PhD-FSTC-2018-26
The Faculty of Sciences, Technology and Communication

DISSERTATION

Defence held on 28/03/2018 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN MATHEMATIQUES

by

Guangqu ZHENG
Born on 21 January 1989 in Wenzhou (China)

RECENT DEVELOPMENTS AROUND THE MALLIAVIN-STEIN APPROACH

—-FOURTH MOMENT PHENOMENA VIA EXCHANGEABLE PAIRS

Dissertation defence committee

Dr Ivan Nourdin, dissertation supervisor
Professor, Université du Luxembourg

Dr Giovanni Peccati, Vice Chairman
Professor, Université du Luxembourg

Dr Guillaume Poly, Referee
Maîtres de conférences, Université du Rennes, I

Dr Anton Thalmaier, Chairman
Professor, Université du Luxembourg

Dr Pierre Vallois, Referee
Professor, Université du Lorraine

To *my grandma* (*waìpó*) *for her silent love*

Acknowledgement

First and foremost, I would like to thank Ivan Nourdin for his supervision and support throughout my Ph.D life in Luxembourg. I first met Ivan in the old campus Kirchberg (as a M2 student from Orsay) and I have enjoyed the conversations and email-communications with Ivan since then.

I would also like to thank Giovanni Peccati, Guillaume Poly, Anton Thalmaier and Pierre Vallois for joining the jury of my defense, and I deeply appreciate their time for reading my manuscript. I thank my Monsieur Vallois (my academic grandpa) for his valuable career advice, and I thank Guillaume for proposing interesting problems to me; special thanks go to Giovanni, who has shared many ideas (in particular those around his conjecture) with me and introduced Nick Cave's music to me.

Three of my papers included in this thesis are the fruits of collaborations with Anna, Atef, Christian, Ivan, to whom I would like to express my deep gratitude and appreciation. These collaborations have produced more than papers. I thank Christian and Ivan for constant discussions that have substantially improved my understanding on Stein's method and Malliavin calculus.

I thank the Fondation Mathématique Jacques Hadamard for supporting my M2 study at Orsay and I also thank my teachers in Paris for their excellent courses and valuable advices, in particular, Thierry Bodineau, Nicolas Curien, Halim Doss, Maxime Fevrier, Jean-François Le Gall, José Trashorras and Marc Yor. Moreover, I cherish very much their generosity, which may be casual to them but for sure has an important impact on my "*trajectoire aléatoire*". I would like to thank Darko Volkov at Worcester Polytechnic Institute for his nice lectures on real analysis, which was my first math course and stimulated my appreciation of the abstract reasoning; I am very grateful to Stephan Sturm for his very kind support through a RAship during my study at WPI.

After my undergraduate study at Wuhan University, I have stayed at various places, where my friends and colleagues therein have given me nice memories. In particular, I thank Nguyenho Ho, Yiqing Li, Grigor Nika, Xiao Shen for many nice lunch conversations at Dragon Dynasty; I thank Linxiao Chen, Rui Chen, Henri Elad-Altman, Qun Wang, Bo Xia and Yilin Wang for many discussions, and I thank Leilei Qu for many coffee breaks together at Dauphine.

Here comes another important group of people that have more explicit connection to my Ph.D life. I thank colleagues from the mathematics unit, University of Luxembourg: they have contributed to very friendly, lively pictures in my mind about the old campus Kirchberg and the new campus Belval. These pictures will almost surely survive for a long time regardless of my bad memory. I thank colorful dots in these pictures: Anna, Andrea, Antoine, Christian, Diu, Federico, François, Gergo, Hugo, James, Lijuan (& little *Duoduo*), Mariagiulia, Massimo, Maurizia, Robert, Ronan, Simon, Liyi and Yannick. Among these • • •, Anna and Maurizia are the most shining, adorable ones, I thank them in my heart for their care and nice Italian food (Italians do it better...) In particular, my heartfelt big **THANKS** goes to the adorable $\text{nana}_{/\sim(1,2)}$, who shares various interests, passion with me and offers me a true friendship[1] despite of the difference and "disagreements". Ah!!! I almost forgot to mention our excellent secretaries Katharina Heil and Marie Leblanc, who are always in the background of these pictures, ensuring a nice environment. I thank them for their help and extremely efficient administrative support throughout my Ph.D.

[1]For sure, it is one of the surest paths out of the illusion of separation.

Lastly, I thank my parents and my younger sister for their support and love; I dedicate this thesis to my lovely grandma, her silent love is always immersed in her kind gestures: love you, waìpó.

Part I

This page is left blank.

Contents of Part I

This page is left blank.

Chapter 1

Introduction

Abstract

This thesis is about recent developments around the Malliavin-Stein approach, in particular, I present an unified and elementary approach for obtaining fourth moment bounds on chaoses. I split this thesis into two parts: **Part I** is written as a small survey to present a tranche of the theory and **Part II** collects all my papers written during my Ph.D in *chronological* order.

P1. Normal approximation and almost sure central limit theorem for non-symmetric Rademacher functionals, *Stoch. Process. Appl.* 127 (5), 2017, 1622–1636.

P2. Convergence of random oscillatory integrals in the presence of long-range dependence and application to homogenization, joint work with Atef Lechiheb, Ivan Nourdin and Ezedine Haouala. To appear in *Probab. Math. Stat.*.

P3. Exchangeable pairs on Wiener chaos, joint work with Ivan Nourdin. *submitted.*

P4. Fourth moment theorems on the Poisson space in any dimension, joint work with Christian Döbler and Anna Vidotto. *to appear in Electron. J. Probab.*

P5. A Peccati-Tudor type theorem for Rademacher chaoses. *submitted.*

When citing my own papers, I will use the labelings **[P1]-[P5]**.

The preparation of **Part I** is mainly based on my most recent papers: **[P3],[P4],[P5]**. My first published work **[P1]** that is closely related to **[P5]** will only be tangentially mentioned, while my second paper **[P2]** with A. Lechiheb, I. Nourdin and E. Haouala is isolated from these four articles, so will not be included in this part, and interested readers can refer to **Part II** of this thesis.

The so-called Malliavin-Stein approach is a tailor-made combination of Paul Malliavin's *stochastic calculus of variation* and Charles Stein's method of *probabilistic approximation*. One of the original purposes for the *tailors*[1] was to obtain the rate of convergence for the so-called fourth moment theorem[2] of D. Nualart and G. Peccati [76]. In the sequel, we use the abbreviation FMT(s) for "fourth moment theorem(s)".

In the first two sections within this introductory chapter, we will begin with a historical overview from Nualart and Peccati's FMT to the birth of the Malliavin-Stein approach. This Malliavin-Stein approach was originally extensively studied in the Gaussian setting and later found its discrete version, mainly on the Poisson space and Rademacher setting. We will come

[1] Tailors = Ivan Nourdin and Giovanni Peccati

[2] This FMT shall not be confused with Tao-Vu's fourth moment theorem for Wigner ensembles.

to this in Section 1.3. The last section is devoted to a brief summary of our new contributions, namely, an unified approach to prove the FMTs on both Gaussian space and Poisson space, as well as its extension to the Rademacher setting. More precisely, we'll prove the FMT on the Poisson space in full generality, thus improving results in [32]; and moreover, we'll provide a significant multivariate extension. This corresponds to **[P4]**. For the Rademacher setting, we will use our elementary approach to establish a multivariate fourth-moment type result and in particular give a simple proof of the univariate result proved recently in [30]. This corresponds to **[P5]**.

The rest of this thesis consists of two more chapters: we will collect relevant basics of Stein's method in Chapter 2 and we will present our new contributions in Chapter 3. Although we are not trying to write a very self-contained survey about the Malliavin-Stein approach[3], we provide enough heuristic ideas as well as pointers to literature to guide readers through this presumably short journey. We hope that we will fulfill this simple goal in the end.

1.1 Around the FMT on Gaussian Wiener chaos

Let us start with an example from Peccati and Yor's work [85] : let W be a standard real Brownian motion, then according to Jeulin's lemma[4],

$$\int_0^1 t^{-2} W_t^2 \, dt = +\infty \quad \text{almost surely.}$$

The authors of [85] studied the fluctuation of $F_\varepsilon = \int_\varepsilon^1 W_t^2 t^{-2}\, dt$, after normalization, as ε goes to zero. And they obtained that

$$\widetilde{F_\varepsilon} := \frac{F_\varepsilon - \mathbb{E}[F_\varepsilon]}{\sqrt{\mathbf{Var} F_\varepsilon}} \xrightarrow[\varepsilon\to 0]{law} \mathcal{N}(0,1)\,. \tag{1.1.1}$$

The proof is roughly sketched as follows:

- It is easy to compute that $\sigma_\varepsilon := \sqrt{\mathbf{Var} F_\varepsilon} = \sqrt{4\log(1/\varepsilon) + 4\varepsilon - 4}$.
- Using Itô's formula and stochastic Fubini's theorem (see *e.g.* [107]), one can rewrite $\widetilde{F_\varepsilon}$ as follows:

$$\begin{aligned}\widetilde{F_\varepsilon} &= \frac{2}{\sigma_\varepsilon}\int_\varepsilon^1 \left(\int_0^1 t^{-2}\mathbf{I}_{(s\le t)} W_s dW_s\right) dt = \frac{2}{\sigma_\varepsilon}\int_0^1 \left(\int_\varepsilon^1 t^{-2}\mathbf{I}_{(s\le t)}\, dt\right) W_s dW_s \\ &= \frac{2}{\sigma_\varepsilon}\int_0^1 \left(\frac{1}{\varepsilon\vee s} - 1\right) W_s\, dW_s = \int_0^1 \varphi_s^\varepsilon\, dW_s\ ,\end{aligned}$$

[3]Interested readers can refer to the tailors' blue monograph [67] as well as the constantly updated website [61].

[4]Here is a simple version from the exercise book [18] by Chaumont and Yor: let $\{R_t, t\in(0,1]\}$ be a jointly measurable process with values in $\mathbb{R}_+$ such that $0 < \mathbb{E}[R_1] < +\infty$ and $R_t \overset{(d)}{=} R_1$ for any $t\in(0,1]$, and let μ be a σ-finite measure on $(0,1]$. Then the following equivalence holds true: $\int_{(0,1]} R_t\,\mu(dt) < +\infty$ *almost surely* if and only if $\int_{(0,1]}\mu(dt) < +\infty$.

with $\varphi_s^\varepsilon := 2\sigma_\varepsilon^{-1}[(\varepsilon \vee s)^{-1} - 1]W_s$, $s \in [0,1]$. It is also clear that $M_t^\varepsilon := \int_0^t \varphi_s^\varepsilon \, dW_s$, $t \in \mathbb{R}_+$, is a square-integrable continuous martingale such that its quadratic variation (at infinity) $\langle M^\varepsilon \rangle_\infty = +\infty$ almost surely[5].

- According to the theorem of Dambis and Dubins-Schwarz (see *e.g.* [104, Chapter V]), there exists a standard real Brownian motion β^ε such that $M_t^\varepsilon = \beta^\varepsilon_{\langle M^\varepsilon \rangle_t}$, $t \in \mathbb{R}_+$. As β_1^ε is a standard Gaussian random variable, in order to show the asymptotic normality of $\widetilde{F_\varepsilon} = M_1^\varepsilon$, it is enough to prove $\langle M^\varepsilon \rangle_1$ is close to 1 in $L^2(\mathbb{P})$, in view of the Burkholder-Davis-Gundy inequality (see *e.g.* [104, Chapter IV]).
- It is easy to obtain $\|\langle M^\varepsilon \rangle_1 - 1\|_{L^2(\mathbb{P})} = O(1/\sigma_\varepsilon)$, as $\varepsilon \downarrow 0$. This finishes our sketchy proof.

For more details, one can refer to the original paper [85] as well as the recent survey [79].

Later, Nualart and Peccati studied another Gaussian functional

$$F_{H,\varepsilon} := \int_\varepsilon^1 \frac{(W_t^H)^2}{t^{1+2H}} \, dt \,, \quad (0 < \varepsilon < 1) \tag{1.1.2}$$

where W^H is a fractional Brownian motion with Hurst parameter $H \in (0,1)$. The F_ε above corresponds to the case where $H = 1/2$, that is, when W^H is a standard Brownian motion. According to Jeulin's lemma, $F_{H,0} = +\infty$ a.s., so it is natural to see whether the Gaussian fluctuation of the renormalized $F_{H,\varepsilon}$ occurs for general H. Starting from this motivating example, Nualart and Peccati published a very surprising result [76] in 2005, which is known as the "*fourth moment theorem*" nowadays.

Theorem 1.1.1 (D. Nualart and G. Peccati, 2005)**.** Fix an integer $p \geq 2$. Let

$$F_n \equiv I_p^W(f_n) := p! \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{p-1}} f_n(t_1, \ldots, t_p) \, dW_{t_p} \cdots dW_{t_1} \tag{1.1.3}$$

be a sequence of pth-order *multiple Wiener-Itô integrals*. Assume that the kernel $f_n \in L^2([0,1]^p, dx)$ is symmetric almost everywhere such that $p!\|f_n\|^2_{L^2([0,1]^p)} \to 1$, as $n \to +\infty$. Then the following statements are equivalent:

(i) F_n converges in law to a standard Gaussian, as $n \to +\infty$.

(ii) $\mathbb{E}[F_n^4]$ converges to 3, as $n \to +\infty$.

