_i, \tag{7.2}$$

where $\{\lambda_i, i \geqslant 1\}$ verifying $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n| \geqslant \cdots$ are the eigenvalues of the Hilbert–Schmidt operator corresponding to f and $\{e_i, i \geqslant 1\}$ are the corresponding eigenvectors forming an orthonormal basis of $\mathfrak{H}$. Then, we have $F = I_2(f) = \sum_{i=1}^{\infty} \lambda_i (I_1(e_i)^2 - 1)$,

$$DF = 2\sum_{i=1}^{\infty} \lambda_i I_1(e_i)e_i \tag{7.3}$$

and

$$\|DF\|_{\mathfrak{H}}^2 = 4\sum_{i=1}^{\infty} \lambda_i^2 I_1(e_i)^2. \tag{7.4}$$

For random variables of the form in (7.4), i.e., $G = (\sum_{i=1}^{\infty} \lambda_i^2 X_i^2)^{\frac{1}{2}}$, Hoffmann-Jørgensen, Shepp and Dudley [6] used the volume of the small ball $B_n(0, \varepsilon)$ (the $\mathbb{R}^n$ ball centered at 0 with radius ε) to control $P(G \leqslant \varepsilon)$ as

$$P(G \leqslant \varepsilon) \leqslant P\left(\sum_{i=1}^{n} \lambda_i^2 X_i^2 \leqslant \varepsilon^2\right) \leqslant (2\pi)^{-\frac{n}{2}} \varepsilon^n \left|B_n(0, 1)\right| \prod_{i=1}^{n} \lambda_i^{-1}. \tag{7.5}$$

They proved that $P(G \leqslant \varepsilon)$ converges to zero at the rate $O(\varepsilon^n)$ for all n as $\varepsilon \to 0$, under some implicit conditions on $\{\lambda_i, i \geqslant 1\}$. This idea can be used here to prove inequality (7.6) in the following lemma. However, our case is much simpler, and we shall use the Gamma function to give an alternative proof which leads to a necessary and sufficient condition for $E[G^{-p}] < \infty$.

Lemma 7.1. *Let $G = (\sum_{i=1}^{\infty} \lambda_i^2 X_i^2)^{\frac{1}{2}}$, where $\{\lambda_i\}_{i \geqslant 1}$ satisfies $|\lambda_i| \geqslant |\lambda_{i+1}|$ for all $i \geqslant 1$ and $\{X_i\}_{i \geqslant 1}$ are i.i.d standard normal. Fix an $\alpha > 1$. Then, $E[G^{-2\alpha}] < \infty$ if and only if there exists an integer $N > 2\alpha$ such that $|\lambda_N| > 0$ and in this case there exists a constant C_α depending only on α such that*

$$E\left[G^{-2\alpha}\right] \leqslant C_\alpha N^{-\alpha} |\lambda_N|^{-2\alpha}. \tag{7.6}$$

Proof. Notice $\lambda^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} e^{-\lambda y} y^{\alpha-1}\, dy$ and $E[e^{-tX_i^2}] = \frac{1}{\sqrt{1+2t}}$ for all $t > 0$. If there exists $N > 2\alpha$ such that $|\lambda_N| > 0$, then

$$\begin{aligned} E\left[G^{-2\alpha}\right] &\leqslant E\left[\left(\sum_{i=1}^{N} \lambda_i^2 X_i^2\right)^{-\alpha}\right] = \frac{1}{\Gamma(\alpha)} E\left[\int_0^{\infty} e^{-y \sum_{i=1}^{N} \lambda_i^2 X_i^2} y^{\alpha-1}\, dy\right] \\ &= \frac{1}{\Gamma(\alpha)} \int_0^{\infty} y^{\alpha-1} \prod_{i=1}^{N} \left(1 + 2\lambda_i^2 y\right)^{-\frac{1}{2}} dy. \end{aligned} \tag{7.7}$$

Since λ_i^2 is non-increasing in i and $N > 2\alpha$, using the change of variables $1 + 2\lambda_N^2 y = z$ we have

$$\begin{aligned} \int_0^{\infty} y^{\alpha-1} \prod_{i=1}^{N} \left(1 + 2\lambda_i^2 y\right)^{-\frac{1}{2}} dy &\leqslant \int_0^{\infty} y^{\alpha-1} \left(1 + 2\lambda_N^2 y\right)^{-\frac{N}{2}} dy \\ &= \left(2\lambda_N^2\right)^{-\alpha} \int_1^{\infty} (z-1)^{\alpha-1} z^{-\frac{N}{2}}\, dz \end{aligned}$$

$$= (2\lambda_N^2)^{-\alpha} \int_1^\infty \left(\frac{z-1}{z}\right)^{\alpha-1} z^{\alpha-1-\frac{N}{2}}\, dz$$

$$= (2\lambda_N^2)^{-\alpha} \int_0^1 (1-x)^{\alpha-1} x^{\frac{N}{2}-\alpha-1}\, dx$$

$$= (2\lambda_N^2)^{-\alpha} \frac{\Gamma(\alpha)\Gamma(\frac{N}{2}-\alpha)}{\Gamma(N/2)},$$

which implies (7.6).

On the other hand, if $|\lambda_i| = 0$ for all $i > 2\alpha$, let $N \leqslant 2\alpha$ be the largest nonnegative integer such that $|\lambda_N| > 0$. Then, the inequality in (7.7) becomes an equality. Using again that $\{\lambda_i^2\}_{i\geqslant 1}$ is a decreasing sequence we have

$$\int_0^\infty y^{\alpha-1} \prod_{i=1}^N (1+2\lambda_i^2 y)^{-\frac{1}{2}}\, dy \geqslant (1+2\lambda_1^2)^{-\frac{N}{2}} \left(\int_0^1 y^{\alpha-1}\, dy + \int_1^\infty y^{\alpha-1-\frac{N}{2}}\, dy\right) = \infty,$$

and we conclude that $E[G^{-2\alpha}] = \infty$. This completes the proof. □

The following theorem describes the distance between the densities of $F = I_2(f)$ and $N(0, E[F^2])$.

Theorem 7.2. *Let $F = I_2(f)$ with $f \in \mathfrak{H}^{\odot 2}$ given in (7.2). Assume that there exists $N > 6m + 6(\lfloor \frac{m}{2} \rfloor \vee 1)$, for some integer $m \geqslant 0$, such that $\lambda_N \neq 0$. Then F admits an m times continuously differentiable density f_F. Furthermore, if $\phi(x)$ denotes the density of $N(0, E[F^2])$, then for $k = 0, 1, \ldots, m$,*

$$\sup_{x\in\mathbb{R}} \left| f_F^{(k)}(x) - \phi^{(k)}(x) \right| \leqslant C \left(\sum_{i=1}^\infty \lambda_i^4\right)^{\frac{1}{2}} \leqslant C\left(E[F^4] - 3(E[F^2])^2\right)^{\frac{1}{2}},$$

where the constant C depends on N and λ_N.

Proof. Taking into account of (7.4), we have

$$\operatorname{Var}(\|DF\|_{\mathfrak{H}}^2) = E\left|4 \sum_{i=1}^\infty \lambda_i^2 (I_1(e_i)^2 - 1)\right|^2 = 32 \sum_{i=1}^\infty \lambda_i^4. \tag{7.8}$$

From (7.4) and Lemma 7.1 it follows that

$$E[\|DF\|_{\mathfrak{H}}^{-\beta}] \leqslant C_{\beta/2} N^{-\beta/2} |\lambda_N|^{-\beta}, \tag{7.9}$$

for all $\beta < N$. Then, the theorem follows from Theorem 4.4, taking into account (7.8). □

Now we are ready to prove convergence of densities of random variables in the second Wiener chaos. Consider a sequence $F_n = I_2(f_n)$ with $f_n \in \mathfrak{H}^{\odot 2}$, which can be written as

$$f_n = \sum_{i=1}^{\infty} \lambda_{n,i} e_{n,i} \otimes e_{n,i}, \tag{7.10}$$

where $\{\lambda_{n,i}, i \geqslant 1\}$ verifies $|\lambda_{n,i}| \geqslant |\lambda_{n,i+1}|$ for all $i \geqslant 1$ and $\{e_{n,i}, i \geqslant 1\}$ are the corresponding eigenvectors.

