Bernoulli **24**(2), 2018, 1463–1496
DOI: 10.3150/16-BEJ904

Characterization of the convergence in total variation and extension of the Fourth Moment Theorem to invariant measures of diffusions

SEIICHIRO KUSUOKA[1] and CIPRIAN A. TUDOR[2,3]

[1]*Research Institute for Interdisciplinary Science, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530, Japan. E-mail: kusuoka@okayama-u.ac.jp*
[2]*Laboratoire Paul Painlevé, Université de Lille 1, F-59655 Villeneuve d'Ascq, France*
[3]*Department of Mathematics, Academy of Economical Studies, Bucharest, Romania.*
E-mail: tudor@math.univ-lille1.fr

We give necessary and sufficient conditions to characterize the convergence in distribution of a sequence of arbitrary random variables to a probability distribution which is the invariant measure of a diffusion process. This class of target distributions includes the most known continuous probability distributions. Precisely speaking, we characterize the convergence in total variation to target distributions which are not Gaussian or Gamma distributed, in terms of the Malliavin calculus and of the coefficients of the associated diffusion process. We also prove that, among the distributions whose associated squared diffusion coefficient is a polynomial of second degree (with some restrictions on its coefficients), the only possible limits of sequences of multiple integrals are the Gaussian and the Gamma laws.

Keywords: convergence in total variation; diffusions; Fourth Moment Theorem; invariant measure; Malliavin calculus; multiple stochastic integrals; Stein's method; weak convergence

1. Introduction

In the seminal paper [14], Nualart and Peccati discovered a surprising central limit theorem (called *The Fourth Moment Theorem*) for sequences of multiple stochastic integrals in a Wiener chaos of fixed order. This result says that the convergence in distribution of such a sequence of random variables to the standard normal law is actually equivalent to the convergence of only the sequence of their fourth moments. A multidimensional version of this result has been given in [15]. Since the publication of these two pathbreaking papers, many improvements and developments on this theme have been considered. Among them is the work [13], giving a new proof only based on Malliavin calculus and the use of integration by parts on the Wiener space.

Another pathbreaking paper is the work [6] by Nourdin and Peccati in which the authors bring together Stein's method with Malliavin calculus and obtain useful estimates for the distance between the law of an arbitrary random variable and the Gaussian distribution in terms of the Malliavin calculus. It turns out that Stein's method and Malliavin calculus fit together admirably well, and that their interaction has led to some remarkable new results involving central and non-central limit theorems for functionals of infinite-dimensional Gaussian fields. We refer

1350-7265 © 2018 ISI/BS

to the recent monographs [8] for an overview of the existing literature and to [16] for various applications of the Stein's method and Malliavin calculus to limit theorems and statistics.

There is also a version of the Fourth Moment Theorem having the Gamma distribution as the target distribution (see [7]). The convergence of a sequence of multiple stochastic integrals toward a Gamma distribution is characterized by the convergence of the sequences of the third and fourth moments; alternatively, one can also characterize the convergence to the Gamma law in terms of the Malliavin derivatives.

In the paper [5], we obtained bounds between the distance of an arbitrary random variable and target distributions which are invariant measures of diffusions processes. This class contains the most common continuous probability distributions, including the Gaussian, Gamma, Beta, Pareto, uniform, Student or log-normal distributions, among others. See also [1,4] for other attempts to extend the theory to more general target distribution via Malliavin calculus. Now, our purpose is to give necessary and sufficient conditions for the convergence of a sequence of random variables (regular enough in the Malliavin sense) to such target distributions. We obtain several results based on the Malliavin derivatives of the sequence and of the diffusion coefficients associated to the target distribution. Precisely speaking, we prove the equivalence between the convergence in total variation of random variables under a suitable condition and the convergence of a quantity consisting of the diffusion coefficients and the Malliavin derivatives. The value is associated to the so-called Stein factor, and the result also characterizes the convergence of the Stein factor. Several situations when this theory can be applied (meaning that the diffusion coefficients have an explicit expression) are presented. Then we treat the case of the convergence in law of special sequences of random variables that belong to a Wiener chaos of fixed order and we characterize their convergence to target distributions that are different from the Gaussian and Gamma laws. For example, we obtain necessary and sufficient conditions in the case when the limit is a product normal distribution (the product of two independent normal random variables) or the sum of a normal and an independent random variable.

We will also focus our analysis on the particular case when the squared diffusion coefficient is a polynomial of (at most) second degree. Several common probability laws are contained in this class. In this situation, the necessary and sufficient conditions for the convergence of a sequence of multiple integrals to the target distribution can be analyzed in details. We actually show that, among the distributions whose associated squared diffusion coefficient is a polynomial of second degree with some restrictions on its coefficients, the only possible limits of sequences of multiple integrals are the Gaussian and the Gamma laws. In particular, we show that a sequence of multiple Wiener–Itô integrals cannot converge toward a beta or uniform distribution. We retrieve the standard Fourth Moment Theorem (Theorem 2.2) and its version for the Gamma law as particular cases.

We organized our paper as follows. Section 2 contains some preliminaries on the Malliavin calculus. In Section 3, we recall and extend several results in [5] concerning the characterization of the random variables whose probability distribution is the invariant measure of a diffusion process. In Section 4, we give necessary and sufficient conditions for the convergence of a sequence of random variables (regular enough in the Malliavin sense), while Section 5 and 6 treat the convergence in distribution of sequence of multiple stochastic integrals.

2. Preliminary: Wiener-chaos and Malliavin derivatives

Here we describe the elements from stochastic analysis that we will need in the paper. Consider H a real separable Hilbert space and $(W(h), h \in H)$ an isonormal Gaussian process on a probability space $(\Omega, \mathcal{A}, P)$, which is a centered Gaussian family of random variables such that $\mathbf{E}[W(\varphi)W(\psi)] = \langle \varphi, \psi \rangle_H$. Denote by I_n the multiple stochastic integral with respect to B (see [12]). This mapping I_n is actually an isometry between the Hilbert space $H^{\odot n}$(symmetric tensor product) equipped with the scaled norm $\frac{1}{\sqrt{n!}} \| \cdot \|_{H^{\otimes n}}$ and the Wiener chaos of order n which is defined as the closed linear span of the random variables $H_n(W(h))$ where $h \in H$, $\|h\|_H = 1$ and H_n is the Hermite polynomial of degree $n \in \mathbb{N}$

$$H_n(x) = \frac{(-1)^n}{n!} \exp\left(\frac{x^2}{2}\right) \frac{d^n}{dx^n} \left(\exp\left(-\frac{x^2}{2}\right)\right), \qquad x \in \mathbb{R}.$$

The isometry of multiple integrals can be written as follows: for m, n positive integers,

$$\begin{aligned} &\mathbf{E}\big(I_n(f) I_m(g)\big) = n! \langle \tilde{f}, \tilde{g} \rangle_{H^{\otimes n}} \qquad \text{if } m = n, \\ &\mathbf{E}\big(I_n(f) I_m(g)\big) = 0 \qquad \text{if } m \neq n. \end{aligned} \tag{2.1}$$

It also holds that

$$I_n(f) = I_n(\tilde{f}),$$

where $\tilde{f}$ denotes the symmetrization of f defined by the formula $\tilde{f}(x_1, \dots, x_n) = \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$.

We recall that any square integrable random variable which is measurable with respect to the σ-algebra generated by W can be expanded into an orthogonal sum of multiple stochastic integrals

$$F = \sum_{n=0}^{\infty} I_n(f_n), \tag{2.2}$$

where $f_n \in H^{\odot n}$ are (uniquely determined) symmetric functions and $I_0(f_0) = \mathbf{E}[F]$.

Let L be the Ornstein–Uhlenbeck operator

$$LF = -\sum_{n \geq 0} n I_n(f_n)$$

if F is given by (2.2) and it is such that $\sum_{n=1}^{\infty} n^2 n! \| f_n \|^2_{\mathcal{H}^{\otimes n}} < \infty$.

For $p > 1$ and $\alpha \in \mathbb{R}$, we introduce the Sobolev–Watanabe space $\mathbb{D}^{\alpha, p}$ as the closure of the set of polynomial random variables with respect to the norm

$$\|F\|_{\alpha, p} = \left\| (I - L)^{\frac{\alpha}{2}} F \right\|_{L^p(\Omega)},$$

where I represents the identity. We denote by D the Malliavin derivative operator that acts on smooth functions of the form $F = g(W(h_1), \dots, W(h_n))$ (g is a smooth function with compact

support and $h_i \in H$)

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

The operator D is continuous from $\mathbb{D}^{\alpha,p}$ into $\mathbb{D}^{\alpha-1,p}(H)$.

We will intensively use the product formula for multiple integrals. It is well known that for $f \in H^{\odot n}$ and $g \in H^{\odot m}$

$$I_n(f)I_m(g) = \sum_{r=0}^{n\wedge m} r!\binom{n}{r}\binom{m}{r} I_{m+n-2r}(f \otimes_r g), \tag{2.3}$$

where $f \otimes_r g$ means the r-contraction of f and g (see, e.g., Section 1.1.2 in [12]).

We recall the expression of the third and fourth moment of a random variable in a fixed Wiener chaos. These formulas play an important role in our proofs.

Lemma 2.1. *Let $F = I_n(f)$ with $n \in \mathbb{N}$ and $f \in H^{\odot n}$. Then*

$$\mathbf{E}\big[F^3\big] = \frac{(n!)^3}{[(n/2)!]^3}\langle f, f \tilde{\otimes}_{n/2} f\rangle 1_{\{n \text{ is even}\}} \tag{2.4}$$

and

$$\begin{aligned}\mathbf{E}\big[F^4\big] &= 3\mathbf{E}\big[F^2\big]^2 + 3n\sum_{p=1}^{n-1}(p-1)!\binom{n-1}{p-1}^2 p!\binom{n}{p}^2 (2n-2p)!\|f_m \tilde{\otimes}_p f_m\|^2 \\ &= 3\mathbf{E}\big[I_n(f)^2\big]^2 + n!^2\sum_{r=1}^{n-1}\big(C_n^r\big)^2\big[\|f \otimes_r f\|^2 + C_{2n-2r}^{n-r}\|f \tilde{\otimes}_r f\|^2\big].\end{aligned} \tag{2.5}$$

Proof. We refer to [7], proof of Theorem 1.2 for the first two relations and to [8], formula (5.2.6) for the last equality. □

The Fourth Moment Theorem states as follows.

Theorem 2.2 ([14] and [13]). *Fix $n \in \mathbb{N}$. Consider a sequence $(F_k = I_n(f_k))_{k\in\mathbb{N}}$ of square integrable random variables in the nth Wiener chaos. Assume that*

$$\lim_{k\to\infty} \mathbf{E}\big[F_k^2\big] = \lim_{k\to\infty} \|f_k\|^2_{H^{\odot n}} = 1. \tag{2.6}$$

Then, the following statements are equivalent.

(i) *The sequence of random variables $(F_k = I_n(f_k))_{k\geq 1}$ converges to the standard normal law in distribution as $k \to \infty$.*

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

(iii) $\lim_{k\to\infty} \|f_k \otimes_l f_k\|_{H^{\otimes 2(n-l)}} = 0$ *for* $l = 1, 2, \ldots, n-1$.
(iv) $\|DF_k\|_H^2$ *converges to* n *in* $L^2(\Omega)$ *as* $k \to \infty$.

We consider the general version of this theorem below.

3. General versions of Stein's method and Stein's bound

In order to discuss the general version of the Fourth Moment Theorem, we review Stein's method and Stein's bound obtained in [5] with small extension. Let us briefly recall the context in [5]. 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, 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 $x \in (k, u)$, $b \in L^1(\mu)$, bp is bounded on S and

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

Define

$$a(x) := \frac{2\int_l^x b(y)p(y)\,dy}{p(x)}, \qquad x \in S. \tag{3.1}$$

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 Theorem 2.4 in [2].

Based on this fact, it is possible to define a so-called Stein's equation for a given function $f \in L^1(\mu)$. In Section 3 of [5], we have considered only the case where $f \in C_0(S)$ (the set of continuous functions on S vanishing at the boundary of S). However, it is easy to see that the argument is valid even if $f \in L^1(\mu)$, as shown below.

For $f \in L^1(\mu)$, let $m_f := \int_l^u f(x)\mu(dx)$ and define $\tilde{g}_f$ by, for every $x \in S$,

$$\tilde{g}_f(x) := \frac{2}{a(x)p(x)} \int_l^x \big(f(y) - m_f\big)p(y)\,dy. \tag{3.2}$$

Then, by Proposition 1 in Section 3.2 of [5], we have

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

The function $g_f(x) := \int_k^x \tilde{g}_f(y)\,dy$ satisfies that $f - m_f = Ag_f$ (A is the generator of the diffusion $(X_t)_{t\ge 0}$) μ-almost everywhere, and

$$f(x) - \mathbf{E}\big[f(X)\big] = \frac{1}{2}a(x)\tilde{g}_f'(x) + b(x)\tilde{g}_f(x), \qquad \mu\text{-a.e. } x, \tag{3.3}$$

where X is a random variable with its law μ. The equation (3.3) is a generalized version of Stein's equation.

Remark 3.1. (i) If $f \in L^1(\mu) \cap C(S)$ (where $C(S)$ denotes the class of continuous functions on S), (3.3) holds for all $x \in S$.

(ii) Since μ has the density function p, (3.3) follows almost everywhere with respect to the Lebesgue measure.

Similarly to the original Stein's equation, (3.3) characterizes the distribution of X as follows. This result will play a crucial role in the proofs of the main results in the next sections.