(iii) For each $r \in \{1, \ldots, p-1\}$, $\|f_n \otimes_r f_n\|_{L^2([0,1]^{2p-2r})}$ converges to 0, as $n \to +\infty$, where the *r-contraction* $\otimes_r$ is defined as follows:

$$\begin{aligned} &(f \otimes_r g)(x_1, \ldots, x_{p+q-2r}) \\ &= \int_{[0,1]^r} f(x_1, \ldots, x_{p-r}, s_1, \ldots, s_r) g(x_{p-r+1}, \ldots, x_{p+q-2r}, s_1, \ldots, s_r) ds_1 \ldots ds_r \,, \end{aligned} \tag{1.1.4}$$

for $f \in L^2([0,1]^p, dx)$, $g \in L^2([0,1]^q, dx)$ and[6] $0 \leq r \leq p \wedge q$, see also (1.2.3).

[5] One can first compute directly that $\sigma_\varepsilon^2 \langle M^\varepsilon \rangle_\infty = 4\int_0^\infty [(\varepsilon \vee s)^{-1} - 1]^2 W_s^2 \, ds \geq \int_2^\infty W_s^2 s^{-2} \, ds$, while applying Jeulin's lemma to the inverse-time Brownian motion, we get $\int_1^\infty W_s^2 s^{-2} \, ds = +\infty$ a.s., so we can conclude that $\langle M^\varepsilon \rangle_\infty = +\infty$ almost surely.

[6] Note $f \otimes_0 g = f \otimes g$ is the usual tensor product of two functions.

The name "fourth moment theorem" comes from the equivalence between (i) and (ii). The implication "(ii) $\Rightarrow$ (i)" is particularly surprising, as it simplifies to a great extent the usual way of proving the central limit theorem through the so-called *method of moments and cumulants*[7]. The original proof of FMT follows essentially the same arguments as for (1.1.1), that is, one sees F_n as the value of some continuous martingale $M^{(n)}$ at $t = 1$ and applies the random-time change:

$$M_t^{(n)} := p! \int_0^t \left(\int_0^{t_1} \cdots \int_0^{t_{p-1}} f_n(t_1, \dots, t_p)\, dW_{t_p} \cdots dW_{t_2} \right) dW_{t_1} = \beta^{(n)}_{\langle M^{(n)} \rangle_t} ,$$

where $\beta^{(n)}$ is some standard Brownian motion and $\langle M^{(n)} \rangle_\bullet$ denotes the quadratic variation process associated to $M^{(n)}$. At a handwaving level, if $\langle M^{(n)} \rangle_1$ is close to 1 in $L^2(\mathbb{P})$, then by Burkholder-Davis-Gundy inequality, $F_n = M_1^{(n)}$ is close to $\beta_1^{(n)}$, which is standard Gaussian. Computing $\mathbb{E}[\langle M^{(n)} \rangle_1^2]$ involves an extensive use of the product formula (1.2.9) for multiple Wiener-Itô integrals and this is where the contractions in (iii) come into the play, while after a lengthy computation of the fourth moment $\mathbb{E}[F_n^4]$ using the product formula (1.2.9), one gets the equivalence between (ii) and (iii). Making these arguments rigorous is essentially what was done in the original article.

By the isometry between any two real separable Hilbert spaces, the above FMT holds inside a general Gaussian Wiener chaos associated to an isonormal process, see [76]. Back to the Gaussian functional (1.1.2), one can treat $F_{H,\varepsilon} - \mathbb{E}[F_{H,\varepsilon}]$ as an element in the second Gaussian Wiener chaos; and according to the FMT, checking the Gaussian fluctuation reduces to computing the related contraction-norms, see [76, Section 3.1]. This application is just *a tip of the iceberg*. This univariate FMT together with its multivariate extension has produced a very efficient strategy for proving Gaussian limits on the Wiener space.

Let us first recall the multivariate FMT [84], established by Peccati and Tudor.

Theorem 1.1.2 (G. Peccati and C. A. Tudor, 2005)**.** Fix integers $d \geq 2$ and $1 \leq q_1 \leq \ldots \leq q_d$. Let $C = (C_{i,j}, 1 \leq i, j \leq d)$ be a $d \times d$ symmetric nonnegative definite matrix and $f_{n,i} \in L^2([0,1]^{q_i}, dx)$ be symmetric for any $n \geq 1$ and every $i \in [d] := \{1, \dots, d\}$. Assume that the d-dimensional random vectors

$$F^{(n)} = (F_1^{(n)}, \dots, F_d^{(n)})^T := (I_{q_1}^W(f_{n,1}), \dots, I_{q_d}^W(f_{n,d}))^T$$

satisfy

$$\lim_{n \to +\infty} \mathbb{E}[F_i^{(n)} F_j^{(n)}] = C_{i,j} \text{ for every } i, j \in [d].$$

Then, as $n \to +\infty$, the following assertions are **equivalent**:

(1) The vector $F^{(n)}$ converges in law to a d-dimensional Gaussian vector $\mathcal{N}(0, C)$;

(2) for every $i \in [d]$, $F_i^{(n)}$ converges in law to a Gaussian random variable $\mathcal{N}(0, C_{i,i})$;

(3) for every $i \in [d]$, $\mathbb{E}[(F_i^{(n)})^4] \to 3C_{i,i}^2$;

(4) for every $i \in [d]$ and each $1 \leq r \leq q_i - 1$, $\|f_{n,i} \otimes_r f_{n,i}\|_{L^2([0,1]^{2q_i - 2r})} \to 0$.

[7] See *e.g.* Section 30 in Billingsley's book [13].

In fact, soon after the appearance of [76], Peccati and Tudor attempted to obtain the FMT for a sequence of random variables belonging to finitely many chaoses. Surprisingly, they obtained the above more interesting result: for random vectors with components in the fixed Gaussian Wiener chaos, the joint convergence to Gaussian (1) is ***equivalent*** to the component-wise convergence (2), and the latter is reduced to verifying (3) or (4) in view of Nualart and Peccati's univariate FMT. This observation provides practitioners with an alternative to the semimartingale approach (see *e.g.* [42]) to prove the central limit theorems on the Wiener space; roughly speaking, one can first decompose the random variable into Gaussian Wiener chaoses and then prove central convergence on each chaos. See, for example, [2, 3, 25] for applications to zeros of random polynomials; [54, 55, 56, 81, 97] for applications to statistical physics. Again, we invite interested readers to visit the website [61] for many more works around the FMTs.

Besides Peccati-Tudor's theorem, there have been several important inputs around the FMTs, notably

(i) Nualart and Ortiz-Latorre [75] gave another proof of FMT based on Malliavin calculus, which paved the road for *Stein to meet Malliavin* [66], see Section 1.2;

(ii) Ledoux took another insightful point-of-view for the FMT in [51] and he treated the Gaussian Wiener chaos as an eigenspace of a symmetric self-adjoint Markov diffusion operator, and then he employed Gamma calculus to derive the FMT in this setting. Such a direction has been further enriched in the papers [4, 5, 17]. Originally Ledoux's spectral viewpoint was taken in the diffusive setting and unexpectedly, after slight adjustment to the discrete case, it turns out to be a crucial ingredient for the obtention of the FMTs on the Poisson space and in the Rademacher setting, see Section 1.4.

1.2 How Stein meets Malliavin?

This section tells the story of how Malliavin calculus was first combined with Stein's method. As mentioned before, we will first sketch Nualart and Ortiz-Latorre's methodological breakthrough, namely, linking the Malliavin operators to the study of limit theorems on the Wiener space. To illustrate better the ideas and also for the sake of later reference, let us now introduce some basics of Malliavin calculus and define the central object in this thesis, *Wiener chaos*. Readers, who have already been familiar with Malliavin calculus, could jump to Section 1.2.1.

The concept of chaos that we consider dates back to Norbert Wiener's 1938 paper [108], in which Wiener first introduced the notion of *multiple Wiener integral* calling it *polynomial chaos*. In 1947, Cameron and Martin [16] obtained the orthogonal development of nonlinear Brownian functionals in terms of *Hermite polynomials*[8], which motivated Itô's 1951 work [40].

In 1951, Itô modified the definition of multiple Wiener integral and made it more convenient for analysis in the point that the multiple Wiener integrals of different orders are orthogonal to each other, while Wiener's polynomial chaos does not possess such a property. Later in 1956, Itô

[8]Hermite polynomials are orthogonal polynomials with respect to the Gaussian measure and it is one of the central ingredients in the Gaussian analysis. The first few Hermite polynomials are given by $H_0(x) = 1$, $H_1(x) = x$, $H_2(x) = x^2 - 1$, and recursively, $H_{p+1}(x) := xH_p(x) - pH_{p-1}(x)$. Alternatively, one can define them via the *Rodrigues' formula* $H_p(x) = (-1)^p \exp(x^2/2)\frac{d^p}{dx^p}\exp(-x^2/2)$, $p \in \mathbb{N}$. See *e.g.* Section 1.4 in [67].

generalized ideas from his previous paper and defined the multiple Wiener integral based on an *independently scattered* random measure. Note that the class of independently scattered random measures includes Gaussian and Poisson random measures. In the following section, we will define the Gaussian and Poisson Wiener chaoses in an unified manner.

From now on, we fix a σ-finite measure space $(\mathcal{Z}, \mathscr{Z}, \mu)$ such that $\{x\} \in \mathscr{Z}$ with $\mu(\{x\}) = 0$ for any $x \in \mathcal{Z}$, that is, μ is nonatomic. And we define $\mathscr{Z}_\mu := \{B \in \mathscr{Z} : \mu(B) < \infty\}$. All random objects in this thesis are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 1.2.1. (i) $W = \{W(A), A \in \mathscr{Z}_\mu\}$ is said to be a Gaussian random measure[9] on $(\mathcal{Z}, \mathscr{Z}, \mu)$ if $W(A) \sim \mathcal{N}(0, \mu(A))$ for each $A \in \mathscr{Z}_\mu$ and for each finite sequence $A_1, \dots, A_m \in \mathscr{Z}_\mu$ of pairwise disjoint sets, the random variables $W(A_1), \dots, W(A_m)$ are independent.

(ii) $\eta = \{\eta(B), B \in \mathscr{Z}\}$ is said to be a Poisson random measure on $(\mathcal{Z}, \mathscr{Z}, \mu)$ if $\eta(B) \sim \mathrm{Poi}(\mu(B))$ for each[10] $B \in \mathscr{Z}$ and for each finite sequence $B_1, \dots, B_m \in \mathscr{Z}$ of pairwise disjoint sets, the random variables $\eta(B_1), \dots, \eta(B_m)$ are independent.

(iii) For $B \in \mathscr{Z}_\mu$, we also define $\widehat{\eta}(B) := \eta(B) - \mu(B)$ and denote by $\widehat{\eta} = \{\widehat{\eta}(B) : B \in \mathscr{Z}_\mu\}$ the *compensated Poisson random measure* associated with η.

The measure μ, appearing in Definition 1.2.1, is called the *intensity measure*. The introduction of Gaussian and Poisson random measures can be treated as the randomization of the underlying intensity measure space $(\mathcal{Z}, \mathscr{Z}, \mu)$, and they constitute the basic building blocks of modern probability theory.

Remark 1.2.1. In fact, one can define the Poisson random measure on a more general measure space. More precisely, if ν is a so-called s-finite measure[11] on $(\mathcal{Z}, \mathscr{Z})$, then there exists a probability space supporting random elements $a_1, a_2 \dots$ in $\mathcal{Z}$ and κ in $\mathbb{N}_0 \cup \{+\infty\}$ such that

$$\eta := \sum_{n=1}^{\kappa} \delta_{a_n} \tag{1.2.1}$$

is a Poisson random measure with intensity ν, see *e.g.* [49, Corollary 3.7].

Back to our particular setting, there exists a partition $\{B_j, j \in \mathbb{N}\}$ of $\mathcal{Z}$ such that $\mu_j := \mu|_{B_j}$, the restriction of μ on B_j, is a nonzero and finite measure for each $j \in \mathbb{N}$. Then on each B_j, one can construct, by extending the probability space if necessary, an i.i.d. sequence $\{X_{1,j}, X_{2,j}, \dots\}$ of random variables with values in B_j and $X_{1,j} \sim \mu(B_j)^{-1}\mu_j$, while one can obtain similarly an independent Poisson random variable κ_j with parameter $\mu(B_j)$, so that $\eta_j := \sum_{n=1}^{\kappa_j} \delta_{X_{n,j}}$ is a Poisson random measure on B_j with intensity μ_j. Therefore, an independent superposition of $\eta_j, j \in \mathbb{N}$ gives us the desired Poisson random measure η on the whole space, which has the representation as in (1.2.1). Note these discussions also justify the existence of Poisson random measures.

In what follows, we will define[12] Gaussian and Poisson Wiener chaos in the spirit of Itô's work. First let us introduce some notation.

[9] The existence of W is guaranteed by Kolmogorov's theorem.

[10] We follow the convention that $\mathrm{Poi}(+\infty) = +\infty$, which makes sense in view of the Laplace transform.

[11] *i.e.* ν is a countable sum of finite measures. In particular, if ν is σ-finite, then it is automatically s-finite.

[12] In fact, one could define Poisson Wiener chaos on the s-finite intensity measure space, see [49, Chapter 12]. We decide to consider only the σ-finite nonatomic case, mainly for the unification as well as the statement of our forthcoming transfer principle (Proposition 1.4.1), while losing such a slight generality makes no *big* difference.