Theorem 7.3. *Let $F_n = I_2(f_n)$ with $f_n \in \mathfrak{H}^{\odot 2}$ given by (7.10). Assume that $\{\lambda_{n,i}\}_{n,i\in\mathbb{N}}$ satisfies*

(i) $\sigma^2 := 2\lim_{n\to\infty} \sum_{i=1}^{\infty} \lambda_{n,i}^2 > 0$;
(ii) $\lim_{n\to\infty} \sum_{i=1}^{\infty} \lambda_{n,i}^4 = 0$;
(iii) $\inf_n (\sup_{i>6m+6(\lfloor \frac{m}{2} \rfloor \vee 1)} |\lambda_{n,i}|\sqrt{i}\,) > 0$ *for some integer* $m \geqslant 0$.

Then, each F_n admits a density function $f_{F_n} \in C^m(\mathbb{R})$. Furthermore, for $k = 0, 1, \ldots, m$ and if ϕ denotes the density of the law $N(0, \sigma^2)$, the derivatives of $f_{F_n}^{(k)}$ converge uniformly to the derivatives of ϕ with a rate given by

$$\sup_{x\in\mathbb{R}} \left| f_{F_n}^{(k)}(x) - \phi^{(k)}(x) \right| \leqslant C \left[\left(\sum_{i=1}^{\infty} \lambda_{n,i}^4 \right)^{\frac{1}{2}} + \left| 2 \sum_{i=1}^{\infty} \lambda_{n,i}^2 - \sigma^2 \right|^{\frac{1}{2}} \right],$$

where C is a constant depending only on m and the infimum appearing in condition (iii).

Proof of Theorem 7.3. Note that $E[(I_1(e_{n,i})^2 - 1)(I_1(e_{n,j})^2 - 1)] = 2\delta_{ij}$. Thus,

$$\sum_{i=1}^{\infty} \lambda_{n,i}^2 = \| f_n \|_{\mathfrak{H}^{\odot 2}}^2 = \frac{1}{2} E\left[F_n^2\right].$$

Then, the result follows from (7.8), (7.9) and Corollary 4.6. □

Condition (iii) in Theorem 7.3 means that there exist a positive constant $\delta > 0$ such that for each n we can find an index $i(n) > 6m + 6(\lfloor \frac{m}{2} \rfloor \vee 1)$ with $|\lambda_{n,i(n)}|\sqrt{i(n)} \geqslant \delta$.

Remark 7.4. It is interesting to compare Theorem 7.3 with the case when

$$\lambda_{n,i} = \begin{cases} \frac{1}{\sqrt{n}} & \text{if } 1 \leqslant i \leqslant n; \\ 0 & \text{if } i \geqslant n+1, \end{cases}$$

which corresponds to classical case of sum of independent and identically distributed random variables. In this case all the conditions of Theorem 7.3 are satisfied with $\sigma^2 = 2$. Moreover we have $\sum_{i=1}^{\infty} \lambda_{n,i}^4 = \frac{1}{n}$ and $\sum_{i=1}^{\infty} \lambda_{n,i}^2 = 1$. Then we obtain a Berry–Essen type bound for the derivatives of the density. Namely, we have $\sup_{x\in\mathbb{R}} |f_{F_n}^{(k)}(x) - \phi^{(k)}(x)| \leqslant \frac{C}{\sqrt{n}}$ for sufficiently large n, which provides the right rate of convergence.

7.2. Parameter estimation in Ornstein–Uhlenbeck processes

Consider the following Ornstein–Uhlenbeck process

$$X_t = -\theta \int_0^t X_s\, ds + \gamma B_t,$$

where $\theta > 0$ is an unknown parameter, $\gamma > 0$ is known and $B = \{B_t,\ 0 \leqslant t < \infty\}$ is a standard Brownian motion. Assume that the process $X = \{X_t,\ 0 \leqslant t \leqslant T\}$ can be observed continuously in the time interval $[0, T]$. Then the least squares estimator (or the maximum likelihood estimator) of θ is given by $\widehat{\theta}_T = \frac{\int_0^T X_t\, dX_t}{\int_0^T X_t^2\, dt}$. It is known (see for example, [11,9]) that, as T tends to infinity, $\widehat{\theta}_T$ converges to θ almost surely and

$$\sqrt{T}(\widehat{\theta}_T - \theta) = -\frac{T F_T}{\int_0^T X_t^2\, dt} \xrightarrow{\mathcal{L}} N(0, 2\theta), \tag{7.11}$$

where

$$F_T = I_2(f_T) = \int_0^T \int_0^T f_T(t, s)\, dB_t\, dB_s, \tag{7.12}$$

with

$$f_T(t, s) = \frac{\gamma^2}{2\sqrt{T}} e^{-\theta|t-s|}. \tag{7.13}$$

Recently, Hu and Nualart [7] extended this result to the case where B is a fractional Brownian motion with Hurst parameter $H \in [\frac{1}{2}, \frac{3}{4})$, which includes the standard Brownian motion case. Since $\frac{1}{T}\int_0^T X_t^2\, dt \to \frac{1}{2}\gamma^2\theta^{-1}$ almost surely as T tends to infinity, the main effort in proving (7.11) is to show the convergence in law of F_T to the normal law $N(0, \frac{\gamma^4}{2\theta})$. We shall prove that the density of F_T converges as T tends to infinity to the density of the normal distribution $N(0, \frac{\gamma^4}{2\theta})$.

Theorem 7.5. *Let F_T be given by (7.13) and let ϕ be the density of the law $N(0, \sigma^2)$, where $\sigma^2 = \frac{\gamma^4}{2\theta}$. Then for each $T > 0$, F_T has a smooth probability density f_{F_T} and for any $k \geqslant 0$,*

$$\sup_{x \in \mathbb{R}} \left| f_{F_T}^{(k)}(x) - \phi^{(k)}(x) \right| \leqslant C T^{-\frac{1}{2}},$$

where the constant C depends on k, γ and θ.

Before proving the theorem, let us first analyze the asymptotic behavior of the eigenvalues of f_T. The Hilbert space corresponding to Brownian motion B is $\mathfrak{H} = L^2([0,T])$. Let $Q_T : L^2([0,T]) \to L^2([0,T])$ be the Hilbert–Schmidt operator associated to f_T, that is,

$$(Q_T\varphi)(t) = \int_0^T f_T(t,s)\varphi(s)\,ds \tag{7.14}$$

for $\varphi \in L^2[0,T]$. The operator Q_T has eigenvalues $\lambda_{T,1} > \lambda_{T,2} > \cdots \geqslant 0$ and $\sum_{i=1}^\infty \lambda_{T,i}^2 < \infty$. The following lemma provides upper and lower bounds for these eigenvalues.

Lemma 7.6. *Fix $T > 0$. Let f_T be given by (7.13) and Q_T be given by (7.14). The eigenvalues $\lambda_{T,i}$ of Q_T (except maybe one) satisfy the following estimates*

$$\frac{\gamma^2\theta}{\sqrt{T}(\theta^2 + (\frac{i\pi+\frac{\pi}{2}}{T})^2)} < \lambda_{T,i} < \frac{\gamma^2\theta}{\sqrt{T}(\theta^2 + (\frac{i\pi-\frac{\pi}{2}}{T})^2)}. \tag{7.15}$$

Proof. Consider the eigenvalue problem $Q_T\varphi = \lambda\varphi$, that is,

$$\int_0^T f_T(t,s)\varphi(s)\,ds = \frac{\gamma^2}{2\sqrt{T}}\left(\int_0^t e^{-\theta(t-s)}\varphi(s)\,ds + \int_t^T e^{-\theta(s-t)}\varphi(s)\,ds\right) = \lambda\varphi(t). \tag{7.16}$$

Then, ϕ is differentiable and

$$\frac{\gamma^2\theta}{2\sqrt{T}}\left(-\int_0^t e^{-\theta(t-s)}\varphi(s)\,ds + \int_t^T e^{-\theta(s-t)}\varphi(s)\,ds\right) = \lambda\varphi'(t). \tag{7.17}$$

Differentiating again we have

$$\frac{\gamma^2\theta}{2\sqrt{T}}\left(-2\varphi(t) + \theta\int_0^t e^{-\theta(t-s)}\varphi(s)\,ds + \theta\int_t^T e^{-\theta(s-t)}\varphi(s)\,ds\right) = \lambda\varphi''(t).$$

