Survey on the fourth moment theorem, Stein's method and related topics

Seiichiro Kusuoka[1] (Tohoku University)

1 Introduction

The fourth moment theorem was originally introduced by Nualart and Peccati [11]. The theorem gives some equivalent conditions for a sequence of random variables belonging to a level of Wiener chaos to convergent to the standard normal distribution. The most surprising part of the theorem is that; if the variances of the sequence converge to 1, then the convergence to the standard normal distribution is equivalent to the convergence of the fourth moments of the sequence to 3. After that, Nualart and Ortiz-Latorre [10] gave another equivalent condition and made a clearer proof in their paper. Stimulated by Nualart and Ortiz-Latorre [10], Nourdin and Peccati [6] discovered a new method to estimate distances between the standard normal distribution and other distributions, and between the centered Gamma distributions and other distributions. The method is a combination of Stein's method and Malliavin calculus. Nourdin and Peccati's method enables us to prove a part of the fourth moment theorem in another way. Now applications and other versions of the fourth moment theorem and Stein's bound are considered.

In this talk, we review the fourth moment theorem and Stein's method mainly, give a short review of further results and related topics.

Now we give some useful information. A textbook [7] written by Nourdin and Peccati was published recently. This book covers from the elementary tools for this topic to the fourth moment theorem and the density estimates obtained by Stein's method. The latest results on this topic are found on the webpage:

http://www.iecn.u-nancy.fr/ nourdin/steinmalliavin.htm

Many of literatures (e.g. lecture notes, articles) are listed up on this webpage.

2 Preliminary on Wiener chaos

First we prepare the elementary things on Wiener chaos.

Let $(T,\mathscr{B})$ be a measurable space, μ a σ-finite measure on $(T,\mathscr{B})$ without atoms, and $H := L^2(T,\mathscr{B},\mu)$. We introduce the isonormal Gaussian process with respect to H. Let $W = \{W(h);\ h \in H\}$ be a family of random variables on a complete probability space $(\Omega,\mathscr{F},P)$.

Definition 2.1. *We call W is an isonormal Gaussian process (or Gaussian process on H) if the following conditions hold.*

(i) *W is a Gaussian family (or a Gaussian system), i.e. for $n \in \mathbb{N}$ and $h_1, h_2, \dots, h_n \in H$, the $\mathbb{R}^n$-valued random variable $(W(h_1), W(h_2), \dots, W(h_n))$ has an n-dimensional Gaussian distribution.*

(ii) *H is the Cameron-Martin space (or the reproducing kernel Hilbert space) of W, i.e.*

$$E[W(h)] = 0, \qquad h \in H, \tag{2.1}$$

$$E[W(g)W(h)] = (g,h)_H, \qquad g,h \in H. \tag{2.2}$$

Let $W(A) := W(\mathbb{I}_A)$ for $A \in \mathscr{B}$ and $\mu(A) < \infty$. Then, the law of $\mathbb{W}$ is also characterized by $\{W(A); A \in \mathscr{B}, \mu(A) < \infty\}$, since L^2-functions are approximated by simple functions (step functions, elementary functions). The following assertions follows immediately from Definition 2.1.

(i) $W(A)$ has the distribution $N(0,\mu(A))$ for $A \in \mathscr{B}$ such that $\mu(A) < \infty$.

[1] e-mail: kusuoka@math.kyoto-u.ac.jp

(ii) $W(A_1)$ and $W(A_2)$ are independent of each other for $A_1, A_2 \in \{A \in \mathscr{B}; \mu(A) < \infty\}$ such that $A_1 \cap A_2 = \emptyset$.

(iii) $A \mapsto W(A)$ is an $L^2(\Omega, \mathscr{F}, P)$-valued finitely additive measure on $(T, \mathscr{B})$.

Note that $A \mapsto W(A)$ is not σ-additive.

Now, we start with the construction of multiple stochastic integrals. Let $m \in \mathbb{N}$ and $\mathscr{B}_0 := \{A \in \mathscr{B}; \mu(A) < \infty\}$. We define the multiple stochastic integral $I_m(f)$ of $f \in L^2(T^m, \mathscr{B}^{\otimes m}, \mu^{\otimes m})$ as follows. Let $\mathcal{E}_m$ be the total set of the functions f such that

$$f(t_1, t_2, \dots, t_m) = \sum_{i_1, i_2, \dots, i_m = 1}^{n} a_{i_1, i_2, \dots, i_m} \mathbb{I}_{A_{i_1} \times A_{i_2} \times \cdots A_{i_m}}(t_1, t_2, \dots, t_m) \tag{2.3}$$

where $n \in \mathbb{N}$, $A_1, A_2, \dots, A_n$ are pairwise-disjoint sets in $\mathscr{B}_0$, and $a_{i_1, i_2, \dots, i_m} \in \mathbb{R}$ such that $a_{i_1, i_2, \dots, i_m} = 0$ if $i_k = i_l$ for some $k, l = 1, 2, \dots, n$. Note that $\mathcal{E}_m$ is a linear space. For f expressed as in (2.3) we define

$$I_m(f) := \sum_{i_1, i_2, \dots, i_m = 1}^{n} a_{i_1, i_2, \dots, i_m} W(A_{i_1}) W(A_{i_2}) \dots W(A_{i_m}).$$