Notation A. For $q \in \mathbb{N}_0 := \{0, 1, 2, \dots\}$, $L^2(\mu^q)$ is an abbreviation for $L^2(\mathcal{Z}^q, \mathscr{Z}^q, \mu^q)$ with the convention $L^2(\mu^0) = \mathbb{R}$. For $p \in \mathbb{N} := \{1, 2, \dots\}$, we denote by $\mathscr{E}_p$ the collection of simple functions of the form

$$f(z_1, \dots, z_p) = \sum_{i_1,\dots,i_p=1}^{m} \beta_{i_1,\dots,i_p} \mathbf{I}_{A_{i_1}\times\cdots\times A_{i_p}}(z_1, \dots, z_p) \tag{1.2.2}$$

where $m \in \mathbb{N}$, $A_1, \dots, A_m \in \mathscr{Z}_\mu$ are *pairwise disjoint*, and the coefficients $\beta_{i_1,\dots,i_p}$ vanish whenever any two of the indices $i_1, \dots, i_p$ are equal. It is well known that $\mathscr{E}_p$ is dense in $L^2(\mu^p)$, see [74, page 10] for a proof. We denote by $\mathfrak{S}_p$ the permutation group over $[p] := \{1, \dots, p\}$, and given $f \in L^2(\mu^p)$, we write $\widetilde{f}$ for the canonical symmetrization of f, that is,

$$\widetilde{f}(z_1, \dots, z_p) = \frac{1}{p!} \sum_{\sigma\in\mathfrak{S}_p} f(z_{\sigma(1)}, \dots, z_{\sigma(p)}) \,.$$

We write $L^2_s(\mu^p)$ for the set of functions $f \in L^2(\mu^p)$ satisfying $f = \widetilde{f}$, μ-almost everywhere. By the Cauchy-Schwarz inequality, it holds that $\|\widetilde{f}\|_{L^2(\mu^p)} \le \|f\|_{L^2(\mu^p)}$ for any $f \in L^2(\mu^p)$. To ease the notation, we just write $\|\cdot\|$ for the Euclidean norm or Hilbert-space norm. Lastly, we define the contractions similarly as in (1.1.4): given $f \in L^2(\mu^p)$ and $g \in L^2(\mu^q)$ $(p, q \in \mathbb{N})$, $f \otimes_r g \in L^2(\mu^{p+q-2r})$ is defined by

$$\begin{aligned}
&(f \otimes_r g)(z_1, \dots, z_{p+q-2r}) \\
&= \int_{\mathcal{Z}^r} f(z_1, \dots, z_{p-r}, s_1, \dots, s_r) g(z_{p-r+1}, \dots, z_{p+q-2r}, s_1, \dots, s_r) \mu(ds_1) \dots \mu(ds_r) \,,
\end{aligned} \tag{1.2.3}$$

for $0 \le r \le p \wedge q$, with the convention $f \otimes_0 g = f \otimes g$. We write $f \widetilde{\otimes}_r g$ for the canonical symmetrization of $f \otimes_r g$.

Now let $\mathfrak{X} \in \{W, \widehat{\eta}\}$, $p \in \mathbb{N}$ and $f \in \mathscr{E}_p$ have the form (1.2.2), we define

$$I_p^{\mathfrak{X}}(f) = \sum_{i_1,\dots,i_p=1}^{m} \beta_{i_1,\dots,i_p} \prod_{j=1}^{p} \mathfrak{X}(A_{i_j}) \,, \tag{1.2.4}$$

with the convention that $I_0^{\mathfrak{X}}(c) = c$ if $c \in L^2(\mu^0)$.

Note by definition, $I_p^{\mathfrak{X}}(f) = I_p^{\mathfrak{X}}(\widetilde{f})$ and the definition of $I_p^{\mathfrak{X}}(f)$ does not depend on the particular representation of f. And, it is easy to verify that for any $f \in \mathscr{E}_p$ and $g \in \mathscr{E}_q$, $(p, q \in \mathbb{N}_0)$

$$\mathbb{E}[I_p^{\mathfrak{X}}(f) I_q^{\mathfrak{X}}(g)] = \mathbf{I}_{\{p=q\}} p! \langle \widetilde{f}, \widetilde{g} \rangle \,, \tag{1.2.5}$$

which, in particular, implies the *isometry relation* when $p = q$. Such a relation enables the extension of $I_p^{\mathfrak{X}}$ to the whole domain $L^2(\mu^p)$ via standard continuity argument, thanks also to the fact that $\mathscr{E}_p$ is dense in $L^2(\mu^p)$. As a consequence, the orthogonality relation (1.2.5) holds true for general $f \in L^2(\mu^p)$ and $g \in L^2(\mu^q)$.

The above discussions give arise to the following definition.

Definition 1.2.2. (i) For $f \in L^2(\mu^p)$, $p \in \mathbb{N}$, we define $I_p^W(f)$ to be the pth order *multiple Wiener-Itô integral* of f with respect to W. We write $\mathbb{C}_p^W = \{I_p^W(f) : f \in L^2(\mu^p)\}$, and call it the pth ***Gaussian Wiener chaos***. As a convention, $\mathbb{C}_0^W = \mathbb{R}$.

(ii) For $f \in L^2(\mu^p)$, $p \in \mathbb{N}$, (with slight abuse of notation) we define $I_p^\eta(f) \equiv I_p^{\widehat{\eta}}(f)$ to be the pth order *multiple Wiener-Itô integral* of f with respect to $\widehat{\eta}$. We write $\mathbb{C}_p^\eta = \{I_p^\eta(f) : f \in L^2(\mu^p)\}$, and call it the pth ***Poisson Wiener chaos***. As a convention, $\mathbb{C}_0^\eta = \mathbb{R}$.

One reason for the importance of Wiener chaos is the *chaotic representation property* enjoyed by W and η:

$$L^2(\Omega, \sigma\{W\}, \mathbb{P}) = \bigoplus_{p\geq 0} \mathbb{C}_p^W \quad \text{and} \quad L^2(\Omega, \sigma\{\eta\}, \mathbb{P}) = \bigoplus_{q\geq 0} \mathbb{C}_q^\eta , \tag{1.2.6}$$

where $\oplus$ indicates an orthogonal sum in the Hilbert space $L^2(\Omega, \mathbb{P})$. One can also read (1.2.6) as follows: every $F \in L^2(\Omega, \sigma\{\mathfrak{X}\}, \mathbb{P})$, with $\mathfrak{X} \in \{W, \eta\}$, admits a unique chaotic decomposition

$$F = \mathbb{E}[F] + \sum_{p\geq 1} I_p^{\mathfrak{X}}(f_p) \ \text{ in } L^2(\mathbb{P}), \quad \text{with } f_p \in L_s^2(\mu^p) \text{ for each } p \in \mathbb{N}. \tag{1.2.7}$$

For the proofs, one can refer to [74, Theorem 1.1.2] and [49, Theorem 18.10].

From the above chaos decomposition, we can see clearly the correspondence between the $L^2(\mathbb{P})$-space and the *symmetric Boson-Fock space*

$$\mathbb{R} \oplus \bigoplus_{p\geq 1} L_s^2(\mu^p)$$

with the identification

$$L_s^2(\mu^p) \simeq p\text{th } \textit{symmetric} \text{ tensor product of } L^2(\mu) .$$

Then, following Nualart and Vives' approach [77], we can define Malliavin operators D, δ, L on Gaussian and Poisson spaces *simultaneously*. One can also refer to D. Nualart's monograph [74] and G. Last's survey [47] for more details.

Malliavin operators. In this paragraph, $\mathfrak{X}$ stands for W or η. Denote by $\mathbb{D}$ the set of $F \in L^2(\Omega, \sigma\{\mathfrak{X}\}, \mathbb{P})$ as in (1.2.7) verifying $\sum_{p\geq 1} pp!\|f_p\|^2 < +\infty$, and for such a $F \in \mathbb{D}$, we define the *Malliavin derivative DF* of F as a random element in $L^2(\mu)$ by

$$D_z F = \sum_{p\geq 1} p I_{p-1}^{\mathfrak{X}}(f_p(z, \cdot)) , \quad z \in \mathcal{Z} .$$

Then it is easy to check by (1.2.5) that $\mathbb{E}[\|DF\|^2] = \sum_{p\geq 1} pp!\|f_p\|^2$ is finite, so we call $\mathbb{D}$ the domain for the Malliavin derivative D.

Now we define the *Skorohod integral* δ via the following duality relation, also known as the integration by parts formula:

$$\mathbb{E}[\langle DF, u\rangle] = \mathbb{E}[F\delta(u)] \ , \quad \text{for any } F \in \mathbb{D} \text{ and } u \in \mathrm{dom}(\delta), \tag{1.2.8}$$

where $\mathrm{dom}(\delta)$ is the set of $L^2(\mu)$-valued random variables u satisfying that $\mathbb{E}[\|u\|^2] < +\infty$ and such that there exists a finite constant c_u only depending on u such that $|\mathbb{E}[\langle DF, u\rangle]| \leq c_u \sqrt{\mathbb{E}[F^2]}$ for every $F \in \mathbb{D}$. In particular, one can check that given

$$u = (u_z, z \in \mathcal{Z}) \ : \ u_z = \sum_{p\geq 1} I^{\mathfrak{X}}_{p-1}(f_p(z,\cdot)) \ \text{ with } f_p \in L^2(\mu^p),\, f_p(z,\cdot) \in L^2_s(\mu^{p-1}),\, \forall p \in \mathbb{N},\, \forall z \in \mathcal{Z},$$

then $u \in \mathrm{dom}(\delta)$ if and only if $\sum_{p\geq 1} p!\|\widetilde{f_p}\|^2 < +\infty$, and in this case, $\delta(u) = \sum_{p\geq 1} I^{\mathfrak{X}}_p(\widetilde{f_p})$.

The operator L, known as the *infinitesimal generator of the Ornstein-Uhlenbeck semigroup*, is defined on the chaos expansion (1.2.6) as

$$L = \sum_{p\geq 0} -pJ_p \ , \text{where } J_p \text{ denotes the projection operator onto } \mathbb{C}^{\mathfrak{X}}_p.$$

Its domain $\mathrm{dom}(L)$ consists of random variables $F \in L^2(\Omega, \sigma\{\mathfrak{X}\}, \mathbb{P})$ satisfying $\sum_{p\geq 1} p^2 \mathbb{E}[J_p(F)^2] < +\infty$. One can check that $F \in \mathrm{dom}(L)$ if and only if $F \in \mathbb{D}$ and $DF \in \mathrm{dom}(\delta)$, and in this case $L = -\delta D$. We also define the *pseudo-inverse* L^{-1} of the operator L as follows: for $F \in L^2(\Omega, \sigma\{\mathfrak{X}\}, \mathbb{P})$ with zero mean, $L^{-1}F := \sum_{p\geq 1} -p^{-1}J_p(F)$. Note that $LL^{-1}(F - \mathbb{E}[F]) = F - \mathbb{E}[F]$ for any $F \in L^2(\Omega, \sigma\{\mathfrak{X}\}, \mathbb{P})$, which explains the name of L^{-1}.

Although the Gaussian space and the Poisson space share the same chaotic structure, their natures are essentially different by virtue of their underlying randomness. Notably,

(1) on the Gaussian space, the Malliavin derivative verifies the following chain rule: for $\phi : \mathbb{R} \to \mathbb{R}$ differentiable and Lipschitz, $F \in \mathbb{D}$, we have $\phi(F) \in \mathbb{D}$ and

$$D\phi(F) = \phi'(F)DF \ ;$$

(2) finitely many Gaussian Wiener chaoses are stable under multiplication: if $f \in L^2_s(\mu^p)$, $g \in L^2_s(\mu^q)$ $(p, q \in \mathbb{N})$, then[13]

$$I^W_p(f)I^W_q(g) = \sum_{r=0}^{p\wedge q} r!\binom{p}{r}\binom{q}{r} I^W_{p+q-2r}(f \otimes_r g) \, . \tag{1.2.9}$$

As a consequence, any random variable in a Gaussian Wiener chaos admit finite moments of all order, and all $L^r(\mathbb{P})$-norms, $r > 1$ are equivalent[14]. Note these are also immediate consequences of Nelson's *hypercontractivity property* for Ornstein-Uhlenbeck semigroups in the Gaussian setting, see *e.g.* [74, Section 1.4.3]. Due to the inherent discreteness, the above neat chain rule and product formula do not hold on the Poisson space, which partially contributes to the difficulty of obtaining FMTs therein.

With the above preliminary knowledge, we are in a position to discuss about Nualart and Ortiz-Latorre's work [75].

[13]For a nice heuristic proof using the concept of diagonal measure, see Theorem 9 in [62]. See also footnote 24 for an analogous proof in the Rademacher setting.

[14]These two facts can be proved by induction and iteratively using (1.2.9)

1.2.1 A methodological breakthrough by Nualart and Ortiz-Latorre

In 2008, Nualart and Ortiz-Latorre provided in [75] another equivalent condition for the FMT. More importantly, this paper contains a methodological breakthrough: the authors used Malliavin calculus tools, in particular the integration by parts formula on the Wiener space.

Theorem 1.2.1 (D. Nualart and S. Ortiz-Latorre, 2008)**.** Let F_n be given as in (1.1.3) and have unit variance, then F_n converges in law to $\mathcal{N}(0,1)$ if and only if

(iv) $\|DF_n\|^2$ converges in $L^2(\mathbb{P})$ to p, as $n \to +\infty$.

Note $\|DF_n\|^2$ is somehow a reminiscence of the quadratic variation $\langle M^{(n)}\rangle_1$ from the random-time change, see page 4. Here we will just sketch the argument for the "if-part":

> By assumption, $F_n = p^{-1}\delta DF_n$ is bounded in $L^2(\mathbb{P})$, hence admitting a subsequence F_{n_k} that converges in law to some random variable Z. It follows that $\mathbb{E}[e^{\mathrm{i}tF_{n_k}}F_{n_k}] = p^{-1}\mathbb{E}[e^{\mathrm{i}tF_{n_k}}\delta DF_{n_k}] = p^{-1}\mathbb{E}[\langle De^{\mathrm{i}tF_{n_k}}, DF_{n_k}\rangle] = p^{-1}\mathbb{E}[\mathrm{i}te^{\mathrm{i}tF_{n_k}}\langle DF_{n_k}, DF_{n_k}\rangle]$. If (iv) holds, then we have $\mathbb{E}[e^{\mathrm{i}tZ}Z] = \mathbb{E}[\mathrm{i}te^{\mathrm{i}tZ}]$. Denote by ϕ_Z the characteristic function of Z, then the previous equality suggests that $\phi_Z'(t) = -t\phi_Z(t)$, subject to the boundary condition $\phi_Z(0) = 1$. Solving this ODE gives us $Z \sim \mathcal{N}(0,1)$. This argument can be applied to any subsequence of (F_n), hence the original sequence indeed converges in law to $\mathcal{N}(0,1)$.