Comparing this expression with (7.16), we obtain

$$\left(\theta^2 - \frac{\gamma^2\theta}{\sqrt{T}\lambda}\right)\varphi(t) = \varphi''(t). \tag{7.18}$$

Also, from (7.16) and (7.17) it follows that

$$\varphi(0) = \theta\varphi'(0), \qquad \varphi(T) = -\theta\varphi'(T). \tag{7.19}$$

Eqs. (7.18) and (7.19) form a Sturm–Liouville system. Its general solution is of the form

$$\varphi(t) = C_1 \sin \mu t + C_2 \cos \mu t,$$

where C_1 and C_2 are constants, and $\mu > 0$ is an eigenvalue of the Sturm–Liouville system. By eliminating the constants C_1 and C_2 from (7.18) and (7.19) we obtain

$$-\mu^2 = \theta^2 - \frac{\gamma^2\theta}{\sqrt{T}\lambda}. \tag{7.20}$$

Then, the desired estimates on the eigenvalues of $Q_T\varphi = \lambda\varphi$ will follow form estimates on μ. Note that the Neumann condition (7.19) yields

$$\left(\mu^2\theta^2 - 1\right) \sin \mu T = 2\mu\theta \cos \mu T.$$

If we write $x = \mu\theta > 0$ (since $\mu, \theta > 0$), the above equation becomes

$$\left(x^2 - 1\right) \sin \frac{x}{\theta} T = 2x \cos \frac{x}{\theta} T.$$

The solution $x = 1$ corresponds to the eigenvalue $\mu = \frac{1}{\theta}$. If $x \neq 1$, then $\cos \frac{x}{\theta} T \neq 0$ and

$$\tan \frac{x}{\theta} T = \frac{2x}{x^2 - 1}. \tag{7.21}$$

For any $i \in \mathbb{Z}_+$, there is exactly one solution x_i to (7.21) such that $\frac{x_i}{\theta} T \in (i\pi - \frac{\pi}{2}, i\pi + \frac{\pi}{2})$. Corresponding to each x_i is an eigenvalue $\mu_i = \frac{x_i}{\theta}$ of the Sturm–Liouville system, satisfying $\frac{i\pi - \frac{\pi}{2}}{T} < \mu_i < \frac{i\pi + \frac{\pi}{2}}{T}$. The corresponding eigenvalue λ_i of Q_T obtained from Eq. (7.20) satisfies the estimate (7.15). □

Proof of Theorem 7.5. For each T, let us compute the second moment of F_T,

$$\begin{aligned} E\left[F_T^2\right] = \|f_T\|^2_{\mathfrak{H}^{\otimes 2}} &= \int_0^T \int_0^T f_T(t,s)^2 \, ds \, dt \\ &= \frac{\gamma^4}{4T} \int_0^T \int_0^t e^{-2\theta(t-s)} \, ds \, dt \\ &= \frac{\gamma^4}{2\theta} - \frac{\gamma^4}{8\theta T}\left(1 - e^{-2\theta T}\right). \end{aligned}$$

Also, noticing that $F_T = I_2(f_T) = \delta^2(f_T)$ and

$$D_s D_t F_T^3 = 3F_T^2 f_T(t,s) + 6F_T I_1\left(f(\cdot,t)\right) \otimes I_1\left(f(\cdot,s)\right),$$

and using the duality between δ and D, we can write

$$\begin{aligned} E\big[F_T^4\big] &= E\big[\langle f_T, D^2 F_T^3\rangle_{\mathfrak{H}^{\otimes 2}}\big] \\ &= 3E\big[F_T^2\langle f_T, f_T\rangle_{\mathfrak{H}^{\otimes 2}}\big] + 6E\big[F_T\big\langle f_T(t,s), I_1\big(f_T(\cdot,t)\big)\otimes I_1\big(f_T(\cdot,s)\big)\big\rangle_{\mathfrak{H}^{\otimes 2}}\big] \\ &= 3\big(E\big[F_T^2\big]\big)^2 + 6A, \end{aligned}$$

where

$$\begin{aligned} A &= E\big[F_T\big\langle f_T(t,s), I_1\big(f_T(\cdot,t)\big)\otimes I_1\big(f_T(\cdot,s)\big)\big\rangle_{\mathfrak{H}^{\otimes 2}}\big] \\ &= \big\langle f_T(u,v), \big\langle f_T(t,s), f_T(u,t)\otimes f_T(v,s)\big\rangle_{\mathfrak{H}^{\otimes 2}}\big\rangle_{\mathfrak{H}^{\otimes 2}} \\ &= \frac{\gamma^8}{16T^2}\int_0^T\int_0^T\int_0^T\int_0^T e^{-\theta(|u-v|+|t-s|+|u-t|+|v-s|)}\,du\,dv\,dt\,ds. \end{aligned}$$

Because the integrand is symmetric, we have

$$A = \frac{\gamma^8}{16T^2}4!\int_0^T du\int_0^u dv\int_0^v ds\int_0^s dt\; e^{-2\theta(u-t)} \leqslant CT^{-1}.$$

Then, in order to complete the proof by applying Corollary 4.6, we only need to verify that condition (iii) of Theorem 7.3 holds for any integer $m \geqslant 1$, which implies the uniform boundedness of the negative moments

$$\sup_{T>0} E\big[\|DF_T\|_{\mathfrak{H}}^{-\beta}\big] < \infty$$

for any $\beta > 0$. Fix $\beta > 0$, and for each T, let $i(T) = \lfloor\beta+1\rfloor + \lfloor T\rfloor$. Then, the lower bound in (7.15) yields

$$\sqrt{i(T)}\lambda_{T,i(T)} \geqslant \frac{\sqrt{i(T)}\gamma^2/\theta}{\sqrt{T}(1+(\frac{(i+1/2)\pi}{T\theta})^2)} \geqslant \frac{\sqrt{i(T)}\gamma^2/\theta}{\sqrt{T}(1+(\frac{i(T)}{T})^2 4\frac{\pi^2}{\theta^2})} \geqslant \frac{\gamma^2/\theta}{\max_{(\beta+2)^{-1}\leqslant r\leqslant 1} g(r)} > 0,$$

where in the last inequality we made the substitution $r^{-1} = \frac{i(T)}{T}$ and set

$$g(r) := \sqrt{r}\left(1 + r^{-2}4\frac{\pi^2}{\theta^2}\right).$$

This implies condition (iii) and the proof of the theorem is complete. □

7.3. Multidimensional case

Now we give an example for random vectors. Let $X = \{X(h), h \in \mathfrak{H}\}$ be an isonormal Gaussian process associated with the Hilbert space $\mathfrak{H}$. Suppose that $\{e_{ij}, 1 \leqslant i \leqslant d, j \geqslant 1\}$ is a sequence of orthonormal elements in $\mathfrak{H}$. Set $e_k = (e_{1k}, \dots, e_{dk})$ for any $k \geqslant 1$. Let A_n be a sequence of $d \times d$ invertible matrices such that $A_n \to I$ as $n \to \infty$. For any $k \geqslant 1$ define

$$\xi_{nk} = \begin{pmatrix} \xi_{1nk} \\ \vdots \\ \xi_{dnk} \end{pmatrix} = A_n \begin{pmatrix} e_{1k} \\ \vdots \\ e_{dk} \end{pmatrix}$$

and, for any $j = 1, \dots, d$ set

$$F_{jn} = \sum_{k=1}^{\infty} \lambda_{jnk} I_2\bigl(\xi_{jnk}^{\otimes 2}\bigr) = \sum_{k=1}^{\infty} \lambda_{jnk} \bigl[\tilde{\xi}_{jnk}^2 - \|\xi_{jnk}\|^2\bigr],$$

where $\tilde{\xi}_{jnk} = I_1(\xi_{jnk}) = X(\xi_{jnk})$ and λ_{jnk} are real numbers which will be specified later. We plan to use Theorem 5.2 to study the convergence of the random vectors $F_n = (F_{1n} \dots F_{dn})$. For this we can follow the approach of Section 7.1, the main extra work being to prove the existence of a uniform bound for the negative moments of the Malliavin covariance matrices. We have

$$DF_{jn} = 2 \sum_{k=1}^{\infty} \lambda_{jnk} \tilde{\xi}_{jnk} \xi_{jnk}.$$

Thus

$$\langle DF_{in}, DF_{jn} \rangle_{\mathfrak{H}} = 4 \sum_{k=1}^{\infty} \lambda_{ink} \lambda_{jnk} \tilde{\xi}_{ink} \tilde{\xi}_{jnk} \alpha_{ijn}, \tag{7.22}$$

where α_{ijn} is the (i, j)th entry of the matrix $\alpha_n = A_n A_n^T$. Consider the matrix $\beta_n := (\beta_{ijn})_{1 \leqslant i,j \leqslant d}$ given by

$$\beta_{ijn} := 4 \sum_{k=1}^{\infty} \lambda_{ink} \lambda_{jnk} \tilde{\xi}_{ink} \tilde{\xi}_{jnk}.$$