For $f \in L^2(T^m, \mathscr{B}^{\otimes m}, \mu^{\otimes m})$, define the symmetrization $\widetilde{f}$ of f by

$$\widetilde{f}(t_1, t_2, \dots, t_m) = \frac{1}{m!} \sum_{\pi \in \mathfrak{S}_m} f(t_{\pi(1)}, t_{\pi(2)}, \dots, t_{\pi(m)})$$

where $\mathfrak{S}_m$ is the group of permutations of $\{1, 2, \dots, m\}$. Note that the mapping $f \mapsto \widetilde{f}$ from $L^2(\mu^{\otimes m})$ to itself is linear and continuous for each $m \in \mathbb{N}$. We call f symmetric if $f = \tilde{f}$. Let $H^{\odot n} := \{f \in H^{\otimes n};\ f$ is symmetric$\}$, and $\|f\|_{H^{\odot n}} := \sqrt{n!}\|f\|_{H^{\otimes n}}$ for $f \in H^{\odot n}$. Then, following properties hold.

Proposition 2.2. (i) *I_m is a linear mapping form $\mathcal{E}_m$ to $L^2(\Omega, \mathscr{F}, P)$.*

(ii) *For $f \in \mathcal{E}_m$, $I_m(f) = I_m(\tilde{f})$.*

(iii) *For $f \in \mathcal{E}_m$ and $g \in \mathcal{E}_q$,*

$$E[I_m(f) I_q(g)] = \begin{cases} 0 & \text{if} \quad m \neq q, \\ m!(\tilde{f}, \tilde{g})_{L^2(\mu^{\otimes m})} & \text{if} \quad m = q. \end{cases}$$

By the property (iii)

$$E[I_m(f)^2] = m! ||\tilde{f}||^2_{L^2(\mu^{\otimes m})} \leq m! ||f||^2_{L^2(\mu^{\otimes m})}. \tag{2.4}$$

By (2.4) we have

$$E[I_m(f)^2] = \|f\|^2_{H^{\odot n}}, \quad f \in H^{\odot n}.$$

The following lemma holds.

Lemma 2.3. *$\mathcal{E}_m$ is dense in $L^2(\mu^{\otimes m})$.*

By (2.4) and Lemma 2.3 we can extend I_m to a bounded linear operator from $L^2(T^m, \mathscr{B}^{\otimes m}, \mu^{\otimes m})$ to $L^2(\Omega, \mathscr{F}, P)$. The extension of I_m also satisfies the properties in Proposition 2.2 again.

By using Hermite polynomial, we have the following theorem. The theorem is called the Wiener-Chaos expansion.

Theorem 2.4. *Assume that $\mathscr{F}$ is the σ-field generated by $W = \{W(h);\ h \in H\}$. Then, for any $F \in L^2(\Omega, \mathscr{F}, P)$, there exist symmetric functions $\{f_n \in H^{\odot n}; n = 0, 1, 2, \dots\}$ such that $f_0 = E[F]$ and*

$$F = \sum_{n=0}^{\infty} I_n(f_n).$$

The functions $\{f_n\}$ are uniquely determined by F.

3 Preliminary on H-derivative

Let $\mathcal{P}$ be the class of the random variables F such that; there exist $n \in \mathbb{N}$, a polynomial function f on $\mathbb{R}^n$, $h_1, h_2, \dots, h_n \in H$, and F is expressed by

$$F = f(W(h_1), W(h_2), \dots, W(h_n)). \tag{3.1}$$

Then, the following lemma holds.

Lemma 3.1. *$\mathcal{P}$ is dense in $L^p(\Omega, \mathscr{F}, P)$ for $p \in [1, \infty)$.*

We define the H-derivative operator D as follows.

Definition 3.2. *For $F \in \mathcal{P}$ expressed as in (3.1), define the H-valued random variable DF of F by*

$$DF = \sum_{i=1}^{n} \partial_i f(W(h_1), W(h_2), \dots, W(h_n)) h_i.$$

We call DF by the H-derivative (or Malliavin's derivative) of F.