Theorem 3.2. *Assume that $\int_S a(x)\mu(dx) < \infty$. Let Y be a S-valued random variable such that $\mathbf{E}[|b(Y)|] < \infty$. Then, the distribution of Y coincides with μ if and only if*

$$\mathbf{E}\left[\frac{1}{2}a(Y)h'(Y) + b(Y)h(Y)\right] = 0 \tag{3.4}$$

for every $h \in C^1(S)$ such that $\mathbf{E}[|b(Y)h(Y)|] < \infty$ and $\mathbf{E}[|a(Y)h'(Y)|] < \infty$.

Proof. Assume that the distribution of Y is μ. Let $h \in C_b^1(S)$ such that $\mathbf{E}[|b(Y)h(Y)|] < \infty$ and $\mathbf{E}[|a(Y)h'(Y)|] < \infty$, and

$$f(x) := \frac{1}{2p(x)}\frac{d}{dx}\big[a(x)p(x)h(x)\big], \qquad x \in S.$$

Here, note that $ap \in C^1(S)$ follows by the definition of a. Since

$$\begin{aligned}\frac{d}{dx}\big[a(x)p(x)h(x)\big] &= \big[a(x)p(x)\big]'h(x) + a(x)p(x)h'(x)\\ &= 2b(x)p(x)h(x) + a(x)p(x)h'(x),\end{aligned}$$

we have

$$\begin{aligned}\int_S \big|f(x)\big|\mu(dx) &= \frac{1}{2}\int_S \left|\frac{d}{dx}\big[a(x)p(x)h(x)\big]\right| dx\\ &\le \int_S \big|b(x)h(x)\big|p(x)\,dx + \frac{1}{2}\int_S a(x)p(x)\big|h'(x)\big|\,dx\\ &= \mathbf{E}\big[\big|b(X)h(X)\big|\big] + \frac{1}{2}\mathbf{E}\big[a(X)\big|h'(X)\big|\big].\end{aligned}$$

Hence, $f \in L^1(\mu)$. Define $\tilde{g}_f$ as in (3.2). Then, we have $h = \tilde{g}_f$ by explicit calculation. Therefore, in view of (3.3), we obtain (3.4) in the case where h is bounded. The case where h is not bounded is obtained by approximation.

Next, we show that the distribution of Y is μ if (3.4) holds. Let $f \in C_K(S)$, where $C_K(S)$ is the total set of continuous functions on S with compact support. Note that the support of f is

strictly included by S, because S is a open set in $\mathbb{R}$. Since $f(x)=0$ for x sufficiently near to u, there exists $u' \in (l,u)$ such that

$$\int_l^x \big[f(y)-m_f\big]p(y)\,dy = m_f\left(1-\int_l^x p(y)\,dy\right) = m_f\int_x^u p(y)\,dy, \qquad x\in\big[u',u\big),$$

where $m_f = \int_l^u f(x)p(x)\,dx$. On the other hand, for sufficiently small $\varepsilon>0$ the assumption on b implies that there exists $u''\in(l,u)$ such that

$$b(x)<-\varepsilon \quad \text{and} \quad \int_x^u b(y)p(y)\,dy<0$$

for $x\in[u'',u)$. Let $\tilde{u} := \max\{u',u''\}$. Hence, by (3.1) we have for $x\in[\tilde{u},u)$

$$\begin{aligned}\big|\tilde{g}_f(x)\big| &= \frac{2}{a(x)p(x)}\left|\int_l^x \big(f(y)-m_f\big)p(y)\,dy\right| \\ &= \frac{|m_f|\int_x^u p(y)\,dy}{|\int_l^x b(y)p(y)\,dy|} = \frac{|m_f|\int_x^u p(y)\,dy}{|\int_x^u b(y)p(y)\,dy|} \\ &\le \frac{|m_f|}{\varepsilon}.\end{aligned}$$

Similarly, there exists $\tilde{l}\in(l,u)$ such that

$$\big|\tilde{g}_f(x)\big| \le \frac{|m_f|}{\varepsilon}, \qquad x\in(l,\tilde{l}].$$

These estimates and the continuity and positivity of $a(x)p(x)$ in S imply that the function $\tilde{g}_f$ defined by (3.2) is bounded and satisfies $\mathbf{E}[|b(Y)\tilde{g}_f(Y)|]<\infty$. The finiteness of $\mathbf{E}[|a(Y)\tilde{g}_f'(Y)|]$ is obtained from (3.3). By (3.3) and (3.4) we have

$$\mathbf{E}\big[f(Y)\big]-\mathbf{E}\big[f(X)\big] = \mathbf{E}\left[\frac{1}{2}a(Y)\tilde{g}_f'(Y)+b(Y)\tilde{g}_f(Y)\right]=0.$$

Since this equality holds for any $f\in C_K(S)$, X and Y have the same distribution. □

An alternative characterization of the random variables Y with distribution μ has been Theorem 2 in Section 3.2 of [5]. It involves operators from Malliavin calculus and a conditional expectation given the σ-field generated by Y. The same conditional expectation will appear in the statement of the main result in the next section.

Theorem 3.3. *Consider a random variable $Y\in\mathbb{D}^{1,2}$ with its values on S which satisfies that $b(Y)\in L^2(\Omega)$. Then, Y has probability distribution μ if and only if $\mathbf{E}[b(Y)]=0$ and*

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

Stein's bounds below for the distance between the law of an arbitrary random variable and the measure μ are based on the Stein equation (3.3) and the bounds of $\tilde{g}_f$ and $\tilde{g}'_f$ are obtained in Section 3.2 of [5]. Now we state the result in [5] for later use.

Assumption 3.4. (i) If $u < \infty$, assume that there exists $u' \in (l, u)$ such that b is non-increasing and Lipschitz continuous on $[u', u)$ and $\liminf_{x \to u} a(x)/(u - x) > 0$. If $u = \infty$, assume that there exists $u' \in (l, u)$ such that b is non-increasing on $[u', u)$ and $\liminf_{x \to u} a(x) > 0$.

(ii) If $l > -\infty$, assume that 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 $l = -\infty$, assume that there exists $l' \in (l, u)$ such that b is non-decreasing on $(l, l']$ and $\liminf_{x \to l} a(x) > 0$.

Theorem 3.5. (i) *Let d be the Fortet–Mourier distance. Under Assumption 3.4, we have for S-valued random variable $Y \in \mathbb{D}^{1,2}$*

$$d\big(\mathcal{L}(Y), \mu\big) \le C\mathbf{E}\bigg[\mathbf{E}\bigg[\bigg|\frac{1}{2}a(Y) + \big\langle D(-L)^{-1}\big\{b(Y) - \mathbf{E}\big[b(Y)\big]\big\}, DY\big\rangle_H \bigg| Y\bigg|\bigg] + C\big|\mathbf{E}\big[b(Y)\big]\big|,$$

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

(ii) *Let d be the Kolmogorov distance or the total variation distance. Assume* $\inf_{x \in S} a(x) > 0$. *Then, we have for S-valued random variable $Y \in \mathbb{D}^{1,2}$*

$$d\big(\mathcal{L}(Y), \mu\big) \le C\mathbf{E}\bigg[\bigg|\mathbf{E}\bigg[\frac{1}{2}a(Y) + \big\langle D(-L)^{-1}\big\{b(Y) - \mathbf{E}\big[b(Y)\big]\big\}, DY\big\rangle_H \bigg| Y\bigg]\bigg|\bigg] + C\big|\mathbf{E}\big[b(Y)\big]\big|,$$

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

We remark that Assumption 3.4 is designed for the estimate of the Kolmogorov distance or the total variation distance, while the estimate of the Fortet-Mourier distance is obtained under a simple assumption. This difference of the assumptions comes from the estimate of $\tilde{g}'_f$. We also remark that the cases that μ has the normal distribution and μ has the Gamma distribution satisfy the Assumption 3.4 under suitable choices of a and b. See Section 3.2 of [5] for the details.

The Fourth Moment Theorem tells us that the Stein's bound is sharp for multiple integrals in the case of Gaussian law, meaning that a sequence of multiple stochastic integrals convergences to a Gaussian distribution if and only if the right-hand side of the Stein's bound vanishes. The purpose of the last three sections of our paper is to give a necessary and sufficient condition for sequences of random variables to converge to the invariant measure μ of a diffusion processes as described above.

4. Necessary and sufficient conditions for the convergence to the invariant measure of a diffusion

The purpose of this section is to provide necessary and sufficient conditions for the convergence of the sequence $(F_m)_{m\in\mathbb{N}}$ to the invariant measure μ of a diffusion (as described in Section 3). We aim to give such a characterization in terms of the squared diffusion coefficient a and of the Malliavin derivatives of F_m. The main result in [5] (see also Section 3) implies that a sufficient condition for the convergence in distribution of $(F_m)_{m\in\mathbb{N}}$ in a certain class to μ is that $\mathbf{E}[\frac{1}{2}a(F_m) - \langle DF_m, D(-L)^{-1}\{F_m - b(F_m)\}\rangle | F_m]$ converges in $L^1(\Omega)$ to zero as $m \to \infty$ and $\mathbf{E}[b(F_m)]$ converges to zero as $m \to \infty$. But it will follow from the results presented below in this section that this condition is sometimes too strong and it is not a necessary condition for the convergence of $(F_m)_{m\geq 1}$ in distribution to μ. Actually, we will consider the convergence in total variation which is strictly stronger than the convergence in distribution. Recall that the total variation distance between the law of two random variables X and Y is defined by

$$d_{\mathrm{TV}}\big(\mathcal{L}(X), \mathcal{L}(Y)\big) = \sup_A \big|P(X \in A) - P(Y \in A)\big|,$$

where the supremum is taken over all Borel sets $A \subset \mathbb{R}$. We also have (see, e.g., Appendix C in [8])

$$d_{\mathrm{TV}}\big(\mathcal{L}(X), \mathcal{L}(Y)\big) = \frac{1}{2}\sup_h \big|\mathbf{E}\big[h(X)\big] - \mathbf{E}\big[h(Y)\big]\big|,$$

where the supremum is considered over all Borel measurable functions h with $\|h\|_\infty \leq 1$. Let us start with the following result which is connected with Theorem 3.3.

Set $S = (l, u)$, μ, X, p, b, a and k as in Section 3. Recall that $\int_S |b(x)|\mu(dx) < \infty$. We assume that $\int_S a(x)\mu(dx) < \infty$. Denote the Lebesgue measure on $\mathbb{R}$ by dx. Additionally, we consider the function ϕ on S given by

$$\phi(x) := \frac{1}{2}a(x) + \big(|k| + |x|\big)\big|b(x)\big|,$$

where k is an element in S which appeared in the assumption of b.

Theorem 4.1. *Let Y be an S-valued random variable in $\mathbb{D}^{1,2}$ such that $b(Y) \in L^2(\Omega)$, and assume that the distribution of Y is absolutely continuous with respect to the Lebesgue measure. Then, for every $p > 1$ we have*

$$\begin{aligned}
&\mathbf{E}\left[\left|\mathbf{E}\left[\frac{1}{2}a(Y) + \big\langle D(-L)^{-1}\big\{b(Y) - \mathbf{E}\big[b(Y)\big]\big\}, DY\big\rangle_H \Big| Y\right]\right|\right] \\
&\quad \leq \big\{1 + \big(1 + \mathbf{E}\big[\big|b(Y)\big|^p\big] + \mathbf{E}\big[\big|b(X)\big|^p\big]\big)\mathbf{E}\big[|Y - k|\big] + \mathbf{E}\big[\phi(Y)^p\big] + \mathbf{E}\big[\phi(X)^p\big]\big\} \\
&\qquad \times d_{\mathrm{TV}}\big(\mathcal{L}(Y), \mu\big)^{1-1/p}.
\end{aligned}$$

Proof. It is sufficient to show the case that $\mathbf{E}[\phi(X)] < \infty$ and $\mathbf{E}[\phi(Y)] < \infty$. First we show that for $M \geq 1$

$$\begin{aligned}
&\mathbf{E}\Big[\Big|\mathbf{E}\Big[\frac{1}{2}a(Y) + \big\langle D(-L)^{-1}\{b(Y) - \mathbf{E}[b(Y)]\}, DY\big\rangle_H \Big| Y\Big]\Big|\Big] \\
&\quad \leq \big(M + \mathbf{E}[|Y-k|]\big) d_{\mathrm{TV}}\big(\mathcal{L}(Y), \mu\big) \\
&\qquad + \mathbf{E}\big[\phi(Y); \phi(Y) > M\big] + \mathbf{E}\big[\phi(X); \phi(X) > M\big] \\
&\qquad + \mathbf{E}\big[|Y-k|\big]\big(\mathbf{E}\big[\big|b(Y)\big|; \big|b(Y)\big| > M\big] + \mathbf{E}\big[\big|b(X)\big|; \big|b(X)\big| > M\big]\big).
\end{aligned} \tag{4.1}$$

Let $h \in L^\infty(S, dx)$ (the space of essentially bounded functions on S with respect to dx). For $x \in S$ define

$$f_h(x) := \frac{1}{2}a(x)h(x) + b(x)\int_k^x h(y)\,dy.$$

We remark that the existence of the density function of X implies $|h(X)| \leq \|h\|_\infty$ almost surely. Since $\mathbf{E}[|a(X)h(X)|] < \infty$ and

$$\begin{aligned}
\mathbf{E}\Big[\Big|b(X)\int_k^X h(y)\,dy\Big|\Big] &\leq \|h\|_\infty \mathbf{E}\big[\big|b(X)X\big|\big] + |k|\|h\|_\infty \mathbf{E}\big[\big|b(X)\big|\big] \\
&< \infty,
\end{aligned}$$