Then from (7.22), we see that $\langle DF_n, DF_n \rangle = (\langle DF_{in}, DF_{jn} \rangle_{\mathfrak{H}})_{1 \leqslant i,j \leqslant d}$ is the Hadamard product of the nonnegative definite matrices α_n and β_n. By the Oppenheim's inequality for Hadamard product, and taking into account that $\det(\alpha_n)$ converges to one, there exists a constant $c > 0$ such that

$$\det\bigl(\langle DF_n, DF_n \rangle\bigr) \geqslant \det(\alpha_n) \prod_{j=1}^{d} \beta_{jjn} \geqslant c \prod_{j=1}^{d} \beta_{jjn},$$

for all n. Note $\beta_{jjn} = 4 \sum_{k=1}^{\infty} \lambda_{jnk}^2 (\tilde{\xi}_{jnk})^2$ and $\xi_{jnk} \to e_{jk}$. Thus we can follow Section 7.1 to verify the conditions that allow us to apply Theorem 5.2. We will write down the theorem and omit the details. In the following we denote by ϕ_σ the density of the law $N(0, \operatorname{diag}(\sigma_1^2, \dots, \sigma_d^2))$ and $\partial_{\alpha_1} \cdots \partial_{\alpha_d} f(x) = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} f(x)$ with $|\alpha| = \alpha_1 + \cdots + \alpha_d$.

Theorem 7.7. *Let A_n be a sequence of $d \times d$ invertible matrices such that $A_n \to I$ and let $F_n = (F_{1n}, \dots, F_{dn})$ be defined as above. We assume the λ_{jnk} satisfy the following conditions for any $1 \leqslant j \leqslant d$.*

(i) $\sigma_j^2 := \lim_{n\to\infty} \sum_{k=1}^{\infty} \lambda_{jnk}^2 > 0$;
(ii) $\lim_{n\to\infty} \sum_{k=1}^{\infty} \lambda_{jnk}^4 = 0$;
(iii) $\inf_{j,n}(\sup_{k>6m+6(\lfloor \frac{m}{2} \rfloor \vee 1)} |\lambda_{jnk}|\sqrt{k}) > 0$ *for some integer $m \geqslant 0$.*

Then, each F_n admits a density function $f_{F_n} \in C^m(\mathbb{R}^d)$. Furthermore, for any $\alpha = (\alpha_1, \dots, \alpha_d)$, with $|\alpha| \leqslant m$, the derivatives of $\partial_{\alpha_1} \cdots \partial_{\alpha_d} f_{F_n}$ converge uniformly to the derivatives of $\partial_{\alpha_1} \cdots \partial_{\alpha_d} \phi_\sigma$ with a rate given by

$$\sup_{x\in\mathbb{R}} \left|\partial_{\alpha_1} \cdots \partial_{\alpha_d} f_{F_n}(x) - \partial_{\alpha_1} \cdots \partial_{\alpha_d} \phi_\sigma(x)\right| \leqslant C \sum_{j=1}^{d} \left[\left(\sum_{k=1}^{\infty} \lambda_{jnk}^4 \right)^{\frac{1}{2}} + \left| \sum_{i=1}^{\infty} \lambda_{jnk}^2 - \sigma^2 \right|^{\frac{1}{2}} \right],$$

where C is a constant depending only on m and the infimum appearing in condition (iii).

Acknowledgments

We would like to thank the anonymous referee's constructive comments on our paper.

Appendix A

In this section, we present the omitted proofs and some technical results.

Proof of Lemma 2.1. Since $\int_{-\infty}^{\infty} \{h(y) - E[h(N)]\} e^{-y^2/(2\sigma^2)}\, dy = 0$, we have

$$\int_{-\infty}^{x} \{h(y) - E[h(N)]\} e^{-y^2/(2\sigma^2)}\, dy = -\int_{x}^{\infty} \{h(y) - E[h(N)]\} e^{-y^2/(2\sigma^2)}\, dy.$$

Hence

$$\left| \int_{-\infty}^{x} \{h(y) - E[h(N)]\} e^{-y^2/(2\sigma^2)}\, dy \right| \leqslant \int_{|x|}^{\infty} \left[a y^k + b + E|h(N)| \right] e^{-y^2/(2\sigma^2)}\, dy.$$

By using the representation (2.18) of f_h and Stein's equation (2.15) we have

$$\left|f_h'(x)\right| \leqslant \left|h(x) - E[h(N)]\right| + \frac{|x|}{\sigma^2} e^{x^2/(2\sigma^2)} \left| \int_{-\infty}^{x} \{h(y) - E[h(N)]\} e^{-y^2/(2\sigma^2)}\, dy \right|$$

$$\leqslant a|x|^k + b + E|h(N)| + \frac{1}{\sigma^2} e^{x^2/(2\sigma^2)} \int_{|x|}^{\infty} y[ay^k + b + E|h(N)|]e^{-y^2/(2\sigma^2)}\,dy$$

$$= a|x|^k + (b + E|h(N)|)\left(1 + \frac{1}{\sigma^2} s_1(x)\right) + \frac{a}{\sigma^2} s_{k+1}(x), \tag{A.1}$$

where we let $s_k(x) = e^{x^2/(2\sigma^2)} \int_{|x|}^{\infty} y^k e^{-y^2/(2\sigma^2)}\,dy$ for any integer $k \geqslant 0$.

Note that $E|h(N)| \leqslant aE|N|^k + b \leqslant C_k a\sigma^k + b$ and

$$s_1(x) = e^{x^2/(2\sigma^2)} \int_x^{\infty} y e^{-\frac{y^2}{2\sigma^2}}\,dy = \sigma^2$$

for all $x \in \mathbb{R}$. Using integration by parts, we see by induction that for any integer $k \geqslant 1$,

$$\begin{aligned} s_{k+1}(x) &= e^{x^2/(2\sigma^2)} \int_{|x|}^{\infty} y^{k+1} e^{-y^2/(2\sigma^2)}\,dy \\ &= \sigma^2 e^{x^2/(2\sigma^2)} \int_{|x|}^{\infty} y^k\, d\big(-e^{-y^2/(2\sigma^2)}\big) \\ &= \sigma^2\big[|x|^k + k\, s_{k-1}(x)\big]. \end{aligned}$$

Then if $k \geqslant 1$ is even, we have

$$s_{k+1}(x) \leqslant C_k\sigma^2\big[|x|^k + \sigma^2|x|^{k-2} + \cdots + \sigma^{k-2} s_1(x)\big] \leqslant C_k\sigma^2 \sum_{i=0}^{k} \sigma^{k-i}|x|^i.$$

If $k \geqslant 1$ is odd, we have

$$s_{k+1}(x) \leqslant C_k\sigma^2\big[|x|^k + \sigma^2|x|^{k-2} + \cdots + \sigma^{k-1}\big(|x| + s_0(x)\big)\big] \leqslant C_k\sigma^2 \sum_{i=0}^{k} \sigma^{k-i}|x|^i,$$

where we used the fact that $s_0(x) \leqslant s_0(0) = \sqrt{\frac{\pi}{2}}\sigma$ for all $x \in \mathbb{R}$ (indeed, when $x \geqslant 0$ we have $s_0'(x) = \frac{x}{\sigma^2} e^{x^2/(2\sigma^2)} \int_x^{\infty} e^{-y^2/(2\sigma^2)}\,dy - 1 \leqslant e^{x^2/(2\sigma^2)} \int_x^{\infty} \frac{y}{\sigma^2} e^{-\frac{y^2}{2\sigma^2}}\,dy - 1 = 0$; similarly when $x < 0$, $s_0'(x) \geqslant 0$). Putting the above estimates into (A.1) we complete the proof. □

Proof of Lemma 3.5. We shall prove these properties by induction. From $T_1 = T_2 = 0$, (3.17) and (3.19) we know that $T_3 = D_u^2\delta_u$, with $J_3 = \{(0, 0, 1)\}$; and $T_4 = \delta_u D_u^2\delta_u + D_u^3\delta_u$, with $J_4 = \{(1, 0, 1, 0), (0, 0, 0, 1)\}$. Now suppose the statement is true for all T_l with $l \leqslant k - 1$ for $k \geqslant 5$. We want to prove the multi-indices of T_k satisfy (a)–(c). This will be done by studying the three operations, $\delta_u T_{k-1}$, $D_u T_{k-1}$ and $\partial_\lambda H_{k-1}(D_u\delta_u, \delta_u) D_u^2\delta_u$, in expression (3.19).