It is easy to see that D is linear on $\mathcal{P}$ and D maps $\mathcal{P}$ into $L^p(\Omega, \mathscr{F}, P)$ for $p \in [1, \infty)$. Moreover, the following lemma holds.

Lemma 3.3. *D is a closable operator on $L^p(\Omega, \mathscr{F}, P)$ for $p \in [1, \infty)$.*

By Lemmas 3.1 and 3.3, D can be extended to a closed (unbounded) linear operator on $L^p(\Omega, \mathscr{F}, P)$ for $p \in [1, \infty)$. We denote the extension by D again.

For $F \in \mathcal{P}$ and $p \in [1, \infty)$, define $\|F\|_{1,p}$ by

$$\|F\|_{1,p} := \left(E[|F|^p] + \|DF\|_H^p\right)^{1/p}.$$

It is easy to see that $\|\cdot\|_{1,p}$ satisfies the properties of norms, and we can consider the closure of $\mathcal{P}$ with respect to $\|\cdot\|_{1,p}$. We denote the closure by $\mathbb{D}^{1,p}$. Note that $\|\cdot\|_{1,p}$ is the operator norm of D in $L^p(\Omega, \mathscr{F}, P)$. The function space $\mathbb{D}^{1,p}$ is the Sobolev space associated with the H-derivative with index $1, p$. Similarly, we can define the Sobolev space $\mathbb{D}^{k,p}$ of higher orders.

Proposition 3.4. *Let $F \in \mathbb{D}^{1,2}$ such that $F = \sum_{n=0}^{\infty} I_n(f_n)$ where $f_n \in H^{\odot n}$. Then,*

$$(DF, h)_H = \sum_{n=1}^{\infty} n \int_T I_{n-1}(f_n(\cdot, t)) h(t) \mu(dt)$$

where $f_n(\cdot, t)$ is the function on T^{n-1} given by

$$[f_n(\cdot, t)](s_1, s, 2, \dots, s_n - 1) := f_n(s_1, s_2, \dots, s_{n-1}, t), \quad s_1, s_2, \dots, s_n \in T.$$

Hence,

$$E[\|DF\|_H^2] = \sum_{n=1}^{\infty} n n! \|f_n\|_{H^{\otimes n}}^2 = \sum_{n=1}^{\infty} n \|f_n\|_{H^{\odot n}}^2.$$

Now we define the operator δ as the follows.

Definition 3.5. *Let δ be the dual operator of $D : L^2(\Omega, \mathscr{F}, P) \to L^2(\Omega, \mathscr{F}, P; H)$.*

The operator δ is called the Skorohod integral. We remark that the Skorohod integral can be regarded as an extension of the stochastic integral (Itô integral).

Let L be the Ornstein-Uhlembeck operator on $L^2(\Omega, \mathscr{F}, P)$ associated with W. There are some ways to define the Ornstein-Uhlembeck operator. For example, in [9] the Ornstein-Uhlembeck operator is defined by using Wiener chaos expansion. On the other hand, in [13], first we define the Ornstein-Uhlembeck semigroup by using the explicit transition semigroup, and the Ornstein-Uhlembeck operator is defined by the generator of the Ornstein-Uhlembeck semigroup. We omit the precise definition of L here, and only remark that the domain of L includes $\mathcal{P}$ and L is characterized by

$$LF = \sum_{i,j=1}^{n} \partial_i\partial_j f(W(h_1), W(h_2), \dots, W(h_n))(h_i, h_j)_H - \sum_{i=1}^{n} \partial_i f(W(h_1), W(h_2), \dots, W(h_n))W(h_i) \tag{3.2}$$

where F is the random variable expressed as (3.1).

The following propositions hold.

Proposition 3.6. *$\delta D = -L$.*

Proposition 3.7. *For $f \in H^{\otimes n}$, $LI_n(f) = -nI_n(f)$.*

We use these facts in the proofs of the fourth moment theorem and the Stein's bound.

4 The fourth moment theorem

In this section, we give the version of the fourth moment theorem given by Nualart and Ortiz-Latorre [10].

First we define the contraction of functions. For $f \in L^2(T^p, \mathscr{B}^{\otimes p}, \mu^{\otimes p})$, $g \in L^2(T^q, \mathscr{B}^{\otimes q}, \mu^{\otimes q})$ and $r = 1, 2, \dots, \min\{p, q\}$, we define $f \otimes g \in L^2(T^{p+q}, \mathscr{B}^{\otimes p+q}, \mu^{\otimes p+q})$ and $f \otimes_r g \in L^2(T^{p+q-2r}, \mathscr{B}^{\otimes p+q-2r}, \mu^{\otimes p+q-2r})$ by