For the "only-if" part, one can first rewrite $\mathbb{E}[(\|DF_n\|^2 - p)^2] = \mathbb{E}[\|DF_n\|^4] - p^2$, which can be further expressed in terms of contractions, using the product formula (1.2.9). By condition (iii) in Nualart-Peccati's FMT, these contractions vanish asymptotically so that one can obtain the desired $L^2(\mathbb{P})$-convergence.

Remark 1.2.2. In 2009, Nourdin and Peccati [65] used exclusively results from Malliavin calculus to derive non-central extensions of the FMT. One of their main results states that for any sequence $\{f_k, k \geq 1\} \subset L^2_s(\mu^p)$ ($p \geq 2$ *even*) verifying $\lim_{k\to+\infty} p!\|f_k\|^2 = 2\nu > 0$, the following statements are equivalent, as $k \to +\infty$:

(a) $I_p^W(f_k)$ converges in law to $F(\nu)$, centered Gamma distribution[15] with parameter ν;

(b) $\mathbb{E}[I_p^W(f_k)^4] - 12\mathbb{E}[I_p^W(f_k)^3] \longrightarrow 12\nu^2 - 48\nu$;

(c) $\|DI_p^W(f_k)\|^2 - 2pI_p^W(f_k)$ converges in $L^2(\mathbb{P})$ to $2p\nu$.

This work exhibits again the power of Malliavin calculus.

A digression. The readers may have wondered "*why the fourth moment is so special for the asymptotic normality inside a fixed Gaussian Wiener chaos*". One immediate answer could be Theorem 1.2.1 and the following consequence of the product formula (1.2.9): for $F \in \mathbb{C}_q^W$,

$$\mathbf{Var}(q^{-1}\|DF\|^2) \leq \frac{q-1}{3q}(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2) \leq (q-1)\mathbf{Var}(q^{-1}\|DF\|^2)\,, \tag{1.2.10}$$

see equation (5.2.7) in [67].

[15] *i.e.* $F(\nu) = 2\gamma_{\nu/2} - \nu$, where $\gamma_{\nu/2}$ has a Gamma law with parameter $\nu/2$.

One can also answer this question using the classic method of moments and cumulants, as Nourdin explained in the short note [63]: for a normalized sequence $\{F_n, n \geq 1\} \subset \mathbb{C}_p^W$ such that $\mathbb{E}[F_n^4] \to 3$, one can prove, with some soft combinatorial arguments and the product formula, that every moment of F_n converges to the corresponding Gaussian moment. Here, let us provide a slightly different argument: as every monomial is a finite linear combination of Hermite polynomials[16], it suffices instead to show that $\mathbb{E}[H_k(F_n)] \to \mathbb{E}[H_k(Z)] = 0$ for[17] every $k \in \mathbb{N}$, $Z \sim \mathcal{N}(0,1)$. Note the expression $\mathbb{E}[H_k(X)]$ bears the particular name "*kth normal moment of X*" and is closely related to the χ^2-distance of Pearson, see [14, Proposition 8.2]. The first few normal moments of F_n are $\mathbb{E}[H_1(F_n)] = \mathbb{E}[F_n] = 0$, $\mathbb{E}[H_2(F_n)] = \mathbb{E}[F_n^2] - 1 = 0$, $\mathbb{E}[H_4(F_n)] = \mathbb{E}[F_n^4 - 6F_n^2 + 3] = \mathbb{E}[F_n^4] - 3$; for a generic $k \in \mathbb{N}$, we write $\mathbb{E}[H_{k+1}(F_n)] = \mathbb{E}[F_n H_k(F_n)] - \mathbb{E}[kH_{k-1}(F_n)]$ using the recursive relation of Hermite polynomials, and then we apply the integration by parts formula and chain rule as follows:

$$\mathbb{E}[F_n H_k(F_n)] = p^{-1}\mathbb{E}[\delta D F_n H_k(F_n)] = p^{-1}\mathbb{E}[\langle DF_n, DH_k(F_n)\rangle] = \mathbb{E}[p^{-1}\|DF_n\|^2 k H_{k-1}(F_n)]\,,$$

thus,

$$\left|\mathbb{E}[H_{k+1}(F_n)]\right| = \left|\mathbb{E}[(p^{-1}\|DF_n\|^2 - 1)kH_{k-1}(F_n)]\right| \leq k\sqrt{\mathbf{Var}(p^{-1}\|DF_n\|^2)}\sqrt{\mathbb{E}[H_{k-1}(F_n)^2]}\,;$$

see also (1.2.11) for similar computations. The above inequality, together with (1.2.10), implies that the fourth moment controls other moments asymptotically.

Lastly, let us mention Rosiński's *intuitive* explanation [95] based on the independence on Wiener space. Let us first recall the seminal result of Üstünel and Zakai [106] that provides the necessary and sufficient condition for independence of two multiple Wiener-Itô integrals with respect to the same Gaussian random measure.

Theorem 1.2.2 (A.S. Üstünel and M. Zakai, 1989)**.** Fix $p, q \in \mathbb{N}$ and $f \in L_s^2(\mu^p)$, $g \in L_s^2(\mu^q)$, we set $F = I_p^W(f)$, $G = I_q^W(g)$. Then F, G are independent if and only if $f \otimes_1 g = 0$.

Later, Rosiński and Samorodnitsky [96] made another interesting observation.

Proposition 1.2.1 (J. Rosiński and G. Samorodnitsky, 1999)**.** Fix $p, q \in \mathbb{N}$, then $\mathbf{Cov}(F^2, G^2) \geq 0$ for any $F \in \mathbb{C}_p^W$, $G \in \mathbb{C}_q^W$, with equality only when F and G are independent.

Here comes Rosiński's nice heuristic arguments: suppose that $F \in \mathbb{C}_p^W$ satisfies $\mathbb{E}[F^2] = 1$ and[18] $\mathbb{E}[F^4] = 3$, let G be an independent copy of F, then $\mathbf{Cov}((F+G)^2, (F-G)^2) = 2\mathbb{E}[F^4] - 6\mathbb{E}[F^2]^2 = 0$, so it follows from Proposition 1.2.1 that $F + G$ is independent of $F - G$. Thus, by Bernstein's theorem, $F \sim \mathcal{N}(0,1)$. In the statement of the FMT, $\mathbb{E}[F_n^4] = 3$ only asymptotically, it is natural to expect $F_n \to \mathcal{N}(0,1)$. Such heuristics have led to several papers on asymptotic independence on the Wiener space [72, 64], with applications to time series analysis [6].

[16] By definition (see footnote 8), Hermite polynomial is a finite linear combination of monomials, then inductively, one can easily verify that every monomial can be expressed as a finite linear combination of Hermite polynomials.

[17] Like to crack a walnut with a sledgehammer: Z is equal in law to $W(h)$ for some unit vector $h \in L^2(\mu)$, and Cameron-Martin's result [16] suggests that $H_k(W(h)) = I_k^W(h^{\otimes k})$, which is an element of kth Gaussian Wiener chaos, and this explains $\mathbb{E}[H_k(Z)] = 0$ for every $k \in \mathbb{N}$.

[18] In fact, if $p \geq 2$, $\mathbb{E}[F^4] > 3$, see *e.g.* [76, Corollary 2] or our Chapter 3. So we view $\mathbb{E}[F^4] = 3$ as an asymptotic relation.

The above ***digression*** around the importance of the fourth moment may also trigger one's desire of quantifying the CLT on Wiener chaos in terms of the fourth moment. It is known, for example from [67, Proposition C.3.2], that a sequence of random variables X_n converges in law to the standard Gaussian N if and only if $d_{\text{Kol}}(X_n, N) \to 0$, where d_{Kol} denotes the Kolmogorov distance[19] appearing in the celebrated Berry-Esséen bound. So is it true that

$$(\#) \qquad d_{\text{Kol}}(F, \mathcal{N}(0, \mathbf{Var}(F))) \le \text{Constant} \times \sqrt{\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2} \text{ for any } F \in \mathbb{C}_p^W \ ?$$

The answer is indeed positive and will be the highlight of the story told in the next section. It is also worth pointing out the paper [5] by Azmoodeh, Malicet, Mijoule and Poly, who provided a generalization of FMT by establishing the equivalence between asymptotic normality and convergence of *even* moments on the Gaussian Wiener chaos; see also (1.4.1).

1.2.2 Nourdin-Peccati analysis

Motivated by the question (#) on the previous section, Nourdin and Peccati took their investigation in search of a quantitative FMT. Their investigation closed with the *somehow* unexpected birth of a new research line, the so-called *Malliavin-Stein approach*[20]. This approach is a tailor-made combination of Paul Malliavin's differential calculus and Charles Stein's method of distributional approximation.

Malliavin's original purpose of initiating his differential calculus was to provide a probabilistic proof for Hörmander's famous "sum of squares" theorem [53]. Since then, Malliavin's theory has been further developed by Malliavin himself, Stroock, Bismut, Watanabe and others. The integration by parts formula lies at the center of this theory. We refer the interested readers to Nualart's monograph [74] for the theory as well as many nice applications. On the other side, Stein's method began with Charles Stein's approach in the sixties to prove Wald-Wolfowitz and Hoeffding's combinatorial central limit theorems; see [24, Section 2.1] for more historical account. Stein's method of normal approximation first appeared in the groundbreaking 1972 paper [103], one of whose fundamental ingredients is the so-called *Stein's lemma* (see Lemma 2.1.1), which is nothing else but a particular case of the integration by parts formula (1.2.8). Since 1972, Stein's method has been ramified and developed by Stein himself and many other mathematicians, for example, A. Barbour, L.H.Y. Chen, P. Diaconis, F. Götze as well as their collaborators and students. We refer the interested readers to Stein's monograph [101] and the more recent book [19] by Chen, Goldstein and Shao, as well as two volumes of Barbour and Chen [11, 12] and Ross' survey [105].

Noticing the existence of integration by parts formulae on both sides of Malliavin calculus and Stein's method, one may conjecture some link between these two fields. I. Nourdin and G. Peccati's investigation shed some new light onto these two fields, by "*steining*"[21] Nualart-Peccati's FMT.

[19] $d_{\text{Kol}}(X, N) := \sup\{|\mathbb{P}(X \le t) - \mathbb{P}(N \le t)| \,:\, t \in \mathbb{R}\}$, see also Section 2.1.1.

[20] As a side note, it is in Airault, Malliavin and Viens' paper [1] that the term "Nourdin-Peccati analysis" was first coined. This explains our title of this subsection.

[21] This is a nice "abuse" of language, following Goldstein, Nourdin and Peccati's paper [37].

Let us briefly illustrate how Stein's method fits into Malliavin calculus and produces the bound in (#):

(i) Stein's heuristic suggests that a real random variable W is close in law to a standard Gaussian N if and only if $\mathbb{E}[f'(W)] \approx \mathbb{E}[Wf(W)]$ for sufficiently nice functions f. Then Stein built his equation $f'(x) - xf(x) = h(x) - \mathbb{E}[h(N)]$ and solved $f := f_h$ for the unknown h. For example, writing $f_z = f_h$ when $h(x) = \mathbf{I}_{\{x\leq z\}}$, one can obtain, after careful analysis, some uniform control of Stein's solution f_z, namely, $0 < f_z(x) \leq \sqrt{\pi/2}$ and $|f_z'(x)| \leq 2$, for every $x \in \mathbb{R}$, see *e.g.* Lemma 2.3 in [19].

(ii) If one replaces the dummy variable x by W and takes expectations at both sides of Stein's equation, it follows from point (i) that $d_{\mathrm{Kol}}(W, N) \leq \sup\{|\mathbb{E}[f'(W) - Wf(W)]| \,:\, \|f\|_\infty \leq \sqrt{\pi/2}\,,\ \|f'\|_\infty \leq 2\,\}$.

(iii) Now let us consider a normalized random variable $F \in \mathbb{C}_q^W$ ($q \in \mathbb{N}$), then by the integration by parts formula (1.2.8) and chain rule, one gets for f Lipschitz and differentiable,

$$\mathbb{E}[f(F)F] = q^{-1}\mathbb{E}[f(F)\delta DF] = q^{-1}\mathbb{E}[\langle Df(F), DF\rangle] = q^{-1}\mathbb{E}[\|DF\|^2 f'(F)]\ . \tag{1.2.11}$$

It follows from point (ii) that $d_{\mathrm{Kol}}(F, N) \leq 2\mathbb{E}[|q^{-1}\|DF\|^2 - 1|] \leq 2\sqrt{\mathbf{Var}(q^{-1}\|DF\|^2)}$, so the bound in (#) is implied by (1.2.10). Moreover, one can also get the fourth moment bound in other distance, like the total-variation, Wasserstein and Fortet-Mourier distances. These bounds are commonly called the *Nourdin-Peccati bounds* nowadays. Following the same computation as in (1.2.11), we can obtain for any normalized $F \in \mathbb{D}$ that $d_{\mathrm{Kol}}(F, N) \leq 2\sqrt{\mathbf{Var}(\langle DF, -DL^{-1}F\rangle)}$.