Theorem 3.2 implies $E[f_h(X)] = 0$. Hence, by integration by parts formula, we have

$$\begin{aligned}
&\mathbf{E}\big[f_h(Y)\big] - \mathbf{E}\big[f_h(X)\big] \\
&\quad = \mathbf{E}\Big[\frac{1}{2}a(Y)h(Y) + b(Y)\int_k^Y h(y)\,dy\Big] \\
&\quad = \mathbf{E}\Big[\frac{1}{2}a(Y)h(Y) + \big(b(Y) - \mathbf{E}[b(Y)]\big)\int_k^Y h(y)\,dy\Big] + \mathbf{E}[b(Y)]\mathbf{E}\Big[\int_k^Y h(y)\,dy\Big] \\
&\quad = \mathbf{E}\Big[\frac{1}{2}a(Y)h(Y) + \big[\delta D(-L)^{-1}\big(b(Y) - \mathbf{E}[b(Y)]\big)\big]\int_k^Y h(y)\,dy\Big] \\
&\qquad + \mathbf{E}[b(Y)]\mathbf{E}\Big[\int_k^Y h(y)\,dy\Big] \\
&\quad = \mathbf{E}\Big[\frac{1}{2}a(Y)h(Y) + \Big\langle D(-L)^{-1}\big(b(Y) - \mathbf{E}[b(Y)]\big), D\int_k^Y h(y)\,dy\Big\rangle_H\Big] \\
&\qquad + \mathbf{E}[b(Y)]\mathbf{E}\Big[\int_k^Y h(y)\,dy\Big] \\
&\quad = \mathbf{E}\Big[h(Y)\Big(\frac{1}{2}a(Y) + \big\langle D(-L)^{-1}\big(b(Y) - \mathbf{E}[b(Y)]\big), DY\big\rangle_H\Big)\Big] \\
&\qquad + \mathbf{E}[b(Y)]\mathbf{E}\Big[\int_k^Y h(y)\,dy\Big].
\end{aligned}$$

Since the duality between the $L^1(S, dx)$ and the $L^\infty(S, dx)$ yields

$$\begin{aligned}
&\mathbf{E}\left[\left|\mathbf{E}\left[\frac{1}{2}a(Y) + \left\langle D(-L)^{-1}\{b(Y) - \mathbf{E}[b(Y)]\}, DY\right\rangle_H \middle| Y\right]\right|\right] \\
&\quad = \sup_{\|h\|_\infty \le 1}\left|\mathbf{E}\left[\left(\frac{1}{2}a(Y) + \left\langle D(-L)^{-1}\{b(Y) - \mathbf{E}[b(Y)]\}, DY\right\rangle_H\right)h(Y)\right]\right|,
\end{aligned}$$

we obtain

$$\begin{aligned}
&\mathbf{E}\left[\left|\mathbf{E}\left[\frac{1}{2}a(Y) + \left\langle D(-L)^{-1}\{b(Y) - \mathbf{E}[b(Y)]\}, DY\right\rangle_H \middle| Y\right]\right|\right] \\
&\quad \le \sup_{\|h\|_\infty \le 1}\left|\mathbf{E}[f_h(Y)] - \mathbf{E}[f_h(X)]\right| + \left|\mathbf{E}[b(Y)]\right|\mathbf{E}[|Y - k|].
\end{aligned} \tag{4.2}$$

Now we consider the estimate of f_h. For $x \in S$, it holds that

$$\begin{aligned}
\left|f_h(x)\right| &\le \left|b(x)\int_k^x h(y)\,dy\right| + \frac{1}{2}a(x)\left|h(x)\right| \\
&\le \|h\|_\infty\left(|k| + |x|\right)\left|b(x)\right| + \frac{1}{2}a(x)\left|h(x)\right|.
\end{aligned}$$

Since the existence of the density function of X and Y implies $|h(X)| \le \|h\|_\infty$ and $|h(Y)| \le \|h\|_\infty$ almost surely, we have

$$\begin{aligned}
\left|f_h(X)\right| &\le \|h\|_\infty\left\{\left(|k| + |X|\right)\left|b(X)\right| + \frac{1}{2}a(X)\right\} = \|h\|_\infty\phi(X), \\
\left|f_h(Y)\right| &\le \|h\|_\infty\left\{\left(|k| + |Y|\right)\left|b(Y)\right| + \frac{1}{2}a(Y)\right\} = \|h\|_\infty\phi(Y)
\end{aligned}$$

almost surely. Hence,

$$\begin{aligned}
&\left|\mathbf{E}[f_h(Y)] - \mathbf{E}[f_h(X)]\right| \\
&\quad \le \left|\mathbf{E}[f_h(Y); \phi(Y) \le M] - \mathbf{E}[f_h(X); \phi(X) \le M]\right| \\
&\qquad + \left|\mathbf{E}[f_h(Y); \phi(Y) > M] - \mathbf{E}[f_h(X); \phi(X) > M]\right| \\
&\quad \le \|h\|_\infty M d_{\mathrm{TV}}\left(\mathcal{L}(Y), \mu\right) \\
&\qquad + \|h\|_\infty\left(\mathbf{E}[\phi(Y); \phi(Y) > M] + \mathbf{E}[\phi(X); \phi(X) > M]\right).
\end{aligned}$$

This inequality and (4.2) yield

$$\begin{aligned}
&\mathbf{E}\left[\left|\mathbf{E}\left[\frac{1}{2}a(Y) + \left\langle D(-L)^{-1}\{b(Y) - \mathbf{E}[b(Y)]\}, DY\right\rangle_H \middle| Y\right]\right|\right] \\
&\quad \le M d_{\mathrm{TV}}\left(\mathcal{L}(Y), \mu\right) + \mathbf{E}[\phi(Y); \phi(Y) > M] + \mathbf{E}[\phi(X); \phi(X) > M] \\
&\qquad + \left|\mathbf{E}[b(Y)]\right|\mathbf{E}[|Y - k|].
\end{aligned} \tag{4.3}$$

Since $\mathbf{E}[b(X)]=0$, we have

$$\begin{aligned}
&\left|\mathbf{E}\left[b(Y)\right]\right| \\
&\quad=\left|\mathbf{E}\left[b(Y)\right]-\mathbf{E}\left[b(X)\right]\right| \\
&\quad\le\left|\mathbf{E}\left[b(Y);\left|b(Y)\right|\le M\right]-\mathbf{E}\left[b(X);\left|b(X)\right|\le M\right]\right| \\
&\qquad+\mathbf{E}\left[\left|b(Y)\right|;\left|b(Y)\right|>M\right]+\mathbf{E}\left[\left|b(X)\right|;\left|b(X)\right|>M\right] \\
&\quad\le M d_{\mathrm{TV}}\left(\mathcal{L}(Y),\mu\right)+\mathbf{E}\left[\left|b(Y)\right|;\left|b(Y)\right|>M\right]+\mathbf{E}\left[\left|b(X)\right|;\left|b(X)\right|>M\right].
\end{aligned}$$

From this inequality and (4.3), we obtain (4.1).

It is easy to see

$$\begin{aligned}
\mathbf{E}\left[\phi(Y);\phi(Y)>M\right]&\le M^{1-p}\mathbf{E}\left[\phi(Y)^p;\phi(Y)>M\right] \\
&\le M^{1-p}\mathbf{E}\left[\phi(Y)^p\right].
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\mathbf{E}\left[\phi(X);\phi(X)>M\right]&\le M^{1-p}\mathbf{E}\left[\phi(X)^p\right], \\
\mathbf{E}\left[\left|b(Y)\right|;\left|b(Y)\right|>M\right]&\le M^{1-p}\mathbf{E}\left[\left|b(Y)\right|^p\right], \\
\mathbf{E}\left[\left|b(X)\right|;\left|b(X)\right|>M\right]&\le M^{1-p}\mathbf{E}\left[\left|b(X)\right|^p\right].
\end{aligned}$$

Hence, (4.1) implies

$$\begin{aligned}
&\mathbf{E}\left[\left|\mathbf{E}\left[\frac{1}{2}a(Y)+\left\langle D(-L)^{-1}\left\{b(Y)-\mathbf{E}\left[b(Y)\right]\right\},DY\right\rangle_H\middle|Y\right]\right|\right] \\
&\quad\le M\left(1+\mathbf{E}\left[|Y-k|\right]\right)d_{\mathrm{TV}}\left(\mathcal{L}(Y),\mu\right) \\
&\qquad+M^{1-p}\left(\mathbf{E}\left[\left|b(Y)\right|^p+\left|b(X)\right|^p\right]\mathbf{E}\left[|Y-k|\right]+\mathbf{E}\left[\phi(Y)^p+\phi(X)^p\right]\right).
\end{aligned}$$

Letting $M=d_{\mathrm{TV}}(\mathcal{L}(Y),\mu)^{-1/p}$ and using the fact that $d_{\mathrm{TV}}(\mathcal{L}(Y),\mu)\le 1$, we obtain the desired estimate. □

Theorems 3.5 and 4.1 give a necessary and sufficient condition of the convergence of in total variation.

Theorem 4.2. *Assume* $\inf_{x\in S}a(x)>0$. *Let* F_m *be a sequence of* S*-valued random variables in* $\mathbb{D}^{1,2}$ *and assume that the distribution of* F_m *is absolutely continuous with respect to the Lebesgue measure for* $m\in\mathbb{N}$. *Assume that there exists* $\varepsilon>0$ *such that*

$$\begin{aligned}
\mathbf{E}\left[|X|\right]+\mathbf{E}\left[\left|b(X)\right|^{1+\varepsilon}\right]+\mathbf{E}\left[\phi(X)^{1+\varepsilon}\right]&<\infty, \\
\sup_{m\in\mathbb{N}}\left(\mathbf{E}\left[|F_m|\right]+\mathbf{E}\left[\left|b(F_m)\right|^{1+\varepsilon}\right]+\mathbf{E}\left[\phi(F_m)^{1+\varepsilon}\right]\right)&<\infty.
\end{aligned}$$

Then, the following statements are equivalent.

(i) *The distribution of F_m converges to the distribution of X in total variation.*
(ii) $\lim_{m\to\infty} \mathbf{E}[|\frac{1}{2}a(F_m) + \mathbf{E}[\langle D(-L)^{-1}\{b(F_m) - \mathbf{E}[b(F_m)]\}, DF_m\rangle_H | F_m]|] = 0$ *and* $\lim_{m\to\infty} E[b(F_m)] = 0$.

Proof. The implication of (i) to (ii) follows from Theorem 4.1, the uniform integrability of $\{b(F_m)\}$ and $\mathbf{E}[b(X)] = 0$. The converse implication follows from Theorem 3.5. □

If the random variables F_m from Theorem 4.2 belong to a Wiener chaos of fixed order and the limit distribution is not trivial, then the convergence in total variation is equivalent to the convergence in distribution (see [10]). Hence, we also have the following theorem.

Theorem 4.3. *Suppose Assumption 3.4. Let F_m be a sequence of S-valued random variables and assume that there exists $p \in \mathbb{N}$ such that F_m is expressed as*

$$F_m = \sum_{k=0}^{p} I_k\big(f_k^{(m)}\big), \qquad f_k^{(m)} \in H^{\odot k}$$

for all $m \in \mathbb{N}$. Furthermore, assume that there exists $\varepsilon > 0$ such that

$$\mathbf{E}\big[|X|\big] + \mathbf{E}\big[\big|b(X)\big|^{1+\varepsilon}\big] + \mathbf{E}\big[\phi(X)^{1+\varepsilon}\big] < \infty,$$
$$\sup_{m\in\mathbb{N}}\Big(\mathbf{E}\big[|F_m|\big] + \mathbf{E}\big[\big|b(F_m)\big|^{1+\varepsilon}\big] + \mathbf{E}\big[\phi(F_m)^{1+\varepsilon}\big]\Big) < \infty.$$

Then, the following statements are equivalent.

(i) *The distribution of F_m converges to the distribution of X in distribution.*
(ii) *The distribution of F_m converges to the distribution of X in total variation.*
(iii) $\lim_{m\to\infty} \mathbf{E}[|\frac{1}{2}a(F_m) + \mathbf{E}[\langle D(-L)^{-1}\{b(F_m) - \mathbf{E}[b(F_m)]\}, DF_m\rangle_H | F_m]|] = 0$ *and* $\lim_{m\to\infty} E[b(F_m)] = 0$.

Proof. The implication of (i) to (ii) follows from Theorem 3.1 in [10]. The implication of (ii) to (iii) follows from Theorem 4.1, the uniform integrability of $\{b(F_m)\}$ and $\mathbf{E}[b(X)] = 0$. The implication of (iii) to (i) follows from Theorem 3.5. □

Remark 4.4. • Our Theorem 4.3 is related to the result stated in Theorem 3.2 in [1]. In this reference, a criterium involving the conditional expectation for the convergence towards a linear combination of chi squared random variable is given.

• The result in Theorem 4.2 indicates that the conditions

$$\lim_{m\to\infty} \mathbf{E}\left[\left|\frac{1}{2}a(F_m) + \mathbf{E}\big[\big\langle DF_m, D(-L)^{-1}\big(b(F_m) - \mathbf{E}\big[b(F_m)\big]\big)\big\rangle_H | F_m\big]\right|\right] = 0 \quad \text{and}$$
$$\lim_{m\to\infty} E\big[b(F_m)\big] = 0$$

are sometimes maybe stronger than the convergence in law.

From now on, for simplicity, we additionally assume that $\mathbf{E}[X] = \int_S x\mu(dx) = 0$ and $b(x) = -x$. In this case, $k = 0 \in S$ and $\phi(x) := \frac{1}{2}a(x) + |x|^2$. Then, we have the following corollaries from Theorems 4.2 and 4.3.