$$\begin{aligned}
&(f \otimes g)(t_1, t_2, \dots, t_{p+q}) \\
&= f(t_1, t_2, \dots, t_p)g(t_{p+1}, t_{p+2}, \dots, t_{p+q}), \\
&(f \otimes_r g)(t_1, t_2, \dots, t_{p+q-2r}) \\
&= \int_{T^r} f(t_1, t_2, \dots, t_{p-r}, s_1, s_2, \dots, s_r)g(t_{p-r+1}, t_{p-r+2}, \dots, t_{p+q-2r}, s_1, s_2, \dots, s_r) \\
&\qquad\qquad \times \mu^{\otimes r}(ds_1, ds_2, \dots, ds_r),
\end{aligned}$$

respectively. We call the operation $(f, g) \mapsto f \otimes_r g$ is called the contraction of f and g of order r. Since $f \otimes g$ can be regarded as $f \otimes_r g$ with $r = 0$, we define $f \otimes_0 g$ by $f \otimes g$.

The tensor product $f \otimes g$ and the contractions $f \otimes_r g$ are not always symmetric even if f and g are symmetric. We denote the symmetrizations of $f \otimes g$ and $f \otimes_r g$ by $f\widetilde{\otimes}g$ and $f\widetilde{\otimes}_r g$, respectively.

By using contraction we can calculate the product of two random variables in some levels of Wiener chaos as follows.

Proposition 4.1. *Let $f \in L^2(\mu^{\otimes p})$ be symmetric and $g \in L^2(\mu)$. Then,*

$$I_p(f)I_1(g) = I_{p+1}(f \otimes g) + pI_{p-1}(f \otimes_1 g). \tag{4.1}$$

The proposition 4.1 is extended as follows.

Proposition 4.2. *Let $f \in L^2(\mu^{\otimes p})$ and $g \in L^2(\mu^{\otimes q})$ are symmetric. Then,*

$$I_p(f)I_q(g) = \sum_{r=0}^{p\wedge q} r! \begin{pmatrix} p \\ r \end{pmatrix} \begin{pmatrix} q \\ r \end{pmatrix} I_{p+q-2r}(f \otimes_r g). \tag{4.2}$$

Note that Proposition 4.2 gives the explicit information of the Wiener chaos expansion of the product of two random variables in some levels of Wiener chaos. The proposition is crucial to prove the fourth moment theorem. In the proof of Proposition 4.2, complicated calculation in combination is needed.

By using Proposition 4.2, we have the fourth moment theorem which is the version given by Nualart and Ortiz-Latorre [10] as follows.

Theorem 4.3. *(The fourth moment theorem) Consider a sequence $\{F_k = I_n(f_k)\}$ of square integrable random variables in the n-th Wiener chaos. Assume that*

$$\lim_{k\to\infty} E[F_k^2] = \lim_{k\to\infty} \|f_k\|^2_{H^{\odot n}} = 1. \tag{4.3}$$

Then, the following statements are equivalent.

(i) $\{F_k = I_n(f_k)\}$ *converges to the standard normal law in distribution as* $k \to \infty$.

(ii) $\lim_{k\to\infty} E[F_k^4] = 3$.

(iii) $\lim_{k\to\infty} \|f_k \otimes_l f_k\|_{H^{\otimes 2(n-l)}} = 0$ *for* $l = 1, 2, \dots, n-1$.

(iv) $\|DF_k\|_H^2$ *converges to* n *in* $\mathbb{L}^2$ *as* $k \to \infty$.

Multidimensional case of the fourth moment theorem is considered in [12], and [10]. In [6] the fourth moment theorem with respect to the centered Gamma distribution is also obtained.

5 Stein's method and Application of Malliavin calculus

Charles Stein considered in order to estimate the reminder term of the central limit theorem (see [14]). He prepared the ordinary differential equation associated with the standard normal distribution satisfies, and obtained a bound of the reminder term by using the solution to the equation. The equation is called Stein's equation, and the method to obtain the bound is called Stein's method. The large deviation principle is also well-known as a method to obtain the convergence rate of the central limit theorem (or the law of large numbers.) The large deviation principle has advantages in analysis to Stein's method, because the large deviation principle is related to the spectral analysis. On the other hand, Stein's method has advantages in computation and in practice, because the bound of the reminder term is obtained by explicit calculations. By using Stein's method, one can estimate the distances between the standard normal distribution and other distributions, where the distances mean, for example, Kolmogorov distance, Wasserstein distance, and total variation distance.