The above illustration basically tells one of the core ideas in the Malliavin-Stein approach, for more general results and a systematic look, one can refer to Nourdin-Peccati's original article [66] and their monograph [67]. This active line of research has been growing quite fast and led to many interesting results, here we just name two aspects that are related to this thesis:

(α) Nourdin, Peccati and Reinert [71] established the universality of Gaussian Wiener chaos using additionally the *Lindeberg invariance principle* from [59]. This motivated several works on the universality of multilinear homogeneous sums involving general independent random variables, see *e.g.* [7, 86, 68];

(β) Since the two papers [70, 82], the Malliavin-Stein approach has been successfully extended to the discrete settings, that is, Stein's method is coupled with suitable versions of discrete Malliavin calculus in order to yield quantitative limit theorems for Poisson functionals and Rademacher functionals over a Poisson random measure and a sequence of independent Bernoulli random variables, respectively. This discrete Malliavin-Stein approach has yielded nice applications to random graphs [44, 45, 46], and stochastic geometry [99, 91, 48], to name a few.

1.3 Discrete Malliavin-Stein approach

Let us begin this section by first addressing more on the above point (α) on the previous page. First, let us explain the term "*multilinear homogeneous sums*" (homogeneous sums in the sequel), after introducing some notation.

1.3.1 Homogeneous sums and Rademacher chaos

Notation B. We write $\mathfrak{H} = \ell^2(\mathbb{N})$, equipped with usual ℓ^2-norm $\|\cdot\|$ and for $p \in \mathbb{N}$, $\mathfrak{H}^{\otimes p}$ means the p-th tensor product of $\mathfrak{H}$ and $\mathfrak{H}^{\odot p}$ its symmetric subspace. We define[22] $\mathfrak{H}_0^{\odot p} := \{f \in \mathfrak{H}^{\odot p} : f|_{\triangle_p^c} = 0\}$ with $\triangle_p = \{(i_1, \dots, i_p) \in \mathbb{N}^p : i_k \neq i_j \text{ for different } k, j\}$.

Let us introduce an important notion concerning the homogeneous sums: for a given kernel $f \in \mathfrak{H}_0^{\odot d}$, we define the *maximal influence* $\mathcal{M}(f)$ of f as follows:

$$(1.3.1) \quad \mathcal{M}(f) := \sup_{k\in\mathbb{N}} \sum_{i_1,\dots,i_{d-1}\in\mathbb{N}} f(i_1, \dots, i_{d-1}, k)^2 \text{ for } d \geq 2 \text{ and } \mathcal{M}(f) := \sup_{k\in\mathbb{N}} f(k)^2 \text{ for } d = 1.$$

This notion is adapted from the boolean analysis (see *e.g.* [78]), in which the class of low-influence functions is often what is interesting or necessary in practice. It is also closely related to the aforementioned Lindeberg invariance principle.

Lastly, we introduce the star-type contractions, similar to the contractions in (1.2.3): for $p, q \in \mathbb{N}$, $0 \leq r \leq \ell \leq p \wedge q$, the (r, ℓ)-star contractions $f \star_\ell^r g$ of (a) $f \in L^2(\mu^p)$, $g \in L^2(\mu^q)$; (b) $f \in \mathfrak{H}^{\otimes p}$, $g \in \mathfrak{H}^{\otimes q}$ are defined by

(a) $(f \star_\ell^r g)(z_1, \dots, z_{p-\ell}, x_1, \dots, x_{\ell-r}, z_{p-\ell+1}, \dots, z_{p+q-2\ell})$

$$:= \int_{\mathcal{Z}^r} f(z_1, \dots, z_{p-\ell}, x_1, \dots, x_{\ell-r}, s_1, \dots, s_r) \times g(z_{p-\ell+1}, \dots, z_{p+q-2\ell}, x_1, \dots, x_{\ell-r}, s_1, \dots, s_r)\mu(ds_1)\dots\mu(ds_r),$$

with $f \star_\ell^\ell g = f \otimes_\ell g$ and $(f \star_\ell^0 g)(z_1, \dots, z_{p-\ell}, x_1, \dots, x_\ell, z_{p-\ell+1}, \dots, z_{p+q-2\ell}) = f(z_1, \dots, z_{p-\ell}, x_1, \dots, x_\ell) \times g(z_{p-\ell+1}, \dots, z_{p+q-2\ell}, x_1, \dots, x_\ell)$;

(b) $(f \star_\ell^r g)(z_1, \dots, z_{p-\ell}, x_1, \dots, x_{\ell-r}, z_{p-\ell+1}, \dots, z_{p+q-2\ell})$

$$:= \sum_{s_1,\dots s_r\in\mathbb{N}} f(z_1, \dots, z_{p-\ell}, x_1, \dots, x_{\ell-r}, s_1, \dots, s_r) g(z_{p-\ell+1}, \dots, z_{p+q-2\ell}, x_1, \dots, x_{\ell-r}, s_1, \dots, s_r)$$

with $f \star_\ell^\ell g = f \otimes_\ell g \in \mathfrak{H}^{\otimes p+q-2r}$ and $f \star_\ell^0 g$ defined as in (a).

In both settings, $f \star_\ell^r g$ is not square-integrable for general $r < \ell$ and the symmetry of f and g is *not* preserved under star-contraction operations. These star-contractions appear naturally in the product formulae on Poisson space ([47, Proposition 5]) and in the Rademacher setting ([43, Proposition 2.2]).

[22] As a convention, $\mathfrak{H}_0^{\odot 0} = \mathfrak{H}^{\otimes 0} = \mathbb{R}$ and $\mathfrak{H}_0^{\odot 1} = \mathfrak{H}$.

Now we give the definition of homogeneous sums and introduce another important chaotic structure.

Definition 1.3.1. Given $f \in \mathfrak{H}_0^{\odot d}$ with $d \in \mathbb{N}$ and $\Xi = (\xi_k, k \in \mathbb{N})$ a generic sequence of independent centered random variables, we define the *homogeneous sum* with order d, based on the kernel f, by setting

$$Q_d(f; \Xi) = \sum_{i_1,\dots,i_d \in \mathbb{N}} f(i_1, \dots, i_d)\xi_{i_1} \cdots \xi_{i_d} \,. \tag{1.3.2}$$

In particular, $I_p^{\mathfrak{X}}(f)$ in (1.2.4) is a pth order homogeneous sum based on the kernel $(\beta_{i_1,\dots,i_p})$ and the random variables $\mathfrak{X}(A_1), \dots, \mathfrak{X}(A_m)$.

If we pick Ξ to be a sequence of Rademacher random variables, we will get some object similar to the Wiener chaos, which will be called Rademacher chaos. In the sequel, we fix $\mathbf{X} = (X_k, k \in \mathbb{N})$ a sequence of independent Rademacher random variables such that

$$\mathbb{P}(X_k = 1) = p_k = 1 - q_k = 1 - \mathbb{P}(X_k = -1) \in (0, 1)$$

for each $k \in \mathbb{N}$. We call it the *symmetric* case, whenever $p_k = 1/2$ for each $k \in \mathbb{N}$; otherwise, we call it the *general* case. We write $\mathbf{Y} = (Y_k, k \in \mathbb{N})$ for the *normalized* version of $\mathbf{X}$, that is,

$$Y_k = \frac{X_k - p_k + q_k}{2\sqrt{p_k q_k}} \,, \tag{1.3.3}$$

for $k \in \mathbb{N}$.

Definition 1.3.2. Following the previous notation and for $f \in \mathfrak{H}_0^{\odot d}$, we call $Q_d(f; \mathbf{Y})$ the dth *discrete multiple integral* of f. We write $\mathbb{C}_d^{\mathbf{Y}} = \{Q_d(f; \mathbf{Y}) \,:\, f \in \mathfrak{H}_0^{\odot d}\}$ and call it the ***Rademacher chaos*** of order d, and as a convention, we put $\mathbb{C}_0^{\mathbf{Y}} = \mathbb{R}$. We always reserve the notation $Q_d(f)$ for $Q_d(f; \mathbf{Y})$. It is not difficult to check that for $f \in \mathfrak{H}_0^{\odot p}$ and $g \in \mathfrak{H}_0^{\odot q}$, it holds that

$$\mathbb{E}[Q_p(f)Q_q(g)] = \mathbf{I}_{\{p=q\}}\, p!\langle f, g\rangle_{\mathfrak{H}^{\otimes p}} \,.$$

This is known as the *orthogonality property* of the discrete multiple integrals and moreover, the Rademacher chaoses generate the space $L^2(\Omega, \sigma\{\mathbf{X}\}, \mathbb{P})$, that is, every $F \in L^2(\Omega, \sigma\{\mathbf{X}\}, \mathbb{P})$ admits a unique chaotic decomposition[23]

$$F = \mathbb{E}[F] + \sum_{p \geq 1} Q_p(h_p) \quad \text{with } h_p \in \mathfrak{H}_0^{\odot p} \text{ for each } p \in \mathbb{N}, \tag{1.3.4}$$

where the above series converge in $L^2(\mathbb{P})$. Mimicking the constructions in Section 1.2, we can define, based on the above chaos expansion, (discrete) Malliavin operators: D, δ, L, L^{-1}.

[23] Compared to the proofs of (1.2.6), the proof of (1.3.4) is much easier and will be sketched as follows: fixing $F \in L^2(\Omega, \sigma\{\mathbf{X}\}, \mathbb{P})$, by the martingale convergence theorem, it suffices to prove that for each $k \in \mathbb{N}$, $F_k := \mathbb{E}[F|\sigma\{Y_1, \dots, Y_k\}]$ belongs to the linear span $\mathcal{S}$ of all Rademacher chaoses. In fact, $L^2(\Omega, \sigma\{Y_1, \dots, Y_k\}, \mathbb{P})$ is isometric to $\mathbb{R}^{2^k}$ and $\{Y_{i_1} \cdots Y_{i_\ell} : (i_1, \dots, i_\ell) \in \triangle_\ell, 0 \leq \ell \leq k\}$ is its orthonormal basis, implying that $F_k \in \mathcal{S}$. Thus, $F \in \bar{\mathcal{S}}$, the $L^2(\mathbb{P})$-closure of $\mathcal{S}$ and this concludes the proof.

Malliavin operators. $\mathbb{D}$ is the collection of random variables $F \in L^2(\Omega, \sigma\{\mathbf{X}\}, \mathbb{P})$ as in (1.3.4) satisfying $\sum_{p\geq 1} pp!\|h_p\|^2 < +\infty$, and for such a $F \in \mathbb{D}$, we define the (discrete) Malliavin derivative $DF = (D_kF, k \in \mathbb{N}) \in \mathfrak{H}$ by

$$D_kF = \sum_{p\geq 1} pQ_{p-1}(h_p(k,\cdot)) .$$

Intuitively, D_kF annihilates the influence of Y_k within the expression of F, see also the following pathwise representation.

> we choose $\Omega = \{+1, -1\}^{\mathbb{N}}$ and define $\mathbb{P} = \bigotimes_{k\in\mathbb{N}} (p_k\delta_{+1} + q_k\delta_{-1})$. Then the coordinate projections $\omega = \{\omega_1, \cdots\} \in \Omega \longmapsto \omega_k =: X_k(\omega)$ form a sequence of independent Rademacher random variables under $\mathbb{P}$. We can define for $F \in L^2(\Omega, \sigma\{\mathbf{X}\}, \mathbb{P})$, $F^{\oplus k} := F(\omega_1, \cdots, \omega_{k-1}, 1, \omega_{k+1}, \cdots)$, that is, by fixing the kth coordinate in the configuration ω to be 1. Similarly, we define $F^{\ominus k} := F(\omega_1, \cdots, \omega_{k-1}, -1, \omega_{k+1}, \cdots)$. It holds that $D_kF = \sqrt{p_kq_k}(F^{\oplus k} - F^{\ominus k})$, see *e.g.* [88, Proposition 7.3]. The following results are clear:
>
> - $|F^{\oplus k} - F| = \mathbf{I}_{(X_k=-1)}\frac{|D_kF|}{\sqrt{p_kq_k}} \leq \frac{|D_kF|}{\sqrt{p_kq_k}}$ and $|F^{\ominus k} - F| = \mathbf{I}_{(X_k=1)}\frac{|D_kF|}{\sqrt{p_kq_k}} \leq \frac{|D_kF|}{\sqrt{p_kq_k}}$.
> - $F \in \mathbb{D}$ if and only if $\sum_{k\in\mathbb{N}} p_kq_k\mathbb{E}[|F^{\oplus k} - F^{\ominus k}|^2] < +\infty$. In particular, if $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, then $f(F) \in \mathbb{D}$.

We can define similarly the divergence operator δ by the duality relation $\mathbb{E}[\langle DF, u\rangle] = \mathbb{E}[F\delta(u)]$ for every $F \in \mathbb{D}$ and we define the Ornstein-Uhlenbeck generator $L = \sum_{p\geq 1} -pJ_p$, with J_p being the projection operator onto $\mathbb{C}_p^{\mathbf{Y}}$. The associated semigroup can be formally defined by $P_t = \exp(tL)$, $t \in \mathbb{R}_+$; and the pseudo-inverse of L is given by $L^{-1} = \sum_{p\geq 1} -p^{-1}J_p$. For a comprehensive treatment, one can refer to N. Privault's survey [88].

In 2010, Nourdin, Peccati and Reinert [70] combined Stein's method with the above discrete Malliavin calculus for the explicit bounds in the normal approximation of Rademacher functionals. Although the paper [70] only concerns the symmetric case (that is, when $\mathbf{X} = \mathbf{Y}$), the method and ideas therein are general enough to allow for extension to the general case.