For the term $\partial_\lambda H_{k-1}(D_u\delta_u,\delta_u)D_u^2\delta_u$, we observe from (3.17) that

$$\partial_\lambda H_{k-1}(D_u\delta_u,\delta_u)D_u^2\delta_u = D_u^2\delta_u \sum_{1\leqslant i\leqslant\lfloor(k-1)/2\rfloor} i c_{k-1,i}\delta_u^{k-1-2i}(D_u\delta_u)^{i-1},$$

whose terms have multi-indices $(k-1-2i, i-1, 1, 0, \dots, 0) \in \mathbb{N}^k$ for $1 \leqslant i \leqslant \lfloor\frac{k-1}{2}\rfloor$. Then, it is straightforward to check that these multi-indices satisfy (a), (b) and (c).

The term $\delta_u T_{k-1}$ shifts the multi-index $(i_0, i_1, \dots, i_{k-2}) \in J_{k-1}$ to $(i_0+1, i_1, \dots, i_{k-2}, 0) \in \mathbb{N}^k$, which obviously satisfies (a), (b) and (c), due to the induction hypothesis.

The third term $D_u T_{k-1}$ shifts the multi-index $(i_0, i_1, \dots, i_{k-2}) \in J_{k-1}$ to either $\alpha = (i_0-1, i_1+1, \dots, i_{k-2}, 0) \in \mathbb{N}^k$ if $i_0 \geqslant 1$, or to

$$\beta = \begin{cases} (i_0, i_1, \dots, i_{j_0}-1, i_{j_0+1}+1, \dots, i_{k-2}, 0), & \text{for } 1 \leqslant j_0 \leqslant k-3; \\ (i_0, i_1, \dots, i_{j_0}-1, 1), & \text{for } j_0 = k-2, \end{cases}$$

if $i_{j_0} \geqslant 1$. It is easy to check that β satisfies properties (a), (b) and (c) and α satisfies properties (b) and (c). We are left to verify that α satisfies property (c). That is, we want to show that

$$1 + \sum_{j=1}^{k-2} i_j \leqslant \left\lfloor \frac{k-1}{2} \right\rfloor. \tag{A.2}$$

If k is odd, say $k = 2m+1$ for some $m \geqslant 2$, (A.2) is true because $(i_0, i_1, \dots, i_{k-2}) \in J_{k-1}$, which implies by induction hypothesis that $\sum_{j=1}^{k-2} i_j \leqslant \lfloor\frac{k-2}{2}\rfloor = m-1$. If k is even, say $k = 2m+2$, (A.2) is true because the following claim asserts that if $i_0 \geqslant 1$, then $\sum_{j=1}^{k-2} i_j < \lfloor\frac{k-2}{2}\rfloor = m$.

Claim. *For $(i_0, i_1, \dots, i_{2m}) \in J_{2m+1}$ with $m \geqslant 1$, if $\sum_{j=1}^{2m} i_j = m$ then $i_0 = 0$.*

Indeed, suppose $(i_0, i_1, \dots, i_{2m}) \in J_{2m+1}$, $\sum_{j=1}^{2m} i_j = m$ and $i_0 \geqslant 1$. We are going to show that leads to a contradiction. First notice that $i_1 \geqslant 1$, otherwise $i_1 = 0$ and $\sum_{j=2}^{2m} i_j = m$, which is not possible because

$$i_0 + 2m \leqslant i_0 + \sum_{j=1}^{2m} j i_j \leqslant 2m.$$

Also, we must have $i_{2m} = 0$, because otherwise property (a) implies $i_{2m} = 1$ and $i_0 = i_1 = \dots = i_{2m-1} = 0$. Now we trace back to its parent multi-indices in J_{2m} by reversing the three operations. Of the three operations, we can exclude $\partial_\lambda H_{2m}(D_u\delta_u,\delta_u)D_u^2\delta_u$ and $\delta_u T_{2m}$, because $\partial_\lambda H_{2m}(D_u\delta_u,\delta_u)D_u^2\delta_u$ generates $(2m-2j, j-1, 1, 0, \dots, 0)$ with $1 \leqslant j \leqslant m$, where j must be m; and $\delta_u T_{2m}$ traces it back to $(i_0-1, i_1, \dots, i_{2m-1}) \in J_{2m}$, where $i_1 + \dots + i_{2m-1} = m > \lfloor\frac{2m-1}{2}\rfloor$. Therefore, its parent multi-index in J_{2m} must come from the operation $D_u T_{2m}$ and hence must be $(i_0+1, i_1-1, \dots, i_{2m-1}) \in J_{2m}$. Note that for this multi-index, $i_1 - 1 + \dots + i_{2m-1} =$

$m-1$. Repeating the above process we will end up at $(i_0+i_1, 0, i_2 \dots, i_{2m-i_1}) \in J_{2m+1-i_1}$ with $i_2+\dots+i_{2m-i_1}=m-i_1$, which contradicts the property (b) of J_{2m+1-i_1} because

$$i_0+2m-i_1 \leqslant i_0+i_1+\sum_{j=2}^{2m-i_1} j i_j \leqslant 2m-i_1. \qquad \square$$

Recall that we denote $D_{DF}w^{-1}=\langle Dw^{-1}, DF\rangle_{\mathfrak{H}}$ and $D^k_{DF}w^{-1}=\langle D(D^{k-1}_{DF}w^{-1}), DF\rangle_{\mathfrak{H}}$ for any $k \geqslant 2$. The following lemma estimates the $L^p(\Omega)$ norms of $D^k_{DF}w^{-1}$.

Lemma A.1. *Let $F=I_q(f)$ with $q \geqslant 2$ satisfying $E[F^2]=\sigma^2$. For any $\beta \geqslant 1$ we define and $M_\beta=(E\|DF\|_{\mathfrak{H}}^{-\beta})^{1/\beta}$. Set $w=\|DF\|^2_{\mathfrak{H}}$.*

(i) *If $M_\beta<\infty$ for some $\beta \geqslant 6$, then for any $1 \leqslant r \leqslant \frac{2\beta}{\beta+6}$*

$$\|D_{DF}w^{-1}\|_r \leqslant CM_\beta^3\|q\sigma^2-w\|_2. \tag{A.3}$$

(ii) *If $k \geqslant 2$ and $M_\beta<\infty$ for some $\beta \geqslant 2k+4$, then for any $1<r<\frac{2\beta}{\beta+2k+4}$*

$$\|D^k_{DF}w^{-1}\|_r \leqslant C(\sigma^{2k-2}\vee 1)(M_\beta^{k+2}\vee 1)\|q\sigma^2-w\|_2. \tag{A.4}$$

(iii) *If $k \geqslant 1$ and $M_\beta<\infty$ for any $\beta>k+2$, then for any $1<r<\frac{\beta}{k+2}$*

$$\|D^k_{DF}w^{-1}\|_r \leqslant C(\sigma^{2k}\vee 1)(M_\beta^{k+2}\vee 1). \tag{A.5}$$

Proof. Note that $D_{DF}w^{-1}=\langle Dw^{-1}, DF\rangle_{\mathfrak{H}}=-2w^{-2}\langle D^2F\otimes_1 DF, DF\rangle$. Then

$$|D_{DF}w^{-1}| \leqslant 2w^{-\frac{3}{2}}\|D^2F\otimes_1 DF\|_{\mathfrak{H}}.$$

Applying Hölder's inequality with $\frac{1}{r}=\frac{1}{p}+\frac{1}{2}$, yields

$$\|D_{DF}w^{-1}\|_r \leqslant 2\big(E\big(w^{-\frac{3p}{2}}\big)\big)^{\frac{1}{p}}\|D^2F\otimes_1 DF\|_2,$$

which implies (A.3) by choosing $p \leqslant \beta/3$ and taking into account (4.3). Notice that we need $1 \geqslant \frac{1}{r} \geqslant \frac{3}{\beta}+\frac{1}{2}=\frac{\beta+6}{2\beta}$.