Corollary 4.5. *Assume* $\inf_{x\in S} a(x) > 0$. *Let F_m be a sequence of S-valued random variables in $\mathbb{D}^{1,2}$ and assume that the distribution of F_m is absolutely continuous with respect to the Lebesgue measure for $m \in \mathbb{N}$. Assume that there exists $\varepsilon > 0$ such that*

$$\mathbf{E}\left[\frac{1}{2}a(X)^{1+\varepsilon} + |X|^{2+\varepsilon}\right] < \infty \quad \textit{and} \quad \sup_{m\in\mathbb{N}} \mathbf{E}\left[\frac{1}{2}a(F_m)^{1+\varepsilon} + |F_m|^{2+\varepsilon}\right] < \infty.$$

Then, the following statements are equivalent.

(i) *The distributions of F_m converges to the distribution of X in total variation.*
(ii) $\lim_{m\to\infty} \mathbf{E}[|\frac{1}{2}a(F_m) - \mathbf{E}[\langle D(-L)^{-1}F_m, DF_m\rangle_H | F_m]|] = 0$ *and* $\lim_{m\to\infty} E[F_m] = 0$.

Corollary 4.6. *Suppose Assumption 3.4. Let F_m be a sequence of S-valued random variables and assume that there exists $p \in \mathbb{N}$ such that F_m is expressed as*

$$F_m = \sum_{k=0}^{p} I_k\big(f_k^{(m)}\big), \qquad f_k^{(m)} \in H^{\odot k}$$

for all $m \in \mathbb{N}$. Furthermore, assume that there exists $\varepsilon > 0$ such that

$$\mathbf{E}\left[\frac{1}{2}a(X)^{1+\varepsilon} + |X|^{2+\varepsilon}\right] < \infty \quad \textit{and} \quad \sup_{m\in\mathbb{N}} \mathbf{E}\left[\frac{1}{2}a(F_m)^{1+\varepsilon}\right] < \infty.$$

Then, the following statements are equivalent.

(i) *The distributions of F_m converges to the distribution of X in distribution.*
(ii) *The distributions of F_m converges to the distribution of X in total variation.*
(iii) $\lim_{m\to\infty} \mathbf{E}[|\frac{1}{2}a(F_m) - \mathbf{E}[\langle D(-L)^{-1}F_m, DF_m\rangle_H | F_m]|] = 0$ *and* $\lim_{m\to\infty} E[F_m] = 0$.

Remark 4.7. The conditional expectation

$$\mathbf{E}\big[\big\langle D(-L)^{-1}F, DF\big\rangle_H | F\big]$$

is called the Stein factor associated with the random variable F, and Corollaries 4.5 and 4.6 imply that the convergence of the distributions in total variation is characterized by the Stein factor. Recently, the Stein factor and the applications are studied as a hot topic, and many results have been obtained (see [9] for details).

We continue the consideration of the necessary and sufficient conditions of the convergence in distributions. We additionally assume $\int_S a(x)^2\mu(dx) < \infty$.

Proposition 4.8. *Consider a sequence $(F_m)_{m\geq 1}$ of S-valued random variables in $\mathbb{D}^{1,4}$, which satisfies that there exists $\varepsilon > 0$ satisfying*

$$\sup_m \mathbf{E}\big[a(F_m)^{2+\varepsilon} + |F_m|^{4+\varepsilon}\big] < \infty. \tag{4.4}$$

Suppose that the distribution of F_m converges to μ as $m \to \infty$. Then the condition that $\frac{1}{2}a(F_m) - \langle DF_m, D(-L)^{-1}F_m\rangle_H$ converges to 0 in $L^2(\Omega)$ as $m \to \infty$ is equivalent to the condition that $\frac{1}{4}\mathbf{E}[a(F_m)^2] - \mathbf{E}[\langle DF_m, D(-L)^{-1}F_m\rangle_H^2]$ converges to 0 as $m \to \infty$.

Proof. Since distribution of F_m converges to μ, by Theorem 3.2 we have

$$\lim_{m\to\infty} \mathbf{E}\left[\frac{1}{2}a(F_m)h'(F_m) - F_m h(F_m)\right] = 0$$

for $h \in C^1(S)$ such that $\sup_m \mathbf{E}[|h(F_m)|^{2+\varepsilon} + |h'(F_m)|^{2+\varepsilon}] < \infty$. On the other hand,

$$\mathbf{E}\big[F_m h(F_m)\big] = \mathbf{E}\big[h(F_m)(\delta D)(-L)^{-1}F_m\big] = \mathbf{E}\big[h'(F_m)\big\langle DF_m, D(-L)^{-1}F_m\big\rangle_H\big].$$

Hence, we have

$$\lim_{m\to\infty} \mathbf{E}\left[\left(\frac{1}{2}a(F_m) - \big\langle DF_m, D(-L)^{-1}F_m\big\rangle_H\right)h'(F_m)\right] = 0$$

for $h \in C^1(S)$ such that $\sup_m \mathbf{E}[|h(F_m)|^{2+\varepsilon} + |h'(F_m)|^{2+\varepsilon}] < \infty$. We have assumed (4.4). Hence, we can choose $h \in C^1(S)$ such that $h'(x) = a(x)$. Thus, we obtain

$$\lim_{m\to\infty}\left(\frac{1}{2}\mathbf{E}\big[a(F_m)^2\big] - \mathbf{E}\big[a(F_m)\big\langle DF_m, D(-L)^{-1}F_m\big\rangle_H\big]\right) = 0. \tag{4.5}$$

On the other hand,

$$\begin{aligned}
&\mathbf{E}\left[\left(\frac{1}{2}a(F_m) - \big\langle DF_m, D(-L)^{-1}F_m\big\rangle_H\right)^2\right] \\
&\quad = \mathbf{E}\big[\big\langle DF_m, D(-L)^{-1}F_m\big\rangle_H^2\big] - \frac{1}{4}\mathbf{E}\big[a(F_m)^2\big] \\
&\qquad + \frac{1}{2}\mathbf{E}\big[a(F_m)^2\big] - \mathbf{E}\big[a(F_m)\big\langle DF_m, D(-L)^{-1}F_m\big\rangle_H\big].
\end{aligned}$$

Hence, applying (4.5), we obtain the required equivalence. □

The necessary and sufficient condition in Theorem 4.3 for the convergence of the sequence $(F_m)_{m\geq 1}$ to the distribution μ in total variation is that the sequence

$$\mathbf{E}\left[\frac{1}{2}a(F_m) + \big\langle D(-L)^{-1}\big\{b(F_m) - \mathbf{E}\big[b(F_m)\big]\big\}, DF_m\big\rangle_H \,\middle|\, F_m\right] \quad \text{and} \quad \mathbf{E}\big[b(F_m)\big]$$

converge to zero in $L^1(\Omega)$ as $m \to \infty$. Due to the appearance of the conditional expectation, this condition is sometimes hard to be checked. Therefore, we give below in Theorem 4.9 an alternative result which does not involve the conditional expectation.

Theorem 4.9. *Suppose Assumption 3.4, and assume that there exists a random variable $G \in \mathbb{D}^{1,4}$ such that the distribution of G is equal to μ and $\langle DG, D(-L)^{-1}G\rangle_H$ is measurable with respect to the σ-field generated by G. Consider a sequence $(F_m)_{m\geq 1}$ of S-valued random variables in $\mathbb{D}^{1,4}$ such that $\mathbf{E}[F_m] = 0$ and (4.4) is satisfied. Then, the following statements are equivalent.*

(i) *The vector valued random variable $(F_m, \langle DF_m, D(-L)^{-1}F_m\rangle_H)$ converges to $(G, \langle DG, D(-L)^{-1}G\rangle_H)$ in distribution as $m \to \infty$, and $\{a(F_m)^2\}$ and $\langle DF_m, D(-L)^{-1}F_m\rangle_H^2$ are uniformly integrable.*

(ii) *$\frac{1}{2}a(F_m) - \langle DF_m, D(-L)^{-1}F_m\rangle_H$ converges to 0 in $L^2(\Omega)$ as $m \to \infty$.*

Proof. Assume (ii). By Theorem 3.5, F_m converges to G in distribution. Since $\langle DG, D(-L)^{-1}G\rangle_H$ is measurable with respect to the σ-field generated by G, by Theorem 3.3, we obtain

$$\frac{1}{2}a(G) = \langle DG, D(-L)^{-1}G\rangle_H \tag{4.6}$$

almost surely. Hence, by the convergence of F_m to G in distribution and (ii), we have for $h_1, h_2 \in C_b(\mathbb{R})$

$$\begin{aligned}
&\limsup_{m\to\infty}\Big|\mathbf{E}\big[h_1(F_m)h_2\big(\langle DF_m, D(-L)^{-1}F_m\rangle_H\big)\big] \\
&\quad - \mathbf{E}\big[h_1(G)h_2\big(\langle DG, D(-L)^{-1}G\rangle_H\big)\big]\Big| \\
&\leq \limsup_{m\to\infty}\left|\mathbf{E}\left[h_1(F_m)h_2\left(\frac{1}{2}a(F_m)\right)\right] - \mathbf{E}\left[h_1(G)h_2\left(\frac{1}{2}a(G)\right)\right]\right| \\
&\quad + \limsup_{m\to\infty}\left|\mathbf{E}\left[h_1(F_m)\left\{h_2\left(\frac{1}{2}a(F_m)\right) - h_2\big(\langle DF_m, D(-L)^{-1}F_m\rangle_H\big)\right\}\right]\right| \\
&= 0.
\end{aligned}$$

Thus, (i) is obtained.

Assume (i). Noting that (4.6) holds, by the assumption (i), we have

$$\begin{aligned}
&\lim_{m\to\infty}\left(\frac{1}{4}\mathbf{E}\big[a(F_m)^2\big] - \mathbf{E}\big[\langle DF_m, D(-L)^{-1}F_m\rangle_H^2\big]\right) \\
&\quad = \frac{1}{4}\mathbf{E}\big[a(G)^2\big] - \mathbf{E}\big[\langle DG, D(-L)^{-1}G\rangle_H^2\big] = 0.
\end{aligned}$$

Thus, by Proposition 4.8, (ii) is obtained. □

Remark 4.10. If we suppose the assumptions in both Corollary 4.5 and Theorem 4.9, then the equivalent conditions in Corollary 4.5 and Theorem 4.9 will be also equivalent. This fact perhaps implies that generally the convergence of distributions to the invariant measures of diffusion processes would be complicated, while the fourth moment theorems with respect to Gaussian distributions and Gamma distributions are simple.

In order to apply the above result, we need to know how to compute the squared diffusion coefficient for a given law and to check the measurability of $\langle DG, D(-L)^{-1}G\rangle_H$ with respect to G.

Proposition 4.11. (i) *Let $F = cW(h)$ where $c \in \mathbb{R}$ and $h \in H$ such that $\|h\|_H = 1$. Then,*

$$\big\langle D(-L)^{-1}F, DF\big\rangle_H = c^2.$$

In particular, $\langle D(-L)^{-1}F, DF\rangle_H$ is measurable with respect to the σ-field generated by F.

(ii) *Let $F = c(W(h)^2 - 1)$ where $c \in \mathbb{R}$ and $h \in H$ such that $\|h\|_H = 1$. Then,*

$$\big\langle D(-L)^{-1}\big(F - \mathbf{E}[F]\big), DF\big\rangle_H = 2cF + 2c^2.$$

In particular, $\langle D(-L)^{-1}F, DF\rangle_H$ is measurable with respect to the σ-field generated by F.

(iii) *Let $F = e^{cW(h)}$ where $c \in \mathbb{R}$ and $h \in H$ such that $\|h\|_H = 1$. Then,*

$$\big\langle D(-L)^{-1}\big(F - \mathbf{E}[F]\big), DF\big\rangle_H = c^2 F \int_0^1 F^v e^{\frac{c^2}{2}(1-v^2)}\,dv.$$

In particular, $\langle D(-L)^{-1}(F - \mathbf{E}[F]), DF\rangle_H$ is measurable with respect to the σ-field generated by F.

(iv) *Let $F = e^{c\sum_{k=1}^n W(h_k)^2}$ where $n \in \mathbb{N}$, $c \in (-\infty, 1/2)$ and $h_1, h_2, \dots, h_n \in H$ such that $\|h_k\|_H = 1$ for $k = 1, 2, \dots, n$, and $\langle h_k, h_l\rangle_H = 0$ for $k, l = 1, 2, \dots, n$. Then,*

$$\big\langle D(-L)^{-1}\big(F - \mathbf{E}[F]\big), DF\big\rangle_H = 4cF(\log F)\int_0^1 \frac{vF^{\frac{v^2}{1-2c(1-v^2)}}}{(1-2c(1-v^2))^{\frac{n}{2}+1}}\,dv.$$

In particular, $\langle D(-L)^{-1}(F - \mathbf{E}[F]), DF\rangle_H$ is measurable with respect to the σ-field generated by F.

(v) *Let $F = W(h)^n - \mathbf{E}[W(h)^n]$ where $n \in \mathbb{N}$ and $h \in H$ such that $\|h\|_H = 1$. Then,*

$$\begin{aligned}&\big\langle D(-L)^{-1}F, DF\big\rangle_H \\ &\quad = \frac{n^2}{2}\sum_{l=0}^{\lfloor (n-1)/2\rfloor} \frac{(n-1)!(2l-1)!!}{(2l)!(n-1-2l)!}\beta\Big(\frac{n}{2} - l, l+1\Big)\big|F + \mathbf{E}\big[W(h)^n\big]\big|^{2-2(l+1)/n},\end{aligned}$$

where $\lfloor (n-1)/2\rfloor$ is the largest integer which is no larger than $(n-1)/2$. In particular, $\langle D(-L)^{-1}F, DF\rangle_H$ is measurable with respect to the σ-field generated by F.