First we give the detail of Stein's equation and Stein's bound. Let Z be a random variable with the standard normal distribution and h be a measurable function on $\mathbb{R}$ such that $E[|h(Z)|] < \infty$. Stein's equation associated with h and Z is

$$h(x) - E[h(Z)] = f'(x) - xf(x), \quad x \in \mathbb{R}. \tag{5.1}$$

The solution f to (5.1) is obtained explicitly as follows:

$$f(x) = e^{\frac{1}{2}x^2} \int_{-\infty}^{x} (h(y) - E[h(Z)])\, e^{-\frac{1}{2}y^2} dy, \quad x \in \mathbb{R}. \tag{5.2}$$

By using (5.1) and (5.2), the following proposition holds.

Proposition 5.1. (i) *Let X be a random variable. Then, X has the standard normal distribution if and only if,*

$$E[f'(X) - Xf(X)] = 0$$

for every continuous and piecewise differentiable function f satisfying $E[|f'(Z)|] < \infty$.

(ii) *If* $h(x) = \mathbb{I}_{(-\infty,z]}(x)$ *where* $z \in \mathbb{R}$*, then the solution* f *to (5.1) exists,* f *is piecewise continuously differentiable,* $\|f\|_\infty \le \sqrt{2\pi}/4$*, and* $\|f'\|_\infty \le 1$.

(iii) *If* h *is bounded by* $1/2$*, the solution* f *to (5.1) exists,* f *is differentiable almost everywhere,* $\|f\|_\infty \le \sqrt{\pi/2}$*, and* $\|f'\|_\infty \le 2$.

(iv) *If* h *is bounded and absolutely continuous, then the solution* f *to (5.1) exists,* f *is bounded and twice differentiable,* $\|f\|_\infty \le \sqrt{\pi/2}\|h(y) - E[h(Z)]\|_\infty$*,* $\|f'\| \le 2\|h(y) - E[h(Z)]\|_\infty$*, and* $\|f''\|_\infty \le 2\|h'\|_\infty$.

(v) *If* h *is absolutely continuous and the derivative is bounded, then the solution* f *to (5.1) exists,* f *is twice differentiable,* $\|f'\|_\infty \le \|h'\|_\infty$*, and* $\|f''\|_\infty \le 2\|h'\|_\infty$.

The proofs of the statements in Proposition 5.1 are found in the references in [6].

By using (5.1) and the bound in Proposition 5.1, we can consider the estimate for the distances between the standard normal distribution and other distributions.

Consider a distance between distributions of random variables F and G on $\mathbb{R}$ defined by

$$d_{\mathcal{H}}(\mathcal{L}(F), \mathcal{L}(G)) := \sup_{f\in\mathcal{H}} |\mathbf{E}[f(F)] - \mathbf{E}[f(G)]|, \tag{5.3}$$

where $\mathcal{H}$ is a set of functions on $\mathbb{R}$. There are many distances between distributions defined by (5.3). For example, by taking $\mathcal{H} = \mathcal{F}_{\mathrm{Kol}} := \{\mathbb{I}_{(-\infty,z]} : z \in \mathbb{R}\}$, one obtains the Kolmogorov distance; by taking $\mathcal{H} = \mathcal{F}_{\mathrm{W}} := \{f : ||f||_L \le 1\}$, where $|| \cdot ||_L$ denotes the usual Lipschitz seminorm, one obtains the Wasserstein (or Kantorovich-Wasserstein) distance; by taking $\mathcal{H} = \mathcal{F}_{\mathrm{FM}} := \{f : ||f||_{BL} \le 1\}$, where $|| \cdot ||_{BL} = || \cdot ||_L + || \cdot ||_\infty$, one obtains the Fortet-Mourier (or bounded Wasserstein) distance; by letting $\mathcal{H} = \mathcal{F}_{\mathrm{TV}}$ be the collection of all indicators $\mathbb{I}_B$ of Borel sets, one obtains the total variation distance. We denote the Kolmogorov distance, the Wasserstein distance, Fortet-Mourier distance and the total variation distance by $d_{\mathrm{Kol}}(\cdot,\cdot)$, $d_{\mathrm{W}}(\cdot,\cdot)$, $d_{\mathrm{FM}}(\cdot,\cdot)$ and $d_{\mathrm{TV}}(\cdot,\cdot)$, respectively.

The following theorem is the result on the estimate of the distances, which is obtained by Nourdin and Peccati [6].

Theorem 5.2. *(Theorem 3.1 of [6]) Let* Z *has the standard normal distribution, and* $F \in \mathbb{D}^{1,2}$ *such that* $E[F] = 0$*. Then,*

$$d_{\mathrm{W}}(F, Z) \le E\left[(1 - (DF, -DL^{-1}F)_H)^2\right]^{1/2},$$
$$d_{\mathrm{FM}}(F, Z) \le E\left[(1 - (DF, -DL^{-1}F)_H)^2\right]^{1/2}.$$

If, in addition, the law of F *is absolutely continuous with respect to Lebesgue measure,*

$$d_{\mathrm{Kol}}(F, Z) \le E\left[(1 - (DF, -DL^{-1}F)_H)^2\right]^{1/2},$$
$$d_{\mathrm{TV}}(F, Z) \le 2E\left[(1 - (DF, -DL^{-1}F)_H)^2\right]^{1/2}.$$

On the other hand, by using Proposition 4.2, we have the following Proposition.