A new product formula (for the symmetric case) is derived[24] in [70] that resembles (1.2.9): if $f \in \mathfrak{H}_0^{\odot p}$, $g \in \mathfrak{H}_0^{\odot q}$ $(p, q \in \mathbb{N})$, then

$$Q_p(f)Q_q(g) = \sum_{r=0}^{p\wedge q} r!\binom{p}{r}\binom{q}{r}Q_{p+q-2r}(f \star_r^r g\mathbf{I}_{\Delta_{p+q-2r}}) . \tag{1.3.5}$$

[24] Observe that in the symmetric case, $Y_k^2 = 1$, so $(i_1, \ldots, i_p) \in \Delta_p$, $(j_1, \ldots, j_q) \in \Delta_q$ share exactly r indices if and only if $Y_{i_1} \cdots Y_{i_p}Y_{j_1} \cdots Y_{j_q} \in \mathbb{C}_{p+q-2r}^{\mathbf{Y}}$. This observation and the symmetry of the kernels give us the combinatorial coefficients $r!\binom{p}{r}\binom{q}{r}$, the rest of the proof follows from the definition of star-contractions. Note that in the general case, one has $Y_k^2 = 1 + \frac{q_k-p_k}{\sqrt{p_kq_k}}Y_k$, then by modifying the proof for the symmetric case, one can still get some product formula for the general case, see [43, Proposition 2.2]: the constant 1 from Y_k^2 contributes to the same expression as the RHS of (2.2.4) while the other part gives arise to some complicated terms, for which one needs to impose more integrability conditions on the kernels.

It follows from this product formula that in the symmetric case, any random variable in a Rademacher chaos admits finite moments of any order. However, this is not true in the general case, as one can see from the product formula from [43].

Now fix two centered Rademacher functionals $F, G \in \mathbb{D}$ and $f \in C^1(\mathbb{R})$ with $\|f'\|_\infty < +\infty$, it follows that $f(F), L^{-1}F \in \mathbb{D}$ and by using duality relation, we have

$$\mathbb{E}[Gf(F)] = \mathbb{E}[\langle -DL^{-1}G, Df(F)\rangle] ;$$

see *e.g.* [**P1**, Lemma 2.1]. Note that the term $Df(F)$ is not equal to $f'(F)DF$ in general, unlike the chain rule on a Gaussian space. The authors of [70] obtained an approximate chain rule as follows: (in the symmetric case) let f be of class C^3 such that $f(F) \in \mathbb{D}$ and $\|f'''\|_\infty < +\infty$, then

$$D_k f(F) = f'(F)D_kF - \frac{1}{2}\Big[f''(F^{\oplus k}) + f''(F^{\ominus k})\Big](D_kF)^2 Y_k + \widetilde{R}_k \tag{1.3.6}$$

with $|\widetilde{R}_k| \le \frac{10}{3}\|f'''\|_\infty |D_kF|^3$. The requirement of high-order derivatives of the test function f forced the authors of [70] to use some smooth distance for the normal approximation.

As the first paper included in this thesis, [**P1**] provided the following neater chain rule that requires less regularity of the test function f.

Lemma 1.3.1 (G. Zheng, 2017)**.** If $F \in \mathbb{D}$ and $f : \mathbb{R} \longrightarrow \mathbb{R}$ is Lipschitz and diffentiable such that f' is Lipschitz continuous, then $D_k f(F) = f'(F)D_kF + R_k$, where the remainder term R_k is bounded by $\frac{\|f''\|_\infty}{2\sqrt{p_kq_k}}|D_kF|^2$ in the general setting[25].

This easy observation enabled us to obtain the same normal approximation in Wasserstein distance, which is one of the main achievements in the paper [**P1**].

Instead of using the chain rule, the authors of [44] carefully used a representation of the discrete Malliavin derivative $D\phi(F)$ and the fundamental theorem of calculus; this turns out to be flexible enough for them to deduce the Berry-Esséen bound in the symmetric case. Later they obtained the Berry-Esséen bound in the general case and provided applications to graph-counting statistics. See [**P1**, Remarks 3.1, 3.2] for connections between their work and our Wasserstein bound.

Recall that the Rademacher chaos belongs to the category of homogeneous sums, so it may be natural to study the normal approximation of homogeneous sums along the lines of the discrete Malliavin-Stein approach. Unfortunately, one can not find a suitable version of Malliavin calculus to be coupled with Stein's method. On the bright side, motivated by conjectures in theoretical computer science and social choice theory, Mossel, O'Donnell and Oleszkiewicz [59] established a generalized Lindeberg invariance principle for homogeneous sums. Roughly speaking, they were able to control distributional distance between homogeneous sums over different sequences of independent random variables in terms of maximal influence, see *e.g.* Theorem 2.1 therein.

[25]It is clear that $f(F) \in \mathbb{D}$, because f is Lipschitz. By Taylor formula in mean-value form, we write $f(y) - f(x) = f'(x)(y-x) + \mathcal{R}(f)$, where the remainder term $\mathcal{R}(f)$ is bounded by $\|f''\|_\infty \cdot |y-x|^2/2$. It remains to use the pathwise representation of D. See [**P1**, Lemma 2.2] for the detailed proof.

Let us consider another example, in which the notion of maximal influence is crucial. Fix $d \geq 2$ and a *divergent* sequence $(N_n, n \geq 1)$ of natural numbers. Consider the kernels $f_n : \{1, \dots, N_n\}^d \to \mathbb{R}$ symmetric and vanishing on diagonals and $d!\|f_n\|^2_{\mathfrak{H}^{\otimes d}} = 1$, then according to (1.3.2),

$$Q_d(f_n; \Xi) = \sum_{i_1,\dots,i_d \leq N_n} f_n(i_1, \dots, i_d)\xi_{i_1} \cdots \xi_{i_d} \,.$$

The following central limit theorem[26] due to de Jong [26] gave sufficient conditions for asymptotic normality of $Q_d(f_n; \Xi)$.

Theorem 1.3.1 (P. de Jong, 1990)**.** Under the above setting, let $\Xi = (\xi_i, i \geq 1)$ be a sequence of independent centered random variables with unit variance and *finite fourth moments*. If $\mathbb{E}[Q_d(f_n; \Xi)^4] \to 3$ and the maximal influence $\mathcal{M}(f_n) \to 0$ as $n \to +\infty$, then $Q_d(f_n; \Xi)$ converges in law to a standard Gaussian.

Let us restrict ourselves to the Gaussian setting for a while: when $\mathbf{G}$ is a sequence of *i.i.d.* standard Gaussians, $Q_d(f_n; \mathbf{G})$ belongs to the dth Gaussian Wiener chaos, and the fourth moment theorem [76] implies that if $Q_d(f_n; \mathbf{G})$ converges in law to a standard Gaussian (or equivalently $\mathbb{E}[Q_d(f_n; \mathbf{G})^4] \to 3$), then $\|f_n \otimes_{d-1} f_n\|_{\mathfrak{H}^{\otimes 2}} \to 0$. While $\mathcal{M}(f_n) \leq \|f_n \otimes_{d-1} f_n\|_{\mathfrak{H}^{\otimes 2}}$ due to [70, Lemma 2.4], so that $\mathcal{M}(f_n) \to 0$. This observation, together with the invariance principle from [59], hints the universality of the Gaussian Wiener chaos; see particularly the following interesting result that is (slightly) adapted from Theorem 7.5 in [71].

Theorem 1.3.2 (I. Nourdin, G. Peccati and G. Reinert, 2010)**.** Fix integers $d \geq 2$ and $q_d \geq \ldots \geq q_1 \geq 2$. For each $j \in [d]$, let $(N_{j,n}, n \geq 1)$ be a sequence of natural numbers diverging to infinity, and let $f_{j,n} : [N_{j,n}]^{q_j} \to \mathbb{R}$ be symmetric and vanishing on diagonals (*i.e.* $f_{j,n} \in \mathfrak{H}_0^{\odot q_j}$ with support contained in $[N_{j,n}]^{q_j}$) such that

$$\lim_{n\to+\infty} \mathbf{I}_{(q_k=q_l)} q_k! \sum_{i_1,\dots,i_{q_k} \leq N_{k,n}} f_{k,n}(i_1, \dots, i_{q_k}) f_{l,n}(i_1, \dots, i_{q_k}) = \Sigma_{k,l} \,,$$

where $\Sigma = (\Sigma_{i,j}, 1 \leq i, j \leq d)$ is a symmetric nonnegative definite d by d matrix. Then the following statements are equivalent:

(A_1) Given a sequence $\mathbf{G}$ of i.i.d. standard Gaussians, $(Q_{q_1}(f_{1,n}; \mathbf{G}), \dots, Q_{q_d}(f_{d,n}; \mathbf{G}))^T$ converges in distribution to $\mathcal{N}(0, \Sigma)$, as $n \to +\infty$.

(A_2) For ***every*** sequence $\Xi = (\xi_i, i \in \mathbb{N})$ of independent centered random variables with unit variance and $\sup_{i\in\mathbb{N}} \mathbb{E}[|\xi_i|^3] < +\infty$, the sequence of d-dimensional random vectors $(Q_{q_1}(f_{1,n}; \Xi), \dots, Q_{q_d}(f_{d,n}; \Xi))^T$ converges in distribution to $\mathcal{N}(0, \Sigma)$, as $n \to +\infty$.

A similar universality result for Poisson Wiener chaos was first established in [86] and refined recently in our work **[P4]**. It is worth noting that by assuming

$$\sup_{n,j} \mathbb{E}[Q_{q_j}(f_{j,n}; \mathbf{G})^2] < +\infty \,, \tag{1.3.7}$$

Bai and Taqqu [7] were able to improve the above universality result by allowing an infinite number of terms in the homogeneous sums. We will follow this to formulate our universality results in Chapter 3.

[26] See also Döbler and Peccati's recent work [31] for the quantitative version.

1.3.2 Stein's method and normal approximation of Poisson functionals

Due to our *lack* of creativity and nevertheless as it conveys the right idea, we simply duplicate the title of [82] here. As already mentioned, Peccati, Solé, Taqqu and Utzet launched the discrete Malliavin-Stein approach on the Poisson space and they used additionally a pathwise representation of the derivative operator D that involves the standard difference operators.

A pathwise representation of D**.** Let us first represent the state space $\mathcal{S}_\sigma$ of our Poisson random measure η:

$$\mathcal{S}_\sigma := \left\{ \mathrm{w} = \sum_{j=0}^{n} \delta_{z_j} \,:\, n \in \mathbb{N} \cup \{+\infty\}\,,\ z_j \in \mathcal{Z} \right\},$$

where δ_z denotes the Dirac measure at $z \in \mathcal{Z}$. We equip $\mathcal{S}_\sigma$ with $\mathscr{S}_\sigma$, the smallest σ-algebra that renders $\mathrm{w} \in \mathcal{S}_\sigma \longmapsto \mathrm{w}(B)$ measurable for all $B \in \mathscr{Z}$. So we can see η as a random element in the measurable space $(\mathcal{S}_\sigma, \mathscr{S}_\sigma)$ and $\mathbb{P}_\eta := \mathbb{P} \circ \eta^{-1}$ denotes its distributional measure, that is, $\mathbb{P}_\eta(A) = \mathbb{P}(\eta^{-1}(A))$, $A \in \mathscr{S}_\sigma$. Under $\mathbb{P}_\eta$, the *canonical mapping* associated to η, $\{\mathrm{w}(B) := \eta(B)(\mathrm{w}),\ B \in \mathscr{Z}\}$ is distributed as a Poisson point process with intensity measure μ. In this framework, one has

$$\mathbb{P}_\eta(\mathrm{w} \in \mathcal{S}_\sigma \,:\, \mathrm{w}(B) < +\infty\,,\ \forall B \in \mathscr{Z}_\mu) = \mathbb{P}_\eta(\mathrm{w} \in \mathcal{S}_\sigma \,:\, \mathrm{w}(\{z\}) < +\infty\,,\ \forall z \in \mathcal{Z}) = 1\,,$$

see [47, 49] for more details. Now we define the so-called *add-one cost* operators $(D_z^+, z \in \mathcal{Z})$: given any $\sigma\{\eta\}$-measurable real-valued random variable F, we can write $F = \mathfrak{f}(\eta)$ for some $\mathscr{S}_\sigma$-measurable function $\mathfrak{f} : \mathcal{S}_\sigma \to \mathbb{R}$ and this *representative* $\mathfrak{f}$ is determined $\mathbb{P}_\eta$-almost surely. Then we define $D_z^+ F = \mathfrak{f}(\eta + \delta_z) - \mathfrak{f}(\eta)$, $z \in \mathcal{Z}$. According to [47, Theorem 3], for $F \in \mathbb{D}$, $D_z^+ F = D_z F$, $\mathbb{P}$-almost surely and μ-almost everywhere. This in particular implies that

$$F = \mathfrak{f}(\eta) \in \mathbb{D} \quad \Leftrightarrow \quad \mathbb{E}[\|DF\|^2] = \int_{\mathcal{Z}} \mathbb{E}[|\mathfrak{f}(\eta + \delta_z) - \mathfrak{f}(\eta)|^2]\, \mu(dz) < +\infty\,,$$

from which we can deduce further that $g(F) \in \mathbb{D}$ whenever $F \in \mathbb{D}$ and $g : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous.