Proof. When F is given as in (i), $\mathbf{E}[F]=0$ and

$$\big\langle D(-L)^{-1}F, DF\big\rangle_H = c^2\big\langle DW(h), DW(h)\big\rangle_H = c^2.$$

Hence, (i) holds.

When F is given as in (ii), $\mathbf{E}[F]=0$ and

$$\begin{aligned}\big\langle D(-L)^{-1}F, DF\big\rangle_H &= \frac{1}{2}c^2\big\langle D\big(W(h)^2-1\big), D\big(W(h)^2-1\big)\big\rangle_H \\ &= 2c^2W(h)^2 = 2cF+2c^2.\end{aligned}$$

Hence, (ii) holds.

Next, we show the case that F is given as in (iii). In this case, we use the following formula proved in [11]: if $Y=f(N)-\mathbf{E}[f(N)]$ where $f\in C_b^1(\mathbb{R}^n;\mathbb{R})$ with bounded derivatives and $N=(N_1,\dots,N_n)$ is a Gaussian vector with zero mean and covariance matrix $K=(K_{i,j})_{i,j=1,\dots,n}$ then

$$\begin{aligned}&\big\langle D(-L)^{-1}\big(Y-\mathbf{E}[Y]\big), DY\big\rangle_H \\ &\quad= \int_0^\infty e^{-u}\,du\mathbf{E}'\left[\sum_{i,j=1}^n K_{i,j}\frac{\partial f}{\partial x_i}(N)\frac{\partial f}{\partial x_j}\big(e^{-u}N+\sqrt{1-e^{-2u}}N'\big)\right].\end{aligned} \tag{4.7}$$

Here N' denotes an independent copy of N, and N and N' are defined on a product probability space $(\Omega\times\Omega', \mathcal{F}\otimes\mathcal{F}, P\times P')$ and $\mathbf{E}'$ denotes the expectation with respect to the probability measure P'.

By (4.7), we have

$$\begin{aligned}&\big\langle D(-L)^{-1}\big(F-\mathbf{E}[F]\big), DF\big\rangle_H \\ &\quad= \int_0^\infty e^{-u}\,du\mathbf{E}'\big[c^2e^{cW(h)}e^{c(e^{-u}W(h)+\sqrt{1-e^{-2u}}W'(h))}\big] \\ &\quad= c^2e^{cW(h)}\int_0^1 e^{cvW(h)}\mathbf{E}'\big[e^{c\sqrt{1-v^2}W'(h)}\big]\,dv \\ &\quad= c^2F\int_0^1 F^v e^{\frac{c^2}{2}(1-v^2)}\,dv.\end{aligned}$$

Hence, (iii) holds.

Next, we show the case that F is given as in (iv). In this case, we use Lemma 1 in Section 4 of [5]. The statement of the lemma is as follows. Let Z be a random variable with the standard normal distribution, $K>-\frac{1}{2}$, $C\in\mathbb{R}$ and $a\in(0,1)$. Then,

$$\mathbf{E}\big[e^{-K(C+\sqrt{1-a^2}Z)^2}\big] = \frac{1}{\sqrt{1+2K(1-a^2)}}e^{-\frac{C^2K}{1+2K(1-a^2)}}, \tag{4.8}$$

$$\mathbf{E}\big[\big(C+\sqrt{1-a^2}Z\big)e^{-K(C+\sqrt{1-a^2}Z)^2}\big]=\frac{C}{(1+2K(1-a^2))^{\frac{3}{2}}}e^{-\frac{C^2K}{1+2K(1-a^2)}}. \tag{4.9}$$

It is sufficient to show the case that $c\neq 0$. By (4.7), we have

$$\begin{aligned}
&\big\langle D(-L)^{-1}\big(F-\mathbf{E}[F]\big), DF\big\rangle_H \\
&\quad=\int_0^\infty e^{-u}\,du\mathbf{E}'\Bigg[\sum_{i=1}^n 2cW(h_i)e^{c\sum_{j=1}^n W(h_j)^2} \\
&\qquad\times 2c\big(e^{-u}W(h_i)+\sqrt{1-e^{-2u}}W'(h_i)\big)e^{c\sum_{j=1}^n(e^{-u}W(h_j)+\sqrt{1-e^{-2u}}W'(h_j))^2}\Bigg] \\
&\quad=4c^2\int_0^\infty e^{-u}\,du\sum_{i=1}^n W(h_i)e^{c\sum_{j=1}^n W(h_j)^2}\prod_{j\neq i}\mathbf{E}'\big[e^{c(e^{-u}W(h_j)+\sqrt{1-e^{-2u}}W'(h_j))^2}\big] \\
&\qquad\times\mathbf{E}'\big[\big(e^{-u}W(h_i)+\sqrt{1-e^{-2u}}W'(h_i)\big)e^{c(e^{-u}W(h_i)+\sqrt{1-e^{-2u}}W'(h_i))^2}\big] \\
&\quad=4c^2\int_0^1 dv\sum_{i=1}^n W(h_i)e^{c\sum_{j=1}^n W(h_j)^2}\Bigg(\prod_{j\neq i}\mathbf{E}'\big[e^{c(vW(h_j)+\sqrt{1-v^2}W'(h_j))^2}\big]\Bigg) \\
&\qquad\times\mathbf{E}'\big[\big(vW(h_i)+\sqrt{1-v^2}W'(h_i)\big)e^{c(vW(h_i)+\sqrt{1-v^2}W'(h_i))^2}\big].
\end{aligned}$$

Applying (4.8) and (4.9) to this equation, we obtain

$$\begin{aligned}
&\big\langle D(-L)^{-1}\big(F-\mathbf{E}[F]\big), DF\big\rangle_H \\
&\quad=4c^2\int_0^1 dv\sum_{i=1}^n W(h_i)e^{c\sum_{j=1}^n W(h_j)^2}\Bigg(\prod_{j\neq i}\frac{\exp(\frac{cv^2W(h_j)^2}{1-2c(1-v^2)})}{\sqrt{1-2c(1-v^2)}}\Bigg) \\
&\qquad\times\frac{vW(h_i)\exp(\frac{cv^2W(h_j)^2}{1-2c(1-v^2)})}{(1-2c(1-v^2))^{\frac{3}{2}}} \\
&\quad=4c^2\Bigg(\sum_{i=1}^n W(h_i)^2\Bigg)e^{c\sum_{j=1}^n W(h_j)^2}\int_0^1\frac{v\exp(\frac{cv^2\sum_{j=1}^n W(h_j)^2}{1-2c(1-v^2)})}{(1-2c(1-v^2))^{\frac{n}{2}+1}}\,dv \\
&\quad=4cF(\log F)\int_0^1\frac{vF^{\frac{v^2}{1-2c(1-v^2)}}}{(1-2c(1-v^2))^{\frac{n}{2}+1}}\,dv.
\end{aligned}$$

Hence, (iv) holds.

Finally, we show (v). By (4.7), we have

$$\begin{aligned}
&\langle DF, D(-L)^{-1}F\rangle_H \\
&\quad = n^2\int_0^\infty e^{-u}\,du\mathbf{E}'\big[W(h)^{n-1}\big(e^{-u}W(h)+\sqrt{1-e^{-2u}}W'(h)\big)^{n-1}\big] \\
&\quad = n^2W(h)^{n-1}\int_0^1 dy\mathbf{E}'\big[\big(yW(h)+\sqrt{1-y^2}W'(h)\big)^{n-1}\big],
\end{aligned}$$

where $W'(h)$ denotes an independent copy of $W(h)$, and $W(h)$ and $W'(h)$ are defined on a product of abstract Wiener spaces $(\Omega\times\Omega, H\otimes H, P\times P')$ and $\mathbf{E}'$ denotes the expectation with respect to the Gaussian measure P'. Hence,

$$\begin{aligned}
&\langle DF, D(-L)^{-1}F\rangle_H \\
&\quad = n^2W(h)^{n-1}\int_0^1 dy\sum_{j=0}^{n-1}\frac{(n-1)!}{j!(n-1-j)!}\big(yW(h)\big)^{n-1-j}\mathbf{E}'\big[\big(\sqrt{1-y^2}W'(h)\big)^j\big] \\
&\quad = n^2\sum_{l=0}^{\lfloor(n-1)/2\rfloor}\frac{(n-1)!(2l-1)!!}{(2l)!(n-1-2l)!}W(h)^{2n-2-2l}\int_0^1 y^{n-1-2l}\big(1-y^2\big)^l\,dy,
\end{aligned}$$

where we regard $(-1)!!=1$. Noting that the beta function $\beta(a,b)$ for $a,b>0$ satisfies

$$\begin{aligned}
\beta(a,b) &= \int_0^1 t^{a-1}(1-t)^{b-1}\,dt \\
&= 2\int_0^1 s^{2a-1}\big(1-s^2\big)^{b-1}\,ds,
\end{aligned}$$

we obtain

$$\begin{aligned}
&\langle DF, D(-L)^{-1}F\rangle_H \\
&\quad = \frac{n^2}{2}\sum_{l=0}^{\lfloor(n-1)/2\rfloor}\frac{(n-1)!(2l-1)!!}{(2l)!(n-1-2l)!}\beta\left(\frac{n}{2}-l, l+1\right)\big|F+\mathbf{E}\big[W(h)^n\big]\big|^{2-2(l+1)/n}.
\end{aligned}$$

□

Now we give several common probability distributions as examples.

Example 4.12 (Gaussian and Gamma). Let $F=cW(h)$ where $c\in\mathbb{R}$ and $h\in H$ such that $\|h\|_H=1$. Then, $\mathbf{E}[F]=0$. By (i) of Proposition 4.11, we have $\langle D(-L)^{-1}F, DF\rangle_H$ is measurable with respect to the σ-field generated by F, and we have

$$\frac{1}{2}a(x)=c^2, \qquad x\in\mathbb{R}.$$

Let $F = c(W(h)^2 - 1)$ where $c \in \mathbb{R}$ and $h \in H$ such that $\|h\|_H = 1$. Then, $\mathbf{E}[F] = 0$. By (ii) of Proposition 4.11, we have $\langle D(-L)^{-1}F, DF\rangle_H$ is measurable with respect to the σ-field generated by F, and

$$\frac{1}{2}a(x) = 2c^2(x+1), \qquad x > -1.$$

Example 4.13 (Log-normal). Let $F = e^{cW(h)}$ where $c \in \mathbb{R} \setminus \{0\}$ and $h \in H$ such that $\|h\|_H = 1$. By (iii) of Proposition 4.11, we have $\langle D(-L)^{-1}(F - \mathbf{E}[F]), DF\rangle_H$ is measurable with respect to the σ-field generated by F, and

$$\frac{1}{2}a(x) = c^2 x \int_0^1 x^v e^{\frac{c^2}{2}(1-v^2)}\, dv, \qquad x > 0.$$

Example 4.14 (Exponential of the sum of Gaussian squares). Let $F = e^{c\sum_{k=1}^n W(h_k)^2}$ where $n \in \mathbb{N}$, $c \in (-\infty, 1/2) \setminus \{0\}$ and $h_1, h_2, \dots, h_n \in H$ such that $\|h_k\|_H = 1$ for $k = 1, 2, \dots, n$, and $\langle h_k, h_l\rangle_H = 0$ for $k, l = 1, 2, \dots, n$. By (iv) of Proposition 4.11, we have $\langle D(-L)^{-1}(F - \mathbf{E}[F]), DF\rangle_H$ is measurable with respect to the σ-field generated by F and

$$\frac{1}{2}a(x) = 4cx(\log x)\int_0^1 \frac{v x^{\frac{v^2}{1-2c(1-v^2)}}}{(1-2c(1-v^2))^{\frac{n}{2}+1}}\, dv$$

for x belonging to the support of the law (that depends on the choice of c). This case includes the uniform distribution $U[0,1]$ (by taking for example, $n = 2$ and $c = -\frac{1}{2}$), the centered Pareto distribution (if we consider $n = 2$ and $c = \frac{1}{4}$ we obtain a centered Pareto distribution with parameter 2) or the centered beta distribution (by taking $c = -1$, $n = 2$ we have a random variable with centered beta law with parameters $\frac{1}{2}$ and 1). We refer to [5] for a review on these probability distributions.

Example 4.15 (Power of Gaussian). Consider the random variable

$$G = W(h)^n - \mathbf{E}\big[W(h)^n\big],$$

where $h \in H$ such that $\|h\|_H = 1$ and $n \in \mathbb{N}$. By (v) of Proposition 4.11, we have $\langle D(-L)^{-1}(F - \mathbf{E}[F]), DF\rangle_H$ is measurable with respect to the σ-field generated by F and

$$\begin{aligned}\frac{1}{2}a(x) &\\ &= \frac{n^2}{2}\sum_{l=0}^{\lfloor (n-1)/2\rfloor} \frac{(n-1)!(2l-1)!!}{(2l)!(n-1-2l)!}\beta\left(\frac{n}{2}-l, l+1\right)\big|x + \mathbf{E}\big[W(h)^n\big]\big|^{2-2(l+1)/n}\end{aligned}$$

for x in the support of the law of G. The case that $n = 1$ and $n = 2$ are associated to the standard normal distribution and the centered Gamma distribution, respectively (see Example 4.12). When

$n = 3$,

$$\frac{1}{2}a(x) = 3|x|^{4/3} + 6|x|^{2/3}, \qquad x \in \mathbb{R}.$$

When $n = 4$,

$$\frac{1}{2}a(x) = 4(x+3)^{3/2} + 12(x+3), \qquad x \in (-3, \infty).$$

Many examples in which $\langle D(-L)^{-1}G, DG\rangle_H$ is not measurable with respect to G can be constructed. We will illustrate below such a situation.