Proposition 5.3. *(Proposition 3.2 of [6]) Let* $n = 2, 3, 4, \dots,$ *and* $F = I_n(f)$ *where* $f \in H^{\odot n}$*. Then,* $(DF, -DL^{-1}F)_H = n^{-1}\|DF\|_H^2$*, and*

$$E\left[(1 - (DF, -DL^{-1}F)_H)^2\right]$$
$$= E\left[(1 - n^{-1}\|DF\|_H^2)^2\right] \tag{5.4}$$
$$= (1 - n!\|f\|^2_{H^{\otimes n}})^2 + n^2 \sum_{r=1}^{n-1} (2n-2r)!\,[(r-1)!]^2 \begin{pmatrix} n-1 \\ r-1 \end{pmatrix}^4 \|f \widetilde{\otimes}_r f\|_{H^{\otimes 2(n-r)}} \tag{5.5}$$
$$\le (1 - n!\|f\|^2_{H^{\otimes n}})^2 + n^2 \sum_{r=1}^{n-1} (2n-2r)!\,[(r-1)!]^2 \begin{pmatrix} n-1 \\ r-1 \end{pmatrix}^4 \|f \otimes_r f\|_{H^{\otimes 2(n-r)}}. \tag{5.6}$$

By using Theorem 5.2 and Proposition 5.3 we can simplify some parts of the proof of the fourth moment theorem (Theorem 4.3).

The case of the centered Gamma distribution is also considered in [6]. They prepared Stein's equation associated with the centered Gamma distribution, and obtained the bound of convergence to the centered Gamma distribution by similar way to the case of the standard normal distribution.

6 Further works on the fourth moment theorem and related topics

In this section we introduce some further studies around the fourth moment theorem.

6.1 The case of the centered Gamma distribution

An analogue of the fourth moment theorem to the centered Gamma distribution is obtained by Nourdin and Peccati [5]. The statement is as follows.

Let $\nu > 0$ and $G(\nu/2)$ be a random variable having the Gamma distribution with parameter $\nu/2$, i.e. $G(\nu/2)$ is a random variable with density function $g(x) = \frac{x^{\nu/2-1}e^{-x}}{\Gamma(\nu/2)}\mathbb{I}_{(0,\infty)}$, where Γ is the Gamma function. Consider a random variable $F(\nu)$ defined by

$$F(\nu) := 2G(\nu/2) - \nu.$$

The following theorem is an analogue of the fourth moment theorem with respect to $F(\nu)$.

Theorem 6.1. *Let $n \in 2\mathbb{N}$ and*

$$c_n := \frac{4}{(n/2)!\begin{pmatrix} n \\ n/2 \end{pmatrix}^2}.$$

Consider a sequence of random variables $G_k = I_n(g_k)$ where $g_k \in H^{\odot n}$ and assume that

$$\lim_{k\to\infty} E[G_k^2] = \lim_{k\to\infty} n!\|g_k\|_{H^{\otimes n}}^2 = 2\nu.$$

Then, the following conditions are equivalent.

(i) *G_k converges to $F(\nu)$ in distribution as $k \to \infty$.*

(ii) *$\lim_{k\to\infty}(E[G_k^4] - 12E[G_k^3]) = 12\nu^2 - 48\nu$.*

(iii) *$\|DG_k\|_H^2 - 2nG_k$ converges to $2n\nu$ in $L^2(P)$ as $k \to \infty$.*

(iv) *$\lim_{k\to\infty}\|g_k\tilde{\otimes}_{n/2}g_k - c_n g_k\|_{H^{\otimes n}} = 0$ and $\lim_{k\to\infty}\|g_k \otimes_r g_k\|_{H^{\otimes 2(n-r)}} = 0$ for $r = 1, 2, \ldots, n-1$ except $r = n/2$.*

In [6], they discuss the case of the centered Gamma distribution in a similar way to the case of the standard normal distribution, and Stein's equation with respect to the centered Gamma distribution is obtained. We omit the version of Stein's equation in this note, because more general version of Stein's equation is in Section 6.2.

6.2 Generalization of Stein's bound

As we have seen in Theorem 5.2, by applying Malliavin calculus to Stein's equation we obtain the estimate of the distances between distributions. The cases of the standard normal distribution and the centered Gamma distribution are considered in [6], and more general argument is also mentioned as a conjecture in [6]. After Nourdin and Peccati [6], in [4] a general argument is constructed in view of the invariant measure of one-dimensional stochastic differential equation.