Consider now the case $k \geqslant 2$. From the pattern indicated by the first three terms,

$$\begin{aligned}
D_{DF}w^{-1} &= \langle Dw^{-1}, DF\rangle_{\mathfrak{H}},\\
D^2_{DF}w^{-1} &= \langle D^2w^{-1}, (DF)^{\otimes 2}\rangle_{\mathfrak{H}^{\otimes 2}}+\langle Dw^{-1}\otimes DF, D^2F\rangle_{\mathfrak{H}^{\otimes 2}},\\
D^3_{DF}w^{-1} &= \langle D^3w^{-1}, (DF)^{\otimes 3}\rangle_{\mathfrak{H}^{\otimes 3}}+3\langle D^2w^{-1}\otimes DF, D^2F\otimes DF\rangle_{\mathfrak{H}^{\otimes 3}}\\
&\quad+\langle Dw^{-1}\otimes D^2F, D^2F\otimes DF\rangle_{\mathfrak{H}^{\otimes 3}}+\langle Dw^{-1}\otimes (DF)^{\otimes 2}, D^3F\rangle_{\mathfrak{H}^{\otimes 3}},
\end{aligned}$$

we can prove by induction that

$$\left|D_{DF}^{k} w^{-1}\right| \leqslant C \sum_{i=1}^{k} \left\| D^{i} w^{-1} \right\|_{\mathfrak{H}^{\otimes i}} \|DF\|_{\mathfrak{H}}^{i} \left(\sum_{\sum_{j=1}^{k} i_j = k-i} \prod_{j=1}^{k} \left\| D^{j} F \right\|_{\mathfrak{H}^{\otimes j}}^{i_j} \right).$$

By (2.12), for any $p > 1$, $\|D^j F\|_p \leqslant C\|F\|_2 = C\sigma$. Applying Hölder's inequality and assuming that $s > r$, we have,

$$\left\| D_{DF}^{k} w^{-1} \right\|_{r} \leqslant C \sum_{i=1}^{k} \left\| \left\| D^{i} w^{-1} \right\|_{\mathfrak{H}^{\otimes i}} \|DF\|_{\mathfrak{H}}^{i} \right\|_{s} \sigma^{k-i}. \tag{A.6}$$

We are going to see that $\|DF\|_{\mathfrak{H}}^{i}$ will contribute to compensate the singularity of $\|D^i w^{-1}\|_{\mathfrak{H}^{\otimes i}}$. First by induction one can prove that for $1 \leqslant i \leqslant m$, $D^i w^{-1}$ has the following expression

$$D^{i} w^{-1} = \sum_{l=1}^{i} (-1)^{l} \sum_{(\alpha,\beta)\in I_{i,l}} w^{-(l+1)} \bigotimes_{j=1}^{l} \left(D^{\alpha_j} F \otimes_1 D^{\beta_j} F \right), \tag{A.7}$$

where $I_{i,l} = \{(\alpha, \beta) \in \mathbb{N}^{2l} : \alpha_j + \beta_j \geqslant 3, \sum_{j=1}^{l} (\alpha_j + \beta_j) = i + 2l\}$. In fact, for $i = 1$,

$$D w^{-1} = -2 w^{-2} D^2 F \otimes_1 DF,$$

which is of the above form because $I_{1,1} = \{(1, 2), (2, 1)\}$. Suppose that (A.7) holds for some $i \leqslant m - 1$. Then,

$$\begin{aligned} D^{i+1} w^{-1} = {} & \sum_{l=1}^{i} (-1)^{l+1} 2(l+1) \sum_{(\alpha,\beta)\in I_{i,l}} w^{-(l+2)} \left(D^2 F \otimes_1 DF \right) \bigotimes_{j=1}^{l} \left(D^{\alpha_j} F \otimes_1 D^{\beta_j} F \right) \\ & + \sum_{l=1}^{i} (-1)^{l} \sum_{(\alpha,\beta)\in I_{i,l}} w^{-(l+1)} \sum_{h=1}^{l} \left(D^{\alpha_j+1} F \otimes_1 D^{\beta_j} F + D^{\alpha_j} F \otimes_1 D^{\beta_j+1} F \right) \\ & \times \bigotimes_{j=1, j\neq h}^{l} \left(D^{\alpha_j} F \otimes_1 D^{\beta_j} F \right), \end{aligned}$$

which is equal to

$$\sum_{l=1}^{i+1} (-1)^{l} \sum_{(\alpha,\beta)\in I_{i+1,l}} w^{-(l+1)} \bigotimes_{j=1}^{l} \left(D^{\alpha_j} F \otimes_1 D^{\beta_j} F \right).$$

From (A.7) for any $i = 1, \dots, k$ we can write

$$\left\| D^{i} w^{-1} \right\|_{\mathfrak{H}^{\otimes i}} \|DF\|_{\mathfrak{H}}^{i} \leqslant \sum_{l=1}^{i} w^{-(l+1)+\frac{i}{2}} \sum_{(\alpha,\beta)\in I_{i,l}} \prod_{j=1}^{l} \left\| D^{\alpha_j} F \otimes_1 D^{\beta_j} F \right\|_{\mathfrak{H}^{\otimes \alpha_j+\beta_j-2}}, \tag{A.8}$$

where $I_{i,l} = \{(\alpha, \beta) \in \mathbb{N}^l \times \mathbb{N}^l : \alpha_j + \beta_j \geqslant 3, \sum_{j=1}^l (\alpha_j + \beta_j) = i + 2l\}$. Note that by (2.12),

$$\left\| D^{\alpha_j} F \otimes_1 D^{\beta_j} F \right\|_p \leqslant C \|F\|_2^2 = C\sigma^2$$

for all $p \geqslant 1$ and all α_j, β_j. This inequality will be applied to all but one of the contraction terms in the product $\prod_{j=1}^l \|D^{\alpha_j} F \otimes_1 D^{\beta_j} F\|_{\mathfrak{H}^{\otimes \alpha_j + \beta_j - 2}}$. We decompose the sum in (A.8) into two parts. If the index l satisfies $l \leqslant \frac{i}{2} - 1$, then the exponent of w is nonnegative, and the p norm of w can be estimated by a constant times σ^2, while for $\frac{i}{2} - 1 < l$ this exponent is negative. Then, using Hölder's inequality and assuming that $\frac{1}{s} = \frac{1}{p} + \frac{1}{2}$, we obtain

$$\begin{aligned} &\left\| \left\| D^i w^{-1} \right\|_{\mathfrak{H}^{\otimes i}} \|DF\|_{\mathfrak{H}}^i \right\|_s \\ &\quad \leqslant C \left(\mathbf{1}_{\{i \geqslant 2\}} \sigma^{i-2} + \sum_{\frac{i}{2} - 1 < l \leqslant i} \left\| w^{-(l+1) + \frac{i}{2}} \right\|_p \sigma^{2(l-1)} \right) \left\| D^{\alpha_1} F \otimes_1 D^{\beta_1} F \right\|_2. \end{aligned} \tag{A.9}$$

Note that for $l \leqslant i \leqslant k$, $l + 1 - \frac{i}{2} \leqslant \frac{k}{2} + 1$. Therefore, for $\frac{i}{2} - 1 < l \leqslant i$,

$$\left\| w^{-(l+1) + \frac{i}{2}} \right\|_p = M_{2(l+1-\frac{i}{2})p}^{2l+2-i} \leqslant M_{(k+2)p}^{2l+2-i} \leqslant M_{(k+2)p}^{k+2} \vee 1.$$

Therefore, using (4.3) we obtain

$$\left\| \left\| D^i w^{-1} \right\|_{\mathfrak{H}^{\otimes i}} \|DF\|_{\mathfrak{H}}^i \right\|_s \leqslant C \left(\left(\sigma^{2i-2} \vee 1 \right) \left(M_{(k+2)p}^{k+2} \vee 1 \right) \right) \left\| q\sigma^2 - w \right\|_2. \tag{A.10}$$

Combining (A.10) and (A.6) and choosing p such that $(k+2)p \leqslant \beta$ we get (A.4). Note that we need

$$1 > \frac{1}{r} > \frac{k+2}{\beta} + \frac{1}{2} = \frac{\beta + 2k + 4}{2\beta},$$

which holds if $1 < r < \frac{2\beta}{\beta + 2k + 4}$. The proof of part (iii) is similar and omitted. □

The next lemma gives estimates on $D_u^k \delta_u$ for $k \geqslant 0$.