Example 4.16. Let $G = I_2(f \tilde{\otimes} g)$ with $f, g \in H$, $\langle f, g\rangle_H = 0$. We have $G = I_1(f)I_1(g)$ and clearly $DG = I_1(f)g + I_1(g)f$ and $D(-L)^{-1}G = \frac{1}{2}(I_1(f)g + I_1(g)f)$. This implies $\langle D(-L)^{-1}G, DG\rangle_H = \frac{1}{2}(I_1(f)^2 + I_1(g)^2)$ and this random variable is not measurable with respect to the σ-algebra generated by $I_1(f)I_1(g)$.

Remark 4.17. A concrete example of application of the theory presented in this section is given in [17] and concerns the convergence of a sequence of Pareto distributed random variables to the exponential law. Indeed, if we consider a sequence of random variables $(F_n)_{n\geq 1}$ such that each F_n follows a Pareto distribution with parameter $n \geq 2$ then it is well known that the sequence $(Y_n)_{n\geq 1}$ given by $Y_n = (n-1)(F_n - 1)$ converges in law, as $n \to \infty$, to the exponential distribution with parameter 1. The rate of convergence has been estimated in [17] using the techniques presented in this section.

5. Extension of the Fourth Moment Theorem

In this section, we analyze the weak convergence of sequences of random variables in a Wiener chaos to $X \sim \mu$. The purpose of the argument is to generalize the Fourth Moment Theorem (Theorem 2.2).

Through this section, μ is assumed to satisfy the hypothesis in the previous section. In addition we set $b(x) := -x$ and assume that $\int a(x)^2\mu(dx) < \infty$ and the expectation of the invariant measure equals zero, that is, $\int x\mu(dx) = \mathbf{E}[X] = 0$.

Let us assume that $F_m = I_p(f_m), m \in \mathbb{N}$ is sequence of multiple integrals in the pth Wiener, with $f_m \in H^{\odot p}$ for every $m \in \mathbb{N}$. Our result in Theorem 4.9 says that the joint convergence in distribution as $m \to \infty$ of $(F_m, \frac{1}{p}\|DF_m\|_H^2)$ to $(G, \langle DG, D(-L)^{-1}G\rangle_H)$, with $G \sim \mu$, is equivalent to the convergence in $L^2(\Omega)$ of $\frac{1}{2}a(F_m) - \frac{1}{p}\|DF_m\|_H^2$ to zero as $m \to \infty$, where a denotes the squared diffusion coefficient associated to the law μ.

The nice results in [14] and [7] show that, in the case when the target distribution μ is Gaussian or (centered) Gamma, then the hypothesis on the convergence of $\|DF_m\|_H$ (or the conditional expectation in Theorem 4.2) can be eliminated. Indeed, by assuming that $\mathbf{E}F_m^2 \to_m \mathbf{E}G^2$, the sequence $(F_m)_{m\geq 1}$ convergence to μ (which is Gaussian or Gamma) if and only if the sequence $\frac{1}{2}a(F_m) - \frac{1}{p}\|DF_m\|_H^2$ converges to zero in $L^2(\Omega)$.

It is then natural to ask if the same holds for other target distributions. We will show in the sequel that the answer is in general negative. We will show that in some special situations, the convergence in law of a sequence of multiple integrals to a law (different from Gaussian and Gamma) can be expressed in terms of the convergence of sequence of the fourth moment and also in terms of some squared diffusion coefficients associated to the law and of the Malliavin derivatives.

5.1. Product normal distribution

Let $X, Y \sim N(0,1)$ be two independent random variables ($X = I_1(h_1)$, $Y = I_1(h_2)$ with $h_1, h_2 \in H$ orthonormal) and consider the random variable

$$G = XY.$$

Then G follows the so-called product normal distribution and its density is given by

$$f_G(x) = K_0\big(|x|\big), \qquad x \in \mathbb{R},$$

where K_0 is the modified Bessel function of the second kind. We remark that $f_G(x)$ diverges at $x = 0$, in particular f_G is not continuous in $\mathbb{R}$. Hence, Theorems in Sections 3 and 4 are not applicable in this case.

We will characterize the convergence in distribution (which is equivalent, on Wiener chaos, to the convergence in total variation, see Remark 5.3) of a particular class of sequences of multiple integrals toward the product normal distribution.

Remark 5.1. Let us recall that, if X, Y are independent multiple integrals, then (X, DX) and (Y, DY) are independent random vectors. Moreover, $\langle DX, DY\rangle_H = 0$. See Lemma 1 in [3].

Theorem 5.2. *Suppose $(X_k = I_p(f_k))_{k\in\mathbb{N}}$ and $(Y_k = I_q(g_k))_{k\in\mathbb{N}}$ are two sequences of multiple integrals in the pth and qth Wiener chaos respectively, such that*

$$\lim_{k\to\infty} \mathbf{E}\big[X_k^2\big] = 1 \quad \textit{and} \quad \lim_{k\to\infty} \mathbf{E}\big[Y_k^2\big] = 1.$$

Assume that for every $k \in \mathbb{N}$, the random variables X_k and Y_k are independent. Then the following are equivalent:

(i) *X_kY_k converges in distribution as $k \to \infty$ to XY.*
(ii) $\lim_{k\to\infty} \mathbf{E}[(X_kY_k)^4] = 9$.
(iii) $\lim_{k\to\infty} \|f_k \otimes_r f_k\|_{H^{\otimes 2(p-r)}} = 0$ *and* $\lim_{k\to\infty} \|g_k \otimes_l g_k\|_{H^{\otimes 2(q-l)}} = 0$ *for every* $r = 1, 2, \dots, p-1$ *and* $l = 1, 2, \dots, q-1$.
(iv) $\lim_{k\to\infty} \|DX_k\|_H^2 = p$ *and* $\lim_{k\to\infty} \|DY_k\|_H^2 = q$ *in* $L^2(\Omega)$.
(v) $\lim_{k\to\infty} \mathbf{E}[X_k^4] = 3$ *and* $\lim_{k\to\infty} \mathbf{E}[Y_k^4] = 3$.
(vi) $X_k \to_{k\to\infty} X$ *and* $Y_k \to_{k\to\infty} Y$ *in distribution.*

Proof. Note that for every $k \in \mathbb{N}$, $X_k Y_k = I_{p+q}(f_k \otimes g_k)$. This comes from the product formula for multiple stochastic integrals (2.3) and the fact that, for every $k \geq 1$, the independence of X_k and Y_k is equivalent to $f_k \otimes_1 g_k = 0$ almost everywhere, see [18]. Note that the implication (vi) $\to$ (i) is trivial, and the equivalence between (iii), (iv), (v) and (vi) follows from Theorem 2.2. Hence, it is sufficient to prove

$$\text{(i)} \to \text{(ii)} \to \text{(iii)}$$

for the first assertion. The implication (i) $\to$ (ii) is obvious since $X_k Y_k$ is a multiple integral in the Wiener chaos of order $p+q$ and thus for every $r \geq 1$, $\sup_k \mathbf{E}[|X_k Y_k|^r] < \infty$. Let us show (ii) $\to$ (iii). We use formula (2.5) for the fourth moment of a multiple integral and we get

$$\mathbf{E}\big[I_n(f)^4\big] = 3\mathbf{E}\big[I_n(f)^2\big]^2 + R_f,$$

where, for every $f \in H^{\odot n}$, we denoted

$$R_f := n!^2 \sum_{r=1}^{n-1} \left(\frac{n!}{r!(n-r)!}\right)^2 \big[\|f \otimes_r f\|^2_{H^{\otimes 2(n-r)}} + C^{n-r}_{2n-2r} \|f \tilde{\otimes}_r f\|^2_{H^{\otimes 2(n-r)}}\big]. \tag{5.1}$$

We have, for every $k \geq 1$,

$$\begin{aligned}\mathbf{E}\big[(X_k Y_k)^4\big] &= \mathbf{E}\big[X_k^4\big]\mathbf{E}\big[Y_k^4\big] \\ &= \big(3\mathbf{E}\big[X_k^2\big]^2 + R_{f_k}\big)\big(3\mathbf{E}\big[Y_k^2\big]^2 + R_{g_k}\big)\end{aligned}$$

with R_{f_k} and R_{g_k} defined above by (5.1). Since $3\mathbf{E}[X_k^2]^2 \times 3\mathbf{E}[Y_k^2]^2 \to 9$, we easily get that $\|f_k \otimes_r f_k\|_{H^{\otimes 2(p-r)}} \to_{k\to\infty} 0$ and $\|g_k \otimes_l g_k\|_{H^{\otimes 2(q-l)}} \to_{k\to\infty} 0$ for every $r = 1, \dots, p-1$ and $l = 1, \dots, q-1$. □

Remark 5.3. The convergence of polynomials of (whose orders are uniformly dominated from above) of multiple integrals to a non-degenerate distribution in distribution is equivalent to the convergence in total variation (see Theorem 3.1 in [10]).

5.2. The sum of a standard normal random variable and an independent random variable

Let $Z \sim N(0,1)$ and let G be a centered random variable with the fourth moment and is independent of Z. We will characterize the convergence (of a special class of sequences of multiple stochastic integrals) to the distribution of the random variable $Z + G$. See [3] for the case when Z, G are independent Gamma distributed random variables. We assume that G has a continuous density function on $\mathbb{R}$ (not necessary strictly positive on $\mathbb{R}$). Below a will denote the squared diffusion coefficient associated to $Z + G$.

Proposition 5.4. *Let $(X_k = I_p(f_k))_{k\in\mathbb{N}}$ be a sequence of random variables in the pth Wiener chaos such that*

$$\lim_{k\to\infty} \mathbf{E}\big[X_k^2\big] = 1.$$

Let $(Y_k = I_q(g_k))_{k\in\mathbb{N}}$ be another sequence of random variables in the qth Wiener chaos such that X_k and Y_k are independent for every $k \in \mathbb{N}$ and

$$\lim_{k\to\infty} \mathbf{E}\big[Y_k^2\big] = \mathbf{E}\big[G^2\big] \quad \textit{and} \quad \lim_{k\to\infty} \mathbf{E}\big[Y_k^4\big] = \mathbf{E}\big[G^4\big].$$

Then the following assertions are equivalent:

(i) *$X_k \to_k Z$ and $Y_k \to_k G$ in distribution*,
(ii) *$X_k + Y_k \to_k Z + G$ in distribution.*

If in addition, a satisfies the assumptions in Corollary 4.6, then the following is also equivalent.

(iii) $\frac{1}{2}a(X_k + Y_k) - \mathbf{E}[\frac{1}{p}\|DY_k\|_H^2 + \frac{1}{q}\|DX_k\|_H^2 | X_k + Y_k] \to_k 0$ *in* $L^1(\Omega)$.

Proof. Clearly (i) → (ii). We show that (ii) → (i). We use the relation

$$\begin{aligned}\mathbf{E}\big[(X_k + Y_k)^4\big] &= \mathbf{E}\big[X_k^4\big] + 6\mathbf{E}\big[X_k^2\big]\mathbf{E}\big[Y_k^2\big] + \mathbf{E}\big[Y_k^4\big] \\ &= 3\mathbf{E}\big[X_k^2\big] + 6\mathbf{E}\big[X_k^2\big]\mathbf{E}\big[Y_k^2\big] + \mathbf{E}\big[Y_k^4\big] + R_{f_k} \\ &\sim 3 + 6\mathbf{E}\big[G^2\big] + \mathbf{E}\big[G^4\big] + R_{f_k},\end{aligned} \tag{5.2}$$

where we used (2.5) and R_{f_k} given by (5.1). Recall that $\sim$ means that the sides have the same limit as $k \to \infty$. On the other hand, from (ii) we have

$$\mathbf{E}\big[(X_k + Y_k)^4\big] \to_k \mathbf{E}\big[(Z + G)^4\big] = 3 + 6\mathbf{E}\big[G^2\big] + \mathbf{E}\big[G^4\big]. \tag{5.3}$$

By combining (5.2) and (5.3), we get $R_{f_k} \to_k 0$ and thus

$$\|f_k \otimes_r f_k\|_{H^{\otimes 2(p-r)}} \to_k 0$$

for every $r = 1, 2, \ldots, p-1$. This gives the converges in distribution of X_k to Z as $k \to \infty$ (see Theorem 2.2), and then we easily get the convergence in law as $k \to \infty$ of Y_k to G.

Theorem 3.1 in [10] implies that (ii) is equivalent to the convergence of $X_k + Y_k$ in total variation. Hence, if we additionally give the assumptions in Corollary 4.6, we obtain the equivalence between (ii) and (iii) by Corollary 4.6. □

We apply the above result to the particular case when G follows a centered chi-squared distribution.

Corollary 5.5. *Consider two sequences $(X_k)_{k\geq 1}$ and $(Y_k)_{k\geq 1}$ as in Proposition 5.4 and assume that the random variable G from Proposition 5.4 is a centered Gamma random variable i.e. $G = W(h)^2 - 1$ where $h \in H$, $\|h\|_H = 1$. Then the following are equivalent*:

(i) $X_k + Y_k \to_k Z + G$ *in distribution.*
(ii) $X_k \to_k Z$ *and* $Y_k \to_k G$ *in distribution.*
(iii) $\frac{1}{p}\|DX_k\|_H^2 \to 1$ *and* $\frac{1}{q}\|DY_k\|_H^2 - 2Y_k \to_k 0$ *in* $L^2(\Omega)$.
(iv) $1 - \frac{1}{p}\|DX_k\|_H^2 + 2Y_k - \frac{1}{q}\|DY_k\|_H^2 \to_k 0$ *in* $L^2(\Omega)$.
(v) $\frac{1}{2}a(X_k + Y_k) - \mathbf{E}[\frac{1}{p}\|DX_k\|_H^2 + \frac{1}{q}\|DY_k\|_H^2 | X_k + Y_k] \to_k 0$ *in* $L^1(\Omega)$.