Let S be the interval (l,u) $(-\infty \le l < u \le \infty)$ and μ be a probability measure on S with a density function p which is continuous, bounded, strictly positive on S, and admits finite variance. Consider a continuous function b on S such that there exists $k \in (l,u)$ such that $b(x) > 0$ for $x \in (l,k)$ and $b(x) < 0$ for $x \in (k,u)$, bp is bounded on S and

$$\int_l^u b(x)p(x)dx = 0.$$

Define

$$a(x) := \frac{2\int_l^x b(y)p(y)dy}{p(x)}, \quad x \in S.$$

Then, the stochastic differential equation:

$$dX_t = b(X_t)dt + \sqrt{a(X_t)}dW_t, \qquad t \ge 0$$

has a unique Markovian weak solution, ergodic with invariant density p (see [1]).

For $f \in C_0(S)$ (the set of continuous functions on S vanishing at the boundary of S), let $m_f := \int_l^u f(x)p(x)dx$ and define $\tilde{g}_f$ by, for every $x \in S$,

$$\tilde{g}_f(x) := \frac{2}{a(x)p(x)}\int_l^x (f(y) - m_f)p(y)dy$$

Then, we have

$$\tilde{g}_f(x) = \int_l^x \frac{2(f(y)-m_f)}{a(y)} \exp\left(-\int_y^x \frac{2b(z)}{a(z)}dz\right)dy, \qquad x \in S.$$

Then, $g_f(x) := \int_0^x \tilde{g}_f(y)dy$ satisfies that $f - m_f = Ag_f$ and

$$f(x) - \mathbf{E}[f(X)] = \frac{1}{2}a(x)\tilde{g}_f'(x) + b(x)\tilde{g}_f(x) \tag{6.1}$$

where X is a random variable with its law μ. The equation (6.1) is a generalized version of Stein's equation.

To obtain the estimate of the distances between the distribution associated with $p(x)dx$ and other distributions, we need the bounds for the functions $\tilde{g}_f$ and $\tilde{g}_f'$. Since we have the explicit form of $\tilde{g}_f$, the following propositions are obtained.

Proposition 6.2. *Assume that a is uniformly positive and there exist $l', u' \in (l,u)$ such that b is non-increasing on (l,l') and (u',u). Then we have*

$$\|\tilde{g}_f\|_\infty \le C_1\|f\|_\infty \text{ and } \|a\tilde{g}_f'\|_\infty \le C_2\|f\|_\infty, \quad f \in C_0^\infty(S),$$

where C_1 and C_2 are strictly positive constants.

Proposition 6.3. *Assume that if $u < \infty$, there exists $u' \in (l,u)$ such that b is non-decreasing and Lipschitz continuous on $[u',u)$ and $\liminf_{x\to u} a(x)/(u-x) > 0$; if $u = \infty$, there exists $u' \in (l,u)$ such that b is non-decreasing on $[u',u)$ and $\liminf_{x\to u} a(x) > 0$. Similarly, assume that if $l > -\infty$, there exists $l' \in (l,u)$ such that b is non-increasing and Lipschitz continuous on $(l,l']$ and $\liminf_{x\to l} a(x)/(x-l) > 0$; if $l = -\infty$, there exists $l' \in (l,u)$ such that b is non-decreasing on $(l,l']$ and $\liminf_{x\to l} a(x) > 0$. Then we have*

$$\|\tilde{g}_f'\|_\infty \le C_4(\|f\|_\infty + \|f'\|_\infty), \quad f \in C_0^\infty(S),$$

where C_4 is a constant.

The estimates in Proposition 6.2 are sufficiently good when a is uniformly bounded and strictly positive. However, when a degenerates at the boundary of S, we need Proposition 6.3. We remark that in Proposition 6.3 the derivative of $\tilde{g}_f$ is dominated by the sum of $\|f\|_\infty$ and $\|f'\|_\infty$. In view of this fact it seems true that the case that a is uniformly positive and the case that a is degenerate are very different. In fact, the result obtained in the case of the standard normal distribution is different from that obtained in the case of the centered Gamma distribution (see Sections 5 and 6.1).

By using the generalized Stein's equation (6.1) and Propositions 6.2 and 6.3, the estimate of the distances between the distribution associated with $\mu = p(x)dx$ and other distributions is obtained in the same way as 5.

Theorem 6.4. (i) *Let d be the Fortet-Mourier distance. Assume the conditions in Proposition 6.3. Then,*

$$d(\mathcal{L}(Y), \mu) \leq CE\left[\left|\frac{1}{2}a(Y) + \langle D(-L)^{-1}\left\{b(Y) - \mathbf{E}[b(Y)]\right\}, DY\rangle_H\right|\right] \\ + C|\mathbf{E}\left[b(Y)\right]|, \quad Y \in \mathbb{D}^{1,2}$$

where C is a positive constant and $\mathcal{L}(Y)$ is the law of Y.