Moreover, given a normalized $F \in \mathbb{D}$ and $\varphi \in C^2(\mathbb{R})$ with $\|\varphi\|_\infty + \|\varphi'\|_\infty + \|\varphi''\|_\infty < +\infty$, we can use the above pathwise representation and similar reasoning as in footnote 25 to obtain the following approximate chain rule: $D_z\varphi(F) = D_z^+\varphi(F) = \varphi'(F) D_z F + R_z(F)$, where $R_z(F)$ is a random variable such that $|R_z(F)| \le \frac{1}{2}\|\varphi''\|_\infty (D_z F)^2$. With Stein's ideas in mind (see Section 1.2.2), we use the integration by parts formula on the Poisson space and obtain that

$$\begin{aligned} \mathbb{E}[F\varphi(F)] = \mathbb{E}[\,-\delta D L^{-1} F \varphi(F)] &= \mathbb{E}[\langle -DL^{-1}F, D\varphi(F)\rangle] \\ &= \mathbb{E}[\langle -DL^{-1}F, DF\rangle \varphi'(F)\rangle] + \mathbb{E}[\langle -DL^{-1}F, R_\cdot(F)\rangle]\,, \end{aligned}$$

thus,

$$\begin{aligned} &\Big|\mathbb{E}[F\varphi(F) - \varphi'(F)]\Big| \\ &\le \|\varphi'\|_\infty \mathbb{E}[|1 - \langle -DL^{-1}F, DF\rangle|] + \frac{1}{2}\|\varphi''\|_\infty \int_{\mathcal{Z}} \mathbb{E}\Big[|D_z L^{-1} F|(D_z F)^2\Big]\, \mu(dz)\,. \end{aligned} \tag{1.3.8}$$

This bound is one of the main results in [82], and from this bound, the authors of [82] obtained sufficient conditions for the asymptotic normality inside a fixed Poisson Wiener chaos after a lengthy computation involving the product formula. This paper and several other follow-up articles have opened up a new chapter of stochastic analysis for Poisson point process in the sense that in the past few years, we have witnessed lively applications of Malliavin calculus to problems from stochastic geometry. Besides the aforementioned references, we refer the interested readers to the book [80] for recent developments.

Unfortunately, since the work of Nualart and Peccati, the natural question

"whether or not the FMT exists on the Poisson space"

had been elusive for years except for some partial results. In 2008, Peccati and Taqqu [83] presented a FMT on the second Poisson Wiener chaos under *some* assumptions; in 2016, Bourguin and Peccati provided in [80, Chapter 6] a FMT on the third Poisson Wiener chaos under similar assumptions. And in the paper [86], Peccati and Zheng[27] established the FMTs for the discrete Poisson Wiener chaos, under some *mild* condition. For the later reference, we state it in the following.

Theorem 1.3.3 (G. Peccati and C. Zheng, 2016)**.** Let $\mathbf{P} := (P_i, i \geq 1)$ be a sequence of independent Poisson random variables with parameter $\lambda_i \in (0, +\infty)$ such that[28] $\alpha := \inf\{\lambda_i : i \geq 1\} > 0$ and denote by $\widehat{\mathbf{P}}$ the normalized version of $\mathbf{P}$, *i.e.* $\widehat{P_i} = (P_i - \lambda_i)/\sqrt{\lambda_i}$ for each $i \in \mathbb{N}$. Now fix integers $d \geq 1$ and $q_d \geq \ldots \geq q_1 \geq 1$, and let Σ, $\{N_{j,n}, f_{j,n} \colon j \in [d], n \in \mathbb{N}\}$ be given as in Theorem 1.3.2. Then the following statements are equivalent as $n \to +\infty$:

(i) $(Q_{q_1}(f_{1,n}; \widehat{\mathbf{P}}), \ldots, Q_{q_d}(f_{d,n}; \widehat{\mathbf{P}}))^T$ converges in law to $\mathcal{N}(0, \Sigma)$.

(ii) For each $j \in [d]$, $Q_{q_j}(f_{j,n}; \widehat{\mathbf{P}})$ converges in law to $\mathcal{N}(0, \Sigma_{j,j})$.

(iii) For each $j \in [d]$, $\mathbb{E}[Q_{q_j}(f_{j,n}; \widehat{\mathbf{P}})^4]$ converges to $3\Sigma_{j,j}^2$.

Remark 1.3.1. In fact, the condition "$\alpha > 0$" is not necessary for the implication (iii)$\Rightarrow$(i). As mentioned in footnote 28, (suppose $q_j \geq 2$) the statement (iii) will force $\|f_{j,n} \star_p^p f_{j,n}\| \to 0$ for every $p \in \{1, \ldots, q_j - 1\}$, then the desired implication follows immediately Peccati-Tudor theorem and Theorem 1.3.2. This line of reasoning gives rise to the universality result for discrete Poisson chaos in [86].

[27] According to Wikipedia, in 2006, the family name "Zheng" ranked 21st in China's list of top 100 most common surnames. Zheng belongs to the second major group of ten surnames that makes up more than 10% of the Chinese population.

[28] The condition "$\alpha > 0$" is very crucial for the proof in [86]: (1) it guarantees the uniform integrability of the sequence $\{Q_{q_j}(f_{j,n}; \widehat{\mathbf{P}})^4, n \in \mathbb{N}\}$ for each $j \in [d]$, so that the implication (ii)$\Rightarrow$ (iii) holds true. (2) One can compute the fourth moment of $Q_{q_j}(f_{j,n}; \widehat{\mathbf{P}})$ by exploiting the orthogonality property, and (suppose $q_j \geq 2$) the statement (iii) will force $\|f_{j,n} \star_p^p f_{j,n}\| \to 0$ for every $p \in [q_j - 1]$, while the condition "$\alpha > 0$" allows one to bound $\|f_{j,n} \star_{q_j}^l f_{j,n}\|^2$ by $\alpha^{l-q_j} \|f_{j,n} \star_l^l f_{j,n}\|^2$ for every $l \in [q_j - 1]$ and bound $\|f_{j,n} \star_r^l f_{j,n}\|^2$ by $\alpha^{l-r} \|f_{j,n} \star_l^l f_{j,n}\|^2$ for every $l \in [r]$ and $r \in [q_j - 1]$. Then Peccati and Zheng applied [82, Theorem 5.1] to obtain the asymptotical normality of $Q_{q_j}(f_{j,n}; \widehat{\mathbf{P}})$; see also Remark 1.3.1.

Theorem 1.3.4 (G. Peccati and C. Zheng, 2014)**.** Let the assumptions of Theorem 1.3.3 prevail and we assume in addition that $d \geq 2$, $q_1 \geq 2$. Then the following assertions are equivalent as $n \to +\infty$:

(A_1) $(Q_{q_1}(f_{1,n}; \widehat{\mathbf{P}}), \ldots, Q_{q_d}(f_{d,n}; \widehat{\mathbf{P}}))^T$ converges in law to $\mathcal{N}(0, \Sigma)$, as $n \to +\infty$.

(A_2) For ***every*** sequence $\Xi = (\xi_i, i \in \mathbb{N})$ of independent centered random variables with unit variance and $\sup_{i\in\mathbb{N}} \mathbb{E}[|\xi_i|^3] < +\infty$, the sequence of d-dimensional random vectors $(Q_{q_1}(f_{1,n}; \Xi), \ldots, Q_{q_d}(f_{d,n}; \Xi))^T$ converges in distribution to $\mathcal{N}(0, \Sigma)$.

In 2017, Döbler and Peccati [32] adopted Ledoux's spectral point-of-view [51] and used a pathwise representation of the *carré-du-champ* operator to successfully prove a FMT on a general Poisson Wiener chaos, under the so-called **Assumption (A)**. Their results are of quantitative nature and Stein's method played a fundamental role there.

Theorem 1.3.5 (Döbler and Peccati, 2017)**.** Fix an integer $q \geq 1$ and let $F \in \mathbb{C}_q^\eta$ be such that $F \in L^4(\mathbb{P})$, $\mathbb{E}[F^2] = 1$ and satisfy the following assumption

(A) $\displaystyle\int_{\mathcal{Z}} \mathbb{E}\Big[|D_z^+ F| + |F D_z^+ F| + (D_z^+ F)^4 + |F^3 D_z^+ F|\Big] \mu(dz) < +\infty$, where D_z^+ is the add-one cost operator.

Then, we have the following fourth moment bound in Wasserstein distance

$$d_{\mathrm{W}}(F, Z) := \sup_{\|h'\|_\infty \leq 1} \Big|\mathbb{E}[h(F) - h(Z)]\Big| \leq (2 + \sqrt{2/\pi}) \sqrt{\mathbb{E}[F^4] - 3}\,,$$

where $Z \sim \mathcal{N}(0, 1)$.

Remark 1.3.2. The main reason for the appearance of **(A)** was that certain intrinsic tools, notably the *Mecke formula*[29] (see *e.g.* [49, Chapter 4]), require these integrability conditions. Nevertheless, it is remarkable for Döbler and Peccati to establish the *exact* FMT on the Poisson space.

In the paper [**P4**] jointly written with Döbler and Vidotto, we were able to remove the **Assumption (A)** in the above theorem, using a very different and elementary approach. Following [32], Döbler and Krokowski [30] established the following fourth-moment-influence bound.

Theorem 1.3.6 (Döbler and Krokowski, 2017)**.** Fix $p \in \mathbb{N}$ and $f \in \mathfrak{H}_0^{\odot p}$ satisfying $p!\|f\|^2_{\mathfrak{H}^{\otimes p}} = 1$. Let Z be a standard Gaussian and $F = Q_p(f; \mathbf{Y}) \in L^4(\mathbb{P})$, then,

$$d_{\mathrm{W}}(F, Z) \leq C_1 \sqrt{\Big|\mathbb{E}[F^4] - 3\Big|} + C_2 \sqrt{\mathcal{M}(f)}\,,$$

where C_1, C_2 are two numerical constants and recall that $\mathcal{M}(f)$ denotes the maximal influence of f. This result echoes de Jong's Theorem 1.3.1.

[29] It is named after Joseph Mecke. One shall not confuse this name with Meckes' theorem, which we name after Elizabeth Meckes.

Remark 1.3.3. Döbler and Krokowski's result is *optimal* in the sense that there are examples, in which the fourth moment condition alone would not guarantee the asymptotic normality, see Example 1.5 and Theorem 1.6 in [30].

In the paper [**P5**], we followed our own approach to re-derive the above bound and more importantly, we provided a multivariate extension. To put in a short way, our approach is inspired by our paper [**P3**] with I. Nourdin, in which we constructed exchangeable pair couplings motivated by *Mehler formula* and used Meckes' theorems to recover Nourdin-Peccati bound in any dimension. This idea is simple while flexible enough to allow us to obtain FMTs on the Poisson space and the fourth-moment-influence theorems in the Rademacher setting.

The core of this thesis is to provide unified proofs for these results and to shed some light, as we wish, on how far we can go.

1.4 What is new?

Let us begin with point (ii) from the last paragraph in Section 1.1. In 2012, M. Ledoux analyzed Nualart and Peccati's FMT from the abstract point-of-view of a Markov diffusion generator. And the carré du champ is at the center of his strategy. Let us now quickly sketch Ledoux's framework: let $\mathbf{L}$ be a Markov diffusion generator on some state space $(E, \mathcal{E})$ with invariant and reversible probability measure γ and symmetric bilinear carré du champ operator

$$\Gamma(f,g) = \frac{1}{2}\big[\mathbf{L}(fg) - f\mathbf{L}g - g\mathbf{L}f\big]$$

acting on functions f, g in a suitable domain $\mathcal{A}$. To ease the notation, we write $\Gamma(f) = \Gamma(f,f)$. By symmetry and invariance of γ, the following integration by parts formula holds

$$\int_E -f\mathbf{L}g \, d\gamma = \int_E -g\mathbf{L}f \, d\gamma = \int_E \Gamma(f,g)\, d\gamma \, .$$

And the *diffusion property* of $\mathbf{L}$ asserts that $\mathbf{L}\phi(f) = \phi'(f)\mathbf{L}f + \phi''(f)\Gamma(f)$, for ϕ smooth and $f \in \mathcal{A}$. For example, when $\mathbf{L}$ is the Ornstein-Uhlenbeck generator on the Wiener space, $\Gamma(f,g) = \langle Df, Dg\rangle$ with D being the Malliavin derivative. So (1.2.10) can be rephrased as

$$\mathbf{Var}(q^{-1}\Gamma(F)) \leq \frac{q-1}{3q}\big(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2\big) \leq (q-1)\mathbf{Var}(q^{-1}\Gamma(F))\, .$$

Putting $\Gamma_1 = \Gamma$, one can define the iterated gradients by setting $\Gamma_m(f,g) = \frac{1}{2}[\mathbf{L}\Gamma_{m-1}(f,g) - \Gamma_{m-1}(f, \mathbf{L}g) - \Gamma_{m-1}(g, \mathbf{L}f)]$, $m \geq 2$. In [51], Ledoux gave the definition of "general chaos eigenfunction" using these iterated gradients, and he was able to provide FMT for the chaos eigenfunctions. The slight drawback of Ledoux's definition is that it only includes the Wiener structure, that is, when γ is the standard Gaussian measure. Soon after Ledoux's insightful idea, Azmoodeh, Campese and Poly [4] generalized the notion of chaos eigenfunctions: assume in addition that the generator $-\mathbf{L}$ has pure spectrum $\{0 = \lambda_0 < \lambda_1 < \lambda_2 < \ldots\}$, they called X a pth order chaos eigenfunction if $\mathbf{L}X = -\lambda_p X$ and

$$X^2 \in \bigoplus_{\alpha \leq \lambda_{2p}} \mathbf{Ker}(\mathbf{L} + \alpha \mathbf{Id})\, .$$

This new definition includes two more fundamental structures: Laguerre structures and Jacobi structures, see [8]. And Azmoodeh, Campese and Poly were able to simplify a lot the previous proofs. This set of techniques were later taken in [5] to establish a generalization of Nualart-Peccati's criterion: given an integer $p \geq 2$, $F_n \in \mathbb{C}_p^W$ with $\mathbf{Var}(F_n) = 1$, F_n converges in law to a standard Gaussian Z if and only if as $n \to +\infty$,

(1.4.1) $$\mathbb{E}[F_n^{2k}] \text{ converges to } \mathbb{E}[Z^{2k}] = (2k-1)!! \text{ for } \textit{some} \text{ integer } k \geq 2.$$

It is also worth mentioning that Ledoux's spectral viewpoint promoted some recent advance around the *real Gaussian product conjecture*, see [52].