Proof. The points (i) and (v) are equivalent due to Proposition 5.4 by noting the squared diffusion coefficient a associated with the law $Z + G$ satisfies the assumptions from Corollary 4.6. The fact that (i) is equivalent to (ii), (iii), (iv) follows from Theorem 1.2 in [7] and Remark 5.1. On the other hand, the explicit expression for a (that can be calculated via (3.1)), is pretty complex. □

Remark 5.6. Condition (iv) above show that the convergence toward the sum $Z + G$ of a Gaussian and an independent Gamma distribution can be characterized (without the appearance of the conditional expectation) in terms of the diffusion coefficients associated with Z and G.

6. The case when the diffusion coefficient is a polynomial of second degree

Since we are studying the convergence of a sequence of multiple stochastic integrals, whose expectation is zero, we will assume that the measure μ is centered and the drift coefficient is $b(x) = -x$. We will also assume that the diffusion coefficient is a polynomial of second degree expressed as

$$a(x) = \alpha x^2 + \beta x + \gamma, \qquad x \in S, \alpha, \beta, \gamma \in \mathbb{R} \tag{6.1}$$

such that $a(x) > 0$ for every $x \in S$. In this case, it is possible to understand better when the necessary and sufficient condition for the weak convergence (or equivalently, the convergence in total variation, see Remark 5.3) of a sequence of multiple integrals toward the law μ is satisfied.

This class contains the known continuous probability distributions. See Table 1 in [2]. It contains among others, the normal, Gamma, uniform, Student, Pareto, inverse Gamma or F distributions. We mention by the way that, except the first laws listed here, the others cannot be limits in distribution of a sequence of multiple integrals, because they do not admit moments of any order. Nevertheless, the class satisfying (6.1) and admitting moments of any order contains an infinity of probability distributions.

We give below some examples qualified as candidate to be limit in distribution of sequences of multiple integrals. Note that in Table 1 in [2] the diffusion coefficient is given for non centered probability distributions. In order to obtain the diffusion coefficient a of the centered measure from the diffusion coefficient a_0 of the non-centered measure, we apply the following rule (see Lemma 2.5 in [2])

$$a(x) = a_0\big(x + \mathbf{E}[X]\big),$$

where $X \sim \mu$.

Example 6.1 (The normal distribution $N(0,\gamma)$, $\gamma > 0$). In this case $a(x) = 2\gamma$.

Example 6.2 (The Gamma $\Gamma(a,\lambda)$, $a,\lambda > 0$ law). Here the density is $f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}$ for $x > 0$ and $f(x) = 0$ otherwise. Also $\mathbf{E}[X] = \frac{a}{\lambda}$ and the centered Gamma law has

$$a(x) = \frac{2}{\lambda}\left(x + \frac{a}{\lambda}\right)$$

meaning that $\alpha = 0$, $\beta = \frac{2}{\lambda}$, $\gamma = \frac{2a}{\lambda^2}$.

Example 6.3 (The uniform $U(0,1)$ distribution). Here the density is $f(x) = 1_{[0,1]}(x)$, the mean is $\mathbf{E}[X] = \frac{1}{2}$ and $U[0,1] - \mathbf{E}[U[0,1]]$ has squared diffusion coefficient

$$a(x) = \left(x + \frac{1}{2}\right)\left(\frac{1}{2} - x\right) = \frac{1}{4} - x^2.$$

So $\alpha = -1$, $\beta = 0$, $\gamma = \frac{1}{4}$.

Example 6.4 (The Beta $\beta(a,b)$ law, $a,b > 0$). In this case, the probability density function is

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} 1_{(0,1)}(x),$$

$\mathbf{E}[X] = \frac{a}{a+b}$ and the centered beta law has

$$a(x) = \frac{2}{a+b}\left(x + \frac{a}{a+b}\right)\left(\frac{b}{a+b} - x\right).$$

Note that the beta distribution with $a = b$ has the fourth cumulant negative and therefore it cannot be limit of a sequence of multiple integrals.

Therefore $\alpha = -\frac{2}{a+b}$, $\beta = \frac{2}{a+b}\frac{b-a}{a+b}$, $\gamma = \frac{2}{a+b}\frac{a}{a+b}\frac{b}{a+b}$.

The polynomial form of the diffusion coefficient a together with Theorem 3.2 implies several consequences on the moments of the probability distribution of $X \sim \mu$. We will present them in the following useful lemma. In particular, we give a recurrence formula for the moments of X.

Lemma 6.5. *Suppose that a is given by (6.1). Then for every $k \in \mathbb{R}$, $k \geq 1$ such that $\mathbf{E}[X^{2k}] < \infty$ one has*

$$\left(1 - \frac{2k-1}{2}\alpha\right)\mathbf{E}\left[X^{2k}\right] = \frac{2k-1}{2}\beta\mathbf{E}\left[X^{2k-1}\right] + \frac{2k-1}{2}\gamma\mathbf{E}\left[X^{2k-2}\right]. \tag{6.2}$$

In particular, if $\alpha \neq 2$,

$$\mathbf{E}\left[X^2\right] = \frac{\gamma}{2-\alpha}, \tag{6.3}$$

if $\alpha \neq 1, 2$,

$$\mathbf{E}\big[X^3\big] = \frac{\beta}{1-\alpha}\mathbf{E}\big[X^2\big] = \frac{\beta\gamma}{(1-\alpha)(2-\alpha)} \tag{6.4}$$

and if $\alpha \neq 2, \frac{2}{3}$, *then*

$$\mathbf{E}\big[X^4\big] = \frac{3(\frac{\beta^2}{1-\alpha}+\gamma)}{2-3\alpha}\mathbf{E}\big[X^2\big] = \frac{3\gamma(\frac{\beta^2}{1-\alpha}+\gamma)}{(2-\alpha)(2-3\alpha)}. \tag{6.5}$$

Proof. Relation (6.2) is obtained by applying Theorem 3.2 with $h(x) = x^{2k-1}$. In particular, for $k = 1$, since $\mathbf{E}[X] = 0$

$$\mathbf{E}\big[X^2\big] = \mathbf{E}\left[\frac{1}{2}a(X)\right] = \frac{1}{2}\big(\alpha\mathbf{E}\big[X^2\big] + \beta\mathbf{E}[X] + \gamma\big)$$

and thus $\mathbf{E}[X^2] = \frac{\gamma}{2-\alpha}$. By applying successively Proposition 3.2 with $h(x) = x^2$ $(k = \frac{3}{2})$ and $h(x) = x^3$ $(k = 2)$, we obtain the expressions (6.4) and (6.5). □

Lemma 6.6. *Assume* (6.1) *with* $\alpha \neq 1, 2, \frac{2}{3}$. *Fix* $n \geq 1$ *and let* $\{F_m = I_n(f_m), m \geq 1\}$ *such that*

$$\mathbf{E}\big[F_m^2\big] \underset{m\to\infty}{\to} \mathbf{E}\big[X^2\big], \qquad \mathbf{E}\big[F_m^4\big] \underset{m\to\infty}{\to} \mathbf{E}\big[X^4\big] \quad \textit{and} \quad \mathbf{E}\big[F_m^3\big] \underset{m\to\infty}{\to} \mathbf{E}\big[X^3\big]. \tag{6.6}$$

Then

$$\begin{aligned}
&\mathbf{E}\left[F_m^4 - \frac{3}{2}a(F_m)F_m^2\right] \\
&\quad \simeq C_0 n!\|f_m\|^2 + 3n \sum_{p=1;\, p\neq\frac{n}{2}}^{n-1} (p-1)!\binom{n-1}{p-1}^2 p!\binom{n}{p}^2 (2n-2p)!\|f_m \tilde{\otimes}_p f_m\|^2 \\
&\qquad + \frac{3}{2}c_n^{-2} n!\|f_m \tilde{\otimes}_{n/2} f_m\|^2 - \frac{3}{2}\beta\mathbf{E}\big[F_m^3\big] \underset{m\to\infty}{\to} 0,
\end{aligned} \tag{6.7}$$

where $\simeq$ *means that the two sides have the same limit as* $m \to \infty$, c_n *is defined by*

$$c_n = \frac{(n/2)!^3}{n!^2} \tag{6.8}$$

and

$$C_0 = \frac{3}{2}\alpha\left[\frac{-4\gamma}{(2-\alpha)(2-3\alpha)} - \frac{3\beta^2}{(1-\alpha)(2-3\alpha)}\right]. \tag{6.9}$$

Proof. From assumption (6.6) and Theorem 3.2,

$$\mathbf{E}\left[F_m^4 - \frac{3}{2}a(F_m)F_m^2\right] \underset{m}{\to} \mathbf{E}\left[X^4 - \frac{3}{2}a(X)X^2\right] = 0.$$

Now, by (6.1)

$$\mathbf{E}\left[F_m^4 - \frac{3}{2}a(F_m)F_m^2\right] = \mathbf{E}\big[F_m^4\big] - \frac{3}{2}\alpha\mathbf{E}\big[F_m^4\big] - \frac{3}{2}\gamma\mathbf{E}\big[F_m^2\big] - \frac{3}{2}\beta\mathbf{E}\big[F_m^3\big]$$

and, on the other hand, using (6.6), (6.5) and (2.5) we get

$$\begin{aligned}
&\left(1 - \frac{3}{2}\alpha\right)\mathbf{E}\big[F_m^4\big] - \frac{3}{2}\gamma\mathbf{E}\big[F_m^2\big] \\
&\quad \simeq 3\mathbf{E}\big[F_m^2\big]^2 + 3n\sum_{p=1}^{n-1}(p-1)!\binom{n-1}{p-1}^2 p!\binom{n}{p}^2 (2n-2p)!\|f_m \tilde{\otimes}_p f_m\|^2 \\
&\qquad - \frac{3}{2}\alpha\frac{3(\frac{\beta^2}{1-\alpha}+\gamma)}{2-3\alpha}\mathbf{E}\big[F_m^2\big] - \frac{3}{2}\gamma\mathbf{E}\big[F_m^2\big] \\
&\quad \simeq C_0\mathbf{E}\big[F_m^2\big] + 3n\sum_{p=1}^{n-1}(p-1)!\binom{n-1}{p-1}^2 p!\binom{n}{p}^2 (2n-2p)!\|f_m \tilde{\otimes}_p f_m\|^2
\end{aligned}$$

with C_0 given by (6.9). Therefore, the desired relation (6.7) is obtained. □

From the above result, we will deduce several restrictions on the probability distributions that can be limits of sequences of multiple stochastic integrals. We will discuss separately the cases $\beta = 0$ and $\beta \neq 0$.

6.1. The case $\beta = 0$

Let us assume that $\beta = 0$ in (6.1). This is the case of the Student, uniform and beta (with parameters $a = b$) distributions. In this situation, since $\mathbf{E}X^3 = 0$ (see (6.4)), the order of the chaos n can be even or odd in principle.

Theorem 6.7. *Assume* (6.1) *with* $\alpha \neq 2, \frac{2}{3}$ *and* $\beta = 0$. *Fix* $n \geq 1$ *and let* $(F_m = I_n(f_m))_{m\geq 1}$ *satisfying* (6.6). *Then* $\alpha = 0, \gamma > 0$ *and if* $S = (-\infty, \infty)$, X *follows a centered normal distribution with variance* γ.

Proof. Let us show that $C_0 \geq 0$. Note that

$$C_0 = -\frac{6\alpha\gamma}{(2-\alpha)(2-3\alpha)}.$$

Relations (6.3) and (6.5) with $\beta = 0$ give

$$\mathbf{E}\big[X^2\big] = \frac{\gamma}{2-\alpha} \quad \text{and} \quad \mathbf{E}\big[X^4\big] = \frac{3\gamma^2}{(2-\alpha)(2-3\alpha)}$$

and this implies in particular

$$(2-\alpha)(2-3\alpha) > 0.$$

On the other hand, from the relation (6.2) with $\beta = 0$ we notice that

$$\left(1 - \frac{2k-1}{2}\alpha\right)\mathbf{E}\left[X^{2k}\right] = \frac{2k-1}{2}\gamma\mathbf{E}\left[X^{2k-2}\right]$$

for every $k \geq 1$ we notice that α, γ have different parities. Indeed, if $\alpha > 0$ then for k large enough $(1 - \frac{2k-1}{2}\alpha)$ becomes negative and thus $\gamma < 0$. If $\alpha < 0$ then $(1 - \frac{2k-1}{2}\alpha)$ is positive and so γ should be positive. So $C_0 \geq 0$.

Next, since $\mathbf{E}[F_m^4 - \frac{3}{2}a(F_m)F_m^2] \to_m 0$ we have from Lemma 6.6

$$C_0\mathbf{E}\left[F_m^2\right] + 3n\sum_{p=1}^{n-1}(p-1)!\binom{n-1}{p-1}^2 p!\binom{n}{p}^2(2n-2p)!\|f_m \tilde{\otimes}_p f_m\|^2 \underset{m\to\infty}{\to} 0$$

and this is not possible unless $C_0 = 0$. In this case, we have $\alpha = 0$ (from (6.9)), $\gamma > 0$ (from (6.3)) and consequently, when $S = (-\infty, \infty)$, X follows a normal distribution with mean zero and variance γ. This can be easily obtained by computing the density of X using Proposition 1 in [5]. □

As a consequence, we notice that several probability distributions cannot be limits in distribution of sequences of multiple stochastic integrals.