Lemma A.2. *Let $F = I_q(f)$ with $q \geqslant 2$ satisfying $E[F^2] = \sigma^2$. For any $\beta \geqslant 1$ we define $M_\beta = (E\|DF\|_{\mathfrak{H}}^{-\beta})^{1/\beta}$ and denote $w = \|DF\|_{\mathfrak{H}}^2$.*

(i) *If $M_\beta < \infty$ for some $\beta > 3$, then for any $1 < s < \frac{\beta}{3}$,*

$$\|\delta_u\|_s \leqslant C \left(\sigma^2 \vee 1 \right) \left(M_\beta^3 \vee 1 \right). \tag{A.11}$$

(ii) *If $k \geqslant 1$ and $M_\beta < \infty$ for some $\beta > 3k + 3$, then for any $1 < s < \frac{\beta}{3k+3}$,*

$$\left\| D_u^k \delta_u \right\|_s \leqslant C_\sigma \left(M_\beta^{3k+3} \vee 1 \right). \tag{A.12}$$

(iii) *If $k \geqslant 2$ and $M_\beta < \infty$ for some $\beta > 6k+6$, then for any $1 < s < \frac{2\beta}{\beta+6k+6}$,*

$$\left\| D_u^k \delta_u \right\|_s \leqslant C_\sigma \left(M_\beta^{3k+3} \vee 1 \right) \left\| q\sigma^2 - w \right\|_2. \tag{A.13}$$

Proof. Recall that $\delta_u = qFw^{-1} - D_{DF} w^{-1}$. Then for any $r > s$,

$$\|\delta_u\|_s \leqslant C\left(\sigma \left\|w^{-1}\right\|_r + \left\|D_{DF} w^{-1}\right\|_s\right).$$

Then, $\|w^{-1}\|_r = M_{2r}^2$ and the result follows by applying Lemma A.1(iii) with $k = 1$ and by choosing $r < \frac{\beta}{3}$.

To show (ii) and (iii) we need to find a useful expression for $D_u^k \delta_u$. Consider the operator $D_u = w^{-1} D_{DF}$. We claim that for any $k \geqslant 1$ the iterated operator D_u^k can be expressed as

$$D_u^k = \sum_{l=1}^{k} w^{-l} \sum_{\mathbf{i} \in I_{l,k}} b_{\mathbf{i}} \left[\prod_{j=1}^{k-l} D_{DF}^{i_j} w^{-1} \right] D_{DF}^{i_0}, \tag{A.14}$$

where $b_{\mathbf{i}} > 0$ are real numbers and

$$I_{l,k} = \left\{ \mathbf{i} = (i_0, i_1, \dots, i_l)\colon\ i_0 \geqslant 1,\ i_j \geqslant 0\ \forall j = 1, \dots, l,\ \sum_{j=0}^{k-l} i_j = k \right\}.$$

In fact, this is clearly true for $k = 1$. Assume (A.14) holds for a given k. Then

$$\begin{aligned}
D_u^{k+1} &= w^{-1} D_{DF} D^k u \\
&= \sum_{l=1}^{k} l w^{-l} D_{DF} w^{-1} \sum_{\mathbf{i} \in I_{l,k}} b_{\mathbf{i}} \left[\prod_{j=1}^{k-l} D_{DF}^{i_j} w^{-1} \right] D_{DF}^{i_0} \\
&\quad + \sum_{l=1}^{k} w^{-l-1} \sum_{\mathbf{i} \in I_{l,k}} b_{\mathbf{i}} \left[\sum_{h=1}^{k-l} D_{DF}^{i_h+1} w^{-1} \prod_{j=1, j\neq h}^{k-l} D_{DF}^{i_j} w^{-1} \right] D_{DF}^{i_0} \\
&\quad + \sum_{l=1}^{k} w^{-l-1} \sum_{\mathbf{i} \in I_{l,k}} b_{\mathbf{i}} \left[\prod_{j=1}^{k-l} D_{DF}^{i_j} w^{-1} \right] D_{DF}^{i_0+1}.
\end{aligned}$$

Shifting the indexes, this can be written as

$$\begin{aligned}
D_u^{k+1} &= \sum_{l=1}^{k} l w^{-l} D_{DF} w^{-1} \sum_{\mathbf{i} \in I_{l,k}} b_{\mathbf{i}} \left[\prod_{j=1}^{k-l} D_{DF}^{i_j} w^{-1} \right] D_{DF}^{i_0} \\
&\quad + \sum_{l=2}^{k+1} w^{-l} \sum_{\mathbf{i} \in I_{l-1,k}} b_{\mathbf{i}} \left[\sum_{h=1}^{k+1-l} D_{DF}^{i_h+1} w^{-1} \prod_{j=1, j\neq h}^{k+1-l} D_{DF}^{i_j} w^{-1} \right] D_{DF}^{i_0}
\end{aligned}$$

$$+\sum_{l=2}^{k+1} w^{-l} \sum_{\mathbf{i}\in I_{l-1,k}} b_{\mathbf{i}} \left[\prod_{j=1}^{k+1-l} D_{DF}^{i_j} w^{-1}\right] D_{DF}^{i_0+1}.$$

It easy to check that this coincides with

$$\sum_{l=1}^{k+1} w^{-l} \sum_{\mathbf{i}\in I_{l,k+1}} b_{\mathbf{i}} \left[\prod_{j=1}^{k+1-l} D_{DF}^{i_j} w^{-1}\right] D_{DF}^{i_0}.$$

Also, note that $\delta_u = qFw^{-1} + D_{DF}w^{-1}$ and

$$D_{DF}\delta_u = q + qFD_{DF}w^{-1} + D_{DF}^2 w^{-1}.$$

By induction we can show that for any $i_0 \geqslant 1$

$$D_{DF}^{i_0}\delta_u = q\delta_{1i_0} + q\sum_{j=1}^{i_0-1} c_{i,j} D_{DF}^{i_0-1-j} w D_{DF}^{j} w^{-1} + qFD_{DF}^{i_0} w^{-1} + D_{DF}^{i_0+1} w^{-1}, \qquad \text{(A.15)}$$

where δ_{1i_0} is the Kronecker symbol. Combining (A.14) and (A.15) we obtain

$$D_u^k\delta_u = \sum_{l=1}^{k} w^{-l} \sum_{\mathbf{i}\in I_{l,k}} b_{\mathbf{i}} \left[\prod_{j=1}^{k-l} D_{DF}^{i_j} w^{-1}\right]$$
$$\times \left[q\delta_{1i_0} + q\sum_{j=1}^{i_0-1} c_{i,0j} D_{DF}^{i_0-1-j} w D_{DF}^{j} w^{-1} + qFD_{DF}^{i_0} w^{-1} + D_{DF}^{i_0+1} w^{-1}\right].$$

Next we shall apply Hölder's inequality to estimate $\|D_u^k\delta_u\|_s$. Notice that for $l=k$, $i_0 = k \geqslant 2$. Therefore,

$$\begin{aligned}
\|D_u^k\delta_u\|_s &\leqslant C_\sigma \sum_{l=1}^{k-1}\sum_{\mathbf{i}\in I_{l,k}} \|w^{-l}\|_p \prod_{j=1}^{k-l} \|D_{DF}^{i_j} w^{-1}\|_{r_j} \Big(\delta_{1i_0} + \max_{1\leqslant h\leqslant i_0+1} \|D_{DF}^{h} w^{-1}\|_{r_0}\Big) \\
&\quad + C_\sigma \|w^{-k}\|_p \max_{1\leqslant h\leqslant k+1} \|D_{DF}^{h} w^{-1}\|_{\rho_0} \\
&= B_1 + B_2,
\end{aligned}$$

assuming that for $l = 1,\dots,k-1$, $\frac{1}{s} > \frac{1}{p} + \sum_{j=0}^{k-l}\frac{1}{r_j}$ and $\frac{1}{s} > \frac{1}{p} + \frac{1}{\rho_0}$, and where C_σ denotes a function of σ of the form $C(1+\sigma^M)$.