(ii) *Let d be the Fortet-Mourier distance, the Kolmogorov distance or the total variation distance. Assume the conditions in Proposition 6.2 and a is uniformly positive. Then,*

$$d(\mathcal{L}(Y), \mu) \leq CE\left[\left|\frac{1}{2}a(Y) + \langle D(-L)^{-1}\left\{b(Y) - \mathbf{E}[b(Y)]\right\}, DY\rangle_H\right|\right] \\ + C|\mathbf{E}\left[b(Y)\right]|, \quad Y \in \mathbb{D}^{1,2}$$

where C is a positive constant.

The bounds in Theorem 6.4 are optimal in the following sense.

Theorem 6.5. *A random variable $Y \in \mathbb{D}^{1,2}$ with its values on S has probability distribution μ and satisfies that $b(Y) \in L^2(\Omega)$ if and only if $E[b(Y)] = 0$ and*

$$E\left[\frac{1}{2}a(Y) + \langle D(-L)^{-1}b(Y), DY\rangle_H \middle| Y\right] = 0.$$

In Theorem 6.4, there is the term:

$$E\left[\left|\frac{1}{2}a(Y) + \langle D(-L)^{-1}\left\{b(Y) - \mathbf{E}[b(Y)]\right\}, DY\rangle_H\right|\right]. \tag{6.2}$$

Generally it is difficult to calculate this term. However, if a and b are given explicitly, and if Y is expressed as an explicit function of a Gaussian random variables, then (6.2) can be calculated by using the Ornstein-Uhlembeck semigroup and its resolvent (see [4] and [8].)

6.3 Other works

There are many other works on this topic.

The analogue in free probability theory has been concerned in [3]. In free probability theory, we also have the analogue of Wiener chaos (so-called Wigner chaos or free chaos). In [3], the analogues of the fourth moment theorem and Stein's bound with respect to the semicircular law are obtained.

In [2], the original Stein's method is applied to the theory of spin glasses. In the paper, the Thouless-Anderson-Palmer equations of the Sherrington-Kirkpatrick model is obtained. Moreover, the upper estimate of the convergence to the Thouless-Anderson-Palmer equations is also obtained by using the original Stein's method. Like this, the applications of Stein's method to statistical mechanics are also considered and some results have been obtained recently. We remark that the argument in the paper is away from the fourth moment theorem and the combination of Stein's method and the Malliavin calculus.

References

[1] B. M. Bibby, I. M. Skovgaard, and M. Sørensen. Diffusion-type models with given marginal distribution and autocorrelation function. *Bernoulli*, 11(2):191–220, 2005.

[2] S. Chatterjee. Spin glasses and Stein's method. *Probab. Theory Related Fields*, 148(3-4):567–600, 2010.

[3] Nourdin I. Peccati G. Kemp, T. and R. Speicher. Wigner chaos and the fourth moment. *Ann. Probab.*, 40(4):1577–1635, 2012.

[4] S. Kusuoka and C. A. Tudor. Stein's method for invariant measures of diffusions via Malliavin calculus. *Stochastic Processes and their Applications*, 122:1627–1651, 2012.

[5] I. Nourdin and G. Peccati. Noncentral convergence of multiple integrals. *Ann. Probab.*, 37(4):1412–1426, 2009.

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

[7] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus.* Cambridge University Press, 2012.

[8] I. Nourdin and F. G. Viens. Density formula and concentration inequalities with Malliavin calculus. *Electron. J. Probab.*, 14:no. 78, 2287–2309, 2009.

[9] D. Nualart. *The Malliavin calculus and related topics.* Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[10] D. Nualart and S. Ortiz-Latorre. Central limit theorems for multiple stochastic integrals and Malliavin calculus. *Stochastic Process. Appl.*, 118(4):614–628, 2008.

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

[12] G. Peccati and C. A. Tudor. Gaussian limits for vector-valued multiple stochastic integrals. In *Séminaire de Probabilités XXXVIII*, volume 1857 of *Lecture Notes in Math.*, pages 247–262. Springer, Berlin, 2005.

[13] I. Shigekawa. *Stochastic analysis*, volume 224 of *Translations of Mathematical Monographs.* American Mathematical Society, Providence, RI, 2004. Translated from the 1998 Japanese original by the author, Iwanami Series in Modern Mathematics.

[14] C. Stein. *Approximate computation of expectations.* Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7. Institute of Mathematical Statistics, Hayward, CA, 1986.