Corollary 6.8. *A sequence of random variables in a fixed Wiener chaos cannot converge to the uniform distribution.*

6.2. The case $\beta \neq 0$

In this paragraph, we study the convergence of a sequence a multiple stochastic integrals to the law of a random variable $X \sim \mu$ where μ is the invariant measure of a diffusion with drift coefficient $b(x) = -x$ (meaning that $\mathbf{E}[X] = 0$) and diffusion coefficient a given by (6.1) with $\beta \neq 0$. This is the case of the Pareto, Gamma, inverse Gamma and F distributions.

Fix $n \geq 1$. Consider throughout this section that $\{F_m, m \geq 1\}$ is a sequence of random variables expressed as $F_m = I_n(f_m)$ with $f_m \in H^{\odot n}$. Since the third moment of a multiple Wiener–Itô integral of odd order is zero, from (6.4) we may assume in this paragraph that *n is even.*

Let us first deduce some consequences on the convergence in law of F_m to X.

Lemma 6.9. *Assume* (6.1) *with* $\alpha \neq 1, 2, \frac{2}{3}$. *Fix* $n \geq 1$ *and let* $\{F_m = I_n(f_m), m \geq 1\}$ *satisfying* (6.6). *Then*

$$\lim_m \langle f_m, f_m \tilde{\otimes}_{n/2} f_m\rangle = \lim_m \frac{\beta}{1-\alpha}c_n\|f_m\|^2,$$

where c_n *is given by* (6.8).

Proof. Condition (6.6) and Proposition 3.2 imply

$$\mathbf{E}\left[F_m^3 - a(F_m)F_m\right] \underset{m\to\infty}{\to} \mathbf{E}\left[X^3 - a(X)X\right] = 0$$

or equivalently

$$(1-\alpha)\mathbf{E}\big[F_m^3\big]-\beta\mathbf{E}\big[F_m^2\big]\underset{m}{\to}0.$$

But from (2.4), we have

$$\begin{aligned}(1-\alpha)\mathbf{E}\big[F_m^3\big]-\beta\mathbf{E}\big[F_m^2\big]&=(1-\alpha)\frac{n!^3}{(n/2)!^3}\langle f_m, f_m\tilde{\otimes}_{\frac{n}{2}} f_m\rangle-\beta n!\|f_m\|^2\\&=\frac{n!^3}{(n/2)!^3}\big[(1-\alpha)\langle f_m, f_m\tilde{\otimes}_{\frac{n}{2}} f_m\rangle-\beta c_n\|f_m\|^2\big]\end{aligned}$$

and since this converges to zero as $m\to\infty$, we obtain the conclusion. Let us also mention that the two limits in the statement of the lemma exist due to Lemma 2.1 and to the convergence of the sequences of the second and third moments of $(F_m)_{m\geq 1}$. □

We will need the following auxiliary lemma.

Lemma 6.10. *Let the assumptions stated in Lemma 6.9 prevail. Then for every $c\in\mathbb{R}$*

$$\begin{aligned}&\mathbf{E}\bigg[F_m^4-\frac{3}{2}a(F_m)F_m^2\bigg]\\&\quad\simeq\frac{3}{2}c_n^{-2}n!\bigg[\frac{2}{3}c_n^2\bigg(C_0-\frac{3}{2}(1-c)\frac{\beta^2}{1-\alpha}\bigg)\|f_m\|^2-\beta cc_n\langle f_m, f_m\tilde{\otimes}_{\frac{n}{2}} f_m\rangle+\|f_m\tilde{\otimes}_{n/2} f_m\|^2\bigg]\\&\qquad+3n\sum_{p=1;\,p\neq\frac{n}{2}}^{n-1}(p-1)!\binom{n-1}{p-1}^2p!\binom{n}{p}^2(2n-2p)!\|f_m\tilde{\otimes}_p f_m\|^2\\&\quad\underset{m\to\infty}{\to}0,\end{aligned}\tag{6.10}$$

where $\simeq$ means that the two sides have the same limit as $m\to\infty$, c_n, C_0 are defined by (6.8) and (6.9), respectively.

Proof. Recall the relation (6.7). For every $c\in\mathbb{R}$, we can write

$$\frac{3}{2}\beta\mathbf{E}\big[F_m^3\big]=\frac{3}{2}(1-c)\beta\mathbf{E}\big[F_m^3\big]+\frac{3}{2}c\beta\mathbf{E}\big[F_m^3\big]$$

and using (6.4) and the convergence of the moments of F_m to those of μ (relation (6.6)),

$$\begin{aligned}\frac{3}{2}(1-c)\beta\mathbf{E}\big[F_m^3\big]&\simeq\frac{3}{2}(1-c)\beta\frac{\beta}{1-\alpha}\mathbf{E}\big[X^2\big]\\&\simeq\frac{3}{2}(1-c)\beta\frac{\beta}{1-\alpha}\mathbf{E}\big[F_m^2\big]=\frac{3}{2}(1-c)\beta\frac{\beta}{1-\alpha}n!\|f_m\|^2.\end{aligned}\tag{6.11}$$

By combining (6.7) and (6.11), we obtain (6.10). □

The next step is to find $c \in \mathbb{R}$ such that

$$\frac{2}{3}\left(C_0 - \frac{3}{2}(1-c)\frac{\beta^2}{1-\alpha}\right) = A^2 \quad \text{and} \quad \beta c = 2A.$$

If such $c \in \mathbb{R}$ exists, then the sequence $\mathbf{E}[F_m^4 - \frac{3}{2}a(F_m)F_m^2]$ (which converges to zero as $m \to \infty$ due to (3.4)) will have the same limit as $\|Ac_n f_m - f_m \tilde{\otimes}_{n/2} f_m\|^2$ plus a positive term (this is consequence of (6.10)). The existence of $c \in \mathbb{R}$ is equivalent to the existence of a real solution to the second degree equation

$$3\beta^2 c^2 - \frac{12\beta^2}{1-\alpha}c - \left(8C_0 - \frac{12\beta^2}{1-\alpha}\right) = 0 \tag{6.12}$$

which has the following discriminant

$$\Delta = -\frac{144\alpha}{2-3\alpha}\left[\frac{\beta^2}{(1-\alpha)^2} + \frac{2\gamma}{2-\alpha}\right]. \tag{6.13}$$

Since $\frac{\beta^2}{(1-\alpha)^2} + \frac{2\gamma}{2-\alpha} > 0$ (see (6.3)) the sign of Δ depends on the sign of $\frac{\alpha}{2-3\alpha}$.

Theorem 6.11. *Assume* (6.1) *with* $\alpha \neq 1, 2, \frac{2}{3}$. *Fix* $n \geq 1$ *and let* $\{F_m = I_n(f_m), m \geq 1\}$ *satisfying* (6.6). *Moreover, let us assume* $\frac{\alpha}{2-3\alpha} \leq 0$ (*that is,* $\alpha \in \mathbb{R} \setminus (0, \frac{2}{3}]$). *Then* $\alpha = 0$ *and, if* $S = (-\frac{a}{\lambda}, \infty)$, *X follows a centered Gamma law* $\Gamma(a, \lambda) - \mathbf{E}[\Gamma(a, \lambda)]$ *where* $\beta = \frac{2}{\lambda}$, $\gamma = \frac{2a}{\lambda^2}$.

Proof. When $\alpha \notin [0, \frac{2}{3})$, then the discriminant Δ (6.13) is positive and the equation (6.12) admits two real solutions (that may coincide). Let us denote by $c_1, c_2 \in \mathbb{R}$ and by $A_i = \frac{\beta c_i}{2}$, $i = 1, 2$. From (6.10) we deduce that, for $i = 1, 2$

$$\begin{aligned}\mathbf{E}\left[F_m^4 - \frac{3}{2}a(F_m)F_m^2\right] &\simeq \frac{3}{2}c_n^{-2}n!\|A_i c_n f_m - f_m \tilde{\otimes}_{n/2} f_m\|^2 \\ &\quad + 3n \sum_{p=1;\, p\neq\frac{n}{2}}^{n-1} (p-1)!\binom{n-1}{p-1}^2 p!\binom{n}{p}^2 (2n-2p)!\|f_m \tilde{\otimes}_p f_m\|^2 \\ &\underset{m}{\to} 0\end{aligned}$$

and consequently

$$\begin{aligned}\|A_i c_n f_m - f_m \tilde{\otimes}_{n/2} f_m\| &\underset{m}{\to} 0 \quad \text{and} \\ \|f_m \otimes_p f_m\| &\underset{m\to\infty}{\to} 0 \qquad \text{for } p = 1, \ldots, n-1,\ p \neq \tfrac{n}{2},\end{aligned} \tag{6.14}$$

where c_n is given by (6.8). Relation (6.14) and Lemma 6.9 immediately imply $A_i = \frac{\beta}{1-\alpha}$ for $i = 1, 2$ and consequently the two solutions to (6.12) must coincide. The discriminant Δ (6.13)

then vanishes and that gives $\alpha = 0$. The fact that μ is a centered Gamma law follows from (6.1) with $\alpha = 0$ and $\beta \neq 0$ by computing the density of X in terms of the squared diffusion coefficient a, see Proposition 1 in [5]. □

Since in the case of the beta distribution $\alpha = \frac{-2}{a+b}$ and $a, b > 0$, we have the following corollary.

Corollary 6.12. *A sequence of random variables in a fixed Wiener chaos cannot converge to the beta distribution.*

Remark 6.13. In the case of the centered Gamma distribution, we obtain from the proof of Theorem 6.11: a sequence $(F_m = I_n(f_m))_{m\geq 1}$ such that $\mathbf{E}[F_m^2] \to_{m\to\infty} \frac{a}{\lambda^2}$ converges to the centered Gamma law $\Gamma(a,\lambda) - \mathbf{E}[\Gamma(a,\lambda)]$ if and only if the following assertions are satisfied:

- $\mathbf{E}[F_m^3] \to_m \frac{2a}{\lambda^3}$ and $\mathbf{E}[F_m^4] \to_m \frac{3a(a+2)}{\lambda^4}$;
- $\|\frac{2}{\lambda} c_n f_m - f_m \tilde{\otimes}_{n/2} f_m\| \to_m 0$ (recall that c_n is given by (6.8));
- $\frac{1}{\lambda} F_m^2 + \frac{a}{\lambda^2} - \frac{1}{n}\|DF_m\|_H^2$ converges to zero in $L^2(\Omega)$.

When $\lambda = \frac{1}{2}$ and $a = \frac{\nu}{2}$ we retrieve the results in [7], Theorem 1.2.

Acknowledgements

Seiichiro Kusuoka is supported by JSPS KAKENHI Grant Number 25800054. Ciprian A. Tudor is supported by the CNCS Grant PN-II-ID-PCCE-2011-2-0015.

References

[1] Azmoodeh, E., Peccati, G. and Poly, G. (2014). Convergence towards linear combinations of chi-squared random variables. Preprint.

[2] Bibby, B.M., Skovgaard, I.M. and Sørensen, M. (2005). Diffusion-type models with given marginal distribution and autocorrelation function. *Bernoulli* **11** 191–220. MR2132002

[3] Bourguin, S. and Tudor, C.A. (2011). Cramér theorem for gamma random variables. *Electron. Commun. Probab.* **16** 365–378. MR2819659

[4] Eden, R. and Viens, F. (2013). General upper and lower tail estimates using Malliavin calculus and Stein's equations. In *Seminar on Stochastic Analysis, Random Fields and Applications VII. Progress in Probability* **67** 55–84. Basel: Birkhäuser/Springer. MR3380095

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

[6] Nourdin, I. and Peccati, G. (2009). Stein's method on Wiener chaos. *Probab. Theory Related Fields* **145** 75–118. MR2520122

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

[8] Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality. Cambridge Tracts in Mathematics* **192**. Cambridge: Cambridge Univ. Press. MR2962301

[9] Nourdin, I., Peccati, G. and Swan, Y. (2014). Entropy and the fourth moment phenomenon. *J. Funct. Anal.* **266** 3170–3207. MR3158721

[10] Nourdin, I. and Poly, G. (2013). Convergence in total variation on Wiener chaos. *Stochastic Process. Appl.* **123** 651–674. MR3003367

[11] Nourdin, I. and Viens, F.G. (2009). Density formula and concentration inequalities with Malliavin calculus. *Electron. J. Probab.* **14** 2287–2309. MR2556018

[12] Nualart, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. *Probability and Its Applications* (*New York*). Berlin: Springer. MR2200233

[13] Nualart, D. and Ortiz-Latorre, S. (2008). Central limit theorems for multiple stochastic integrals and Malliavin calculus. *Stochastic Process. Appl.* **118** 614–628. MR2394845

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

[15] Peccati, G. and Tudor, C.A. (2005). Gaussian limits for vector-valued multiple stochastic integrals. In *Séminaire de Probabilités XXXVIII. Lecture Notes in Math.* **1857** 247–262. Berlin: Springer. MR2126978

[16] Tudor, C.A. (2013). *Analysis of Variations for Self-Similar Processes: A Stochastic Calculus Approach. Probability and Its Applications* (*New York*). Cham: Springer. MR3112799

[17] Tudor, C.A. (2014). Chaos expansion and asymptotic behavior of the Pareto distribution. *Statist. Probab. Lett.* **91** 62–68. MR3208117

[18] Üstünel, A.S. and Zakai, M. (1989). On independence and conditioning on Wiener space. *Ann. Probab.* **17** 1441–1453. MR1048936

Received November 2015 and revised May 2016