Let us consider first the term B_1. Note that if $i_0 = 1$ there is at least one factor of the form $\|D_{DF}^{r_j} w^{-1}\|_{r_j}$ in the above product, because $\sum_{j=1}^{k-l} i_j = k-1 \geqslant 1$. Then, we will apply the inequality (A.4) to one of these factors and the inequality (A.5) to the remaining ones. The estimate (A.5) requires $\frac{1}{r_j} > \frac{i_j+2}{\beta}$ for $j = 1,\dots,k-l$ and $\frac{1}{r_0} > \frac{i_0+3}{\beta}$. On the other hand, the estimate (A.4) requires $\frac{1}{r_j} > \frac{i_j+2}{\beta} + \frac{1}{2}$ for $j = 1, \dots, k-l$ and $\frac{1}{r_0} > \frac{i_0+3}{\beta} + \frac{1}{2}$. Then, choosing p such that $2pl < \beta$, and taking into account that $\sum_{j=0}^{k-l} i_j = k$ we obtain the inequalities

$$\frac{1}{s} > \frac{1}{p} + \sum_{j=1}^{k-l} \frac{i_j+2}{\beta} + \frac{i_0+3}{\beta} + \frac{1}{2} > \frac{3k+3}{\beta} + \frac{1}{2}.$$

Hence, if $s < \frac{2\beta}{\beta+6k+6}$ we can write

$$\begin{aligned} B_1 &\leqslant C_\sigma \sum_{l=1}^{k-1} M_\beta^{2l} \prod_{j=1}^{k-l} \left(M_\beta^{i_j+2} \vee 1\right)\left(M_\beta^{i_0+3} \vee 1\right) \left\| q\sigma^2 - w^{-1} \right\|_2 \\ &\leqslant C_\sigma \left(M_\beta^{3k+3} \vee 1\right) \left\| q\sigma^2 - w^{-1} \right\|_2. \end{aligned}$$

For the term B_2 we use the estimate (A.4) assuming $2pk < \beta$ and

$$\frac{1}{s} > \frac{1}{p} + \frac{k+3}{\beta} + \frac{1}{2} > \frac{3k+3}{\beta} + \frac{1}{2}.$$

This leads to the same estimate and the proof of (A.13) is complete. To show the estimate (A.12) we proceed as before but using the inequality (A.5) for all the factors. In this case the summand $\frac{1}{2}$ does not appear and we obtain (A.12). □

References

[1] A.D. Barbour, Asymptotic expansions based on smooth functions in the central limit theorem, Probab. Theory Related Fields 72 (2) (1986) 289–303.
[2] A.R. Barron, O. Johnson, Fisher information inequalities and the central limit theorem, Probab. Theory Related Fields 129 (3) (2004) 391–409.
[3] S.G. Bobkov, G.P. Chistyakov, F. Götze, Fisher information and the central limit theorem, preprint, arXiv:1204.6650v1, 2012.
[4] S. Chatterjee, E. Meckes, Multivariate normal approximation using exchangeable pairs, ALEA Lat. Am. J. Probab. Math. Stat. 4 (2008) 257–283.
[5] L.H.Y. Chen, L. Goldstein, Q.-M. Shao, Normal Approximation by Stein's Method, Probab. Appl. (N. Y.), Springer-Verlag, Heidelberg, 2011.
[6] J. Hoffmann-Jøgensen, L.A. Shepp, R.M. Dudley, On the lower tail of Gaussian seminorms, Ann. Probab. 7 (1979) 319–342, MR0525057.
[7] Y. Hu, D. Nualart, Parameter estimation for fractional Ornstein–Uhlenbeck processes, Statist. Probab. Lett. 80 (11–12) (2010) 1030–1038.
[8] O. Johnson, Information Theory and the Central Limit Theorem, Imperial College Press, London, 2004, xiv+209 pp.
[9] Yu.A. Kutoyants, Statistical Inference for Ergodic Diffusion Processes, Springer Ser. Statist., Springer-Verlag, 2004.
[10] W.V. Li, Q.-M. Shao, Gaussian processes: inequalities, small ball probabilities and applications, in: Stochastic processes: theory and methods, in: Handbook of Statist., vol. 19, North-Holland, Amsterdam, 2001, pp. 533–597.
[11] R.S. Liptser, A.N. Shiryaev, Statistics of Random Processes. II. Applications, second edition, Appl. Math. (N. Y.), vol. 6, Springer-Verlag, 2001.
[12] E. Meckes, An infinitesimal version of Stein's method of exchangeable pairs, PhD thesis, Stanford University, 2006.
[13] S. Noreddine, I. Nourdin, On the Gaussian approximation of vector-valued multiple integrals, J. Multivariate Anal. 102 (6) (2011) 1008–1017.
[14] I. Nourdin, D. Nualart, G. Poly, Absolute continuity and convergence of densities for random vectors on Wiener chaos, Electron. J. Probab. 18 (22) (2013) 1–19.

[15] I. Nourdin, G. Peccati, Stein's method on Wiener chaos, Probab. Theory Related Fields 145 (1–2) (2009) 75–118.
[16] I. Nourdin, G. Peccati, Normal approximations with Malliavin calculus, in: From Stein's method to universality, in: Cambridge Tracts in Math., Cambridge University Press, 2012.
[17] I. Nourdin, G. Peccati, G. Reinert, Second order Poincaré inequalities and CLTs on Wiener space, J. Funct. Anal. 257 (2) (2009) 593–609.
[18] I. Nourdin, G. Peccati, G. Reinert, Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos, Ann. Probab. 38 (5) (2010) 1947–1985.
[19] I. Nourdin, G. Peccati, A. Réveillac, Multivariate normal approximation using Stein's method and Malliavin calculus, Ann. Inst. Henri Poincaré Probab. Stat. 46 (1) (2010) 45–58.
[20] I. Nourdin, G. Peccati, Y. Swan, Entroy and the fourth moment phenomenon, Preprint, arXiv:1304.1255v1, 2013.
[21] I. Nourdin, G. Poly, Convergence in total variation on Wiener chaos, Stochastic Process. Appl. 123 (2) (2013) 651–674.
[22] D. Nualart, The Malliavin Calculus and Related Topics, 2nd edition, Springer-Verlag, 2006.
[23] D. Nualart, S. Ortiz-Latorre, Central limit theorems for multiple stochastic integrals and Malliavin calculus, Stochastic Process. Appl. 118 (4) (2008) 614–628.
[24] D. Nualart, G. Peccati, Central limit theorems for sequences of multiple stochastic integrals, Ann. Probab. 33 (1) (2005) 177–193.
[25] N. Privault, Cumulant operators and moments of the Itô and Skorohod integrals, preprint.
[26] M. Raič, A multivariate CLT for decomposable random vectors with finite second moments, J. Theoret. Probab. 17 (3) (2004) 573–603.
[27] G. Reinert, A. Röllin, Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition, Ann. Probab. 37 (6) (2009) 2150–2173.
[28] Y. Rinott, V. Rotar, On Edgeworth expansions for dependency-neighborhoods chain structures and Stein's method, Probab. Theory Related Fields 126 (4) (2003) 528–570.
[29] R. Shimizu, On Fishers amount of information for location family, in: G.P. Patil, et al. (Eds.), Statistical Distributions in Scientific Work, vol. 3, Reidel, 1975, pp. 305–312.
[30] C. Stein, 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. Vol. II: Probability theory, Univ. California Press, Berkeley, CA, 1972, pp. 583–602.