Back to our chaoses (Gaussian, Poisson Wiener chaoses and Rademacher chaos), a nice Ornstein-Uhlenbeck structure exists. We can define the carré du champ Γ as above:

$$\Gamma(F, G) := \frac{1}{2}\big(L(FG) - FLG - GLF\big),$$

whenever the above expressions are well defined. In the work [32], Döbler and Peccati took a pathwise representation of the carré du champ and combined it with the usual arguments in the Malliavin-Stein approach. As already mentioned, they applied Mecke formula in a delicate way so as to obtain an exact FMT on the Poisson space, see Theorem 1.3.5. Soon later, Döbler and Krokowski adapted the ideas from [32] and provided the so-called fourth-moment-influence theorem.

One obvious difficulty in view of applying the Malliavin-Stein techniques is that in the discrete settings (*i.e.* Poisson space, Rademacher setting), the Ornstein-Uhlenbeck operator L does not possess the diffusion property. Let us assume that we have instead an approximate-diffusion property as follows:

$$L\phi(F) = \phi'(F)LF + \phi''(F)\Gamma(F) + R(F) \quad \text{for nice } \phi \text{ and } F,$$

where $R(F)$ is a remainder. Note that for a *normalized* $F \in \mathbb{C}_p^{\mathfrak{X}}$, with $\mathfrak{X} \in \{W, \eta, \mathbf{Y}\}$, one has $LF = -pF$ and $\mathbb{E}[L\phi(F)] = 0$ for nice ϕ. If we have nice control over $\mathbb{E}[R(F)]$, say $|\mathbb{E}[R(F)]| \leq c\sqrt{\mathbb{E}[F^4] - 3}$, then taking expectations at both sides gives us

$$\Big| -p\mathbb{E}[F\phi'(F)] + p\mathbb{E}[\phi''(F)] + \mathbb{E}[\phi''(F)(\Gamma(F) - p)]\Big| \leq c\sqrt{\mathbb{E}[F^4] - 3}\,.$$

In view of Stein's ideas, if $\mathbf{Var}(\Gamma(F))$ can be further bounded by a multiple of $\sqrt{\mathbb{E}[F^4] - 3}$, then the above inequality will give us the desired fourth moment bound.

To put in a simple way, the rest of this thesis is mainly devoted to demonstrating the use of Stein's method of exchangeable pairs for controlling the expectation $\mathbb{E}[R(F)]$. As mentioned before, we will prove the following FMT on the Poisson space under the *weakest* possible assumption of finite fourth moment.

Theorem 1.4.1 (C. Döbler, A. Vidotto and G. Zheng, 2017)**.** Fix an integer $q \geq 1$ and let $F \in \mathbb{C}_q^\eta$ be such that $F \in L^4(\mathbb{P})$ and $\mathbb{E}[F^2] = 1$. Then,

$$d_{\mathrm{W}}(F, Z) \leq (2 + \sqrt{2/\pi})\sqrt{\mathbb{E}[F^4] - 3}\,,$$

where $Z \sim \mathcal{N}(0, 1)$. See also [**P4**].

Corollary 1.4.1 (C. Döbler, A. Vidotto and G. Zheng, 2017)**.** Let $q_n \in \mathbb{N}$ and $F_n \in \mathbb{C}_{q_n}^{\mathfrak{X}_n}$, with $\mathfrak{X}_n \in \{W, \eta\}$ for each $n \in \mathbb{N}$. If $\mathbb{E}[F_n^2] \to 1$ and $\mathbb{E}[F_n^4] \to 3$, as $n \to +\infty$, then F_n converges in law to a standard Gaussian distribution.

We will provide the multivariate extension of the above results in Chapter 3. Along similar lines, we will prove the following Peccati-Tudor type theorem in the Rademacher setting.

Theorem 1.4.2 (G. Zheng, 2017)**.** Fix integers $d \geq 2$ and $1 \leq q_1 \leq \ldots \leq q_d$, and consider the sequence of random vectors

$$F^{(n)} = (F_1^{(n)}, \ldots, F_d^{(n)})^T := \big(Q_{q_1}(f_{1,n}), \ldots, Q_{q_d}(f_{d,n})\big)^T$$

with kernels $f_{j,n}$ in $\mathfrak{H}_0^{\odot q_j}$ for each $n \in \mathbb{N}, j \in [d]$. Assume that the covariance matrix Σ_n of $F^{(n)}$ converges in Hilbert-Schmidt norm to a nonnegative definite symmetric matrix $\Sigma = (\Sigma_{i,j}, i, j \in [d])$, as $n \to +\infty$. Suppose that the following condition holds:

$$\lim_{n\to+\infty} \sum_{j=1}^{d} \mathcal{M}(f_{j,n}) = 0 \ .$$

If for each $j \in [d]$, $\mathbb{E}[(F_j^{(n)})^4]$ converges to $3\Sigma_{j,j}^2$, as $n \to +\infty$, then $F^{(n)}$ converges in distribution to $Z \sim \mathcal{N}(0, \Sigma)$, as $n \to +\infty$.

Remark 1.4.1. Although we just presented the above qualitative result in [**P5**], its proof therein is of quantitative nature. We will provide a quantitative version of Theorem 3.2.2 in Chapter 3.

As a byproduct of our proofs, we establish an interesting transfer principle (see [**P4**]), which is closely related to the universality phenomena of homogeneous sums.

Proposition 1.4.1 (C. Döbler, A. Vidotto and G. Zheng, 2017)**.** Given $p \in \mathbb{N}$, $f_n \in L_s^2(\mu^p)$ for each $n \in \mathbb{N}$ such that

$$\lim_{n\to+\infty} p!\,\|f_n\|^2 = 1 \ ,$$

then the following implications holds (with $Z \sim \mathcal{N}(0,1)$)

$$\lim_{n\to+\infty} \mathbb{E}[I_p^\eta(f_n)^4] = 3 \Longrightarrow \lim_{n\to+\infty} \mathbb{E}[I_p^W(f_n)^4] = 3 \Longrightarrow \lim_{n\to+\infty} d_{\mathrm{TV}}(I_p^W(f_n), Z) = 0 \ ,$$

where d_{TV} denotes the total-variation distance that is much stronger than the Kolmogorov distance; see Chapter 2 for more details.

Remark 1.4.2. One can see the above result as a transfer principle from Poisson to Gaussian, and it is worth pointing out that the transfer principle "from-Gaussian-to-Poisson" does *not* hold true, due to a counterexample given in [15]; see Proposition 5.4 therein. For the proof of Proposition 1.4.1, see Remark 3.1.6-(4).

We close this introductory chapter with the following universality result that is a blend of results from [7, 71, 86] and [**P4, P5**].

Theorem 1.4.3. Fix integers $d \geq 2$ and $q_d \geq \ldots \geq q_1 \geq 2$. For each $j \in [d]$ and $n \in \mathbb{N}$, let the kernels $f_{j,n} \in \mathfrak{H}_0^{\odot q_j}$ satisfy $\sup_{n,j} \|f_{j,n}\|^2 < +\infty$ and for $k, l \in [d]$

$$\lim_{n\to+\infty} \mathbf{I}_{(q_k=q_l)} q_k! \sum_{i_1,\ldots,i_{q_k}\in\mathbb{N}} f_{k,n}(i_1,\ldots,i_{q_k}) f_{l,n}(i_1,\ldots,i_{q_k}) = \Sigma_{k,l}\,,$$

where $\Sigma = (\Sigma_{i,j}, i, j \in [d])$ is a symmetric nonnegative definite d by d matrix. Then the following statements are equivalent, as $n \to +\infty$:

(C_1) Given a sequence $\mathbf{G}$ of i.i.d. standard Gaussians,

$$(Q_{q_1}(f_{1,n}; \mathbf{G}), \ldots, Q_{q_d}(f_{d,n}; \mathbf{G}))^T$$

converges in law to $\mathcal{N}(0, \Sigma)$.

(C_2) Given a sequence $\mathbf{V}$ of i.i.d. random variables with $V_1 + 1 \sim \mathrm{Poi}(1)$,

$$(Q_{q_1}(f_{1,n}; \mathbf{V}), \ldots, Q_{q_d}(f_{d,n}; \mathbf{V}))^T$$

converges in law to $\mathcal{N}(0, \Sigma)$.

(C_3) In the ***symmetric*** Rademacher setting,

$$(Q_{q_1}(f_{1,n}; \mathbf{Y}), \ldots, Q_{q_d}(f_{d,n}; \mathbf{Y}))^T$$

converges in law to $\mathcal{N}(0, \Sigma)$ and $\mathcal{M}(f_{1,n}) + \ldots + \mathcal{M}(f_{d,n}) \to 0$.

(C_4) For ***every*** sequence $\Xi = (\xi_i, i \in \mathbb{N})$ of independent centered random variables with unit variance and $\sup_{i\in\mathbb{N}} \mathbb{E}[|\xi_i|^3] < +\infty$,

$$(Q_{q_1}(f_{1,n}; \Xi), \ldots, Q_{q_d}(f_{d,n}; \Xi))^T$$

converges in law to $\mathcal{N}(0, \Sigma)$.

This page is left blank.

Chapter 2

Preliminaries: *$\mathcal{E}$xchangeable pairs* $\rightsquigarrow \bullet \cdots \bullet \leftrightsquigarrow$ *Carré du champ*

Abstract

This chapter consists of two sections. Section 2.1 is devoted to basics of Stein's method and in particular, we provide in Section 2.1.2 *generalized* Meckes' theorems, given an infinitesimal version of exchangeable pairs. In Section 2.2, we construct exchangeable pairs on Gaussian, Poisson spaces and in Rademacher setting; finally we link them to the Ornstein-Uhlenbeck operator and carré du champ. Our construction is naturally motivated by the classic *Mehler formulae*.

2.1 Stein's method of normal approximation

Let us start with the classic *central limit theorem*: suppose $\{X_j, j \geq 1\}$ is a sequence of *i.i.d* random variables with zero mean and unit variance, then

$$\frac{X_1 + \ldots + X_n}{\sqrt{n}} \quad \text{converges in law to } \mathcal{N}(0,1), \text{ as } n \to +\infty.$$

A standard proof consists of using the Fourier transform so as to establish the convergence of characteristic functions. However, this analytic proof strongly relies on the independence and leaves us no clue about how fast this distributional convergence happens. Moreover, this Fourier-based approach, in general, can not be applied to situations where some dependence arises. The *Stein's method* is a very powerful toolbox of techniques that can be used not only to prove limit theorems, but also to provide rates of convergence in some chosen metrics, for instance the *Berry-Esséen bound* in the *Kolmogorov distance*, see *e.g.* [67, Theorem 3.7.1].

The Stein's method is named after Charles Stein, one of the leading statisticians in the last century. In 1972, Charles Stein published his method concerning normal approximation in [103], and after its first appearance, this method has been modified and further developed by Stein himself as well as many other mathematicians. We refer interested readers to the treatise [19].

2.1.1 Basics on Stein's method

We will begin with probability distributions on $\mathbb{R}$ and later the generalization to multivariate setting will be presented. Here is an easy observation: if $Z \sim \mathcal{N}(0,1)$, $f \in C_b^\infty(\mathbb{R})$, then

$$\mathbb{E}[Zf(Z)] = \int_{-\infty}^{\infty} xf(x)\frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx = -\int_{-\infty}^{\infty} f(x)\frac{d(e^{-x^2/2})}{\sqrt{2\pi}} = \int_{-\infty}^{\infty} f'(x)\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx = \mathbb{E}[f'(Z)]\,.$$

This is essentially a Gaussian integration-by-parts formula, see (1.2.8). The converse statement is also true: if Z is integrable and $\mathbb{E}[f'(Z)] = \mathbb{E}[Zf(Z)]$ for all $f \in C_b^\infty(\mathbb{R})$, then Z must be distributed as the standard Gaussian. Indeed, considering $f(x) = \sin(\lambda x)$ and $f(x) = \cos(\lambda x)$, $\lambda \in \mathbb{R}$, we have $\mathbb{E}[Z\sin(\lambda Z)] = \lambda\mathbb{E}[\cos(\lambda Z)]$ and $\mathbb{E}[Z\cos(\lambda Z)] = -\lambda\mathbb{E}[\sin(\lambda Z)]$, from which we deduce that $\mathbb{E}[Ze^{\mathrm{i}\lambda Z}] = \mathrm{i}\lambda\mathbb{E}[e^{\mathrm{i}\lambda Z}]$ for any $\lambda \in \mathbb{R}$. Recognize that $\varphi_Z(\lambda) = \mathbb{E}[e^{\mathrm{i}\lambda Z}]$ is the characteristic function of Z and $\mathrm{i}\mathbb{E}[Ze^{\mathrm{i}\lambda Z}] = \varphi_Z'(\lambda)$, which gives us an ordinary differential equation $\varphi_Z'(\lambda) + \lambda\varphi_Z(\lambda) = 0$ subject to $\varphi_Z(0) = 1$. This ODE has a unique solution $\varphi_Z(\lambda) = \exp(-\lambda^2/2)$, which is the characteristic function of $\mathcal{N}(0,1)$.

The following result, known as *Stein's lemma*, summarizes the above discussion.

Lemma 2.1.1. Let Z be an integrable random variable, then $Z \sim \mathcal{N}(0,1)$ if and only if $\mathbb{E}[f'(Z)] = \mathbb{E}[Zf(Z)]$ for any continuously differentiable function $f : \mathbb{R} \to \mathbb{R}$ such that $f'(Z), Zf(Z) \in L^1(\mathbb{P})$, and f, f' have at most *polynomial growth at infinity*[1].

Heuristically speaking, the above characterization hints that a real random variable Z is *close* to a standard Gaussian, whenever $\mathbb{E}[f'(Z) - Zf(Z)]$ is close to zero for every $f \in \mathscr{H}$, with $\mathscr{H}$ being some rich enough class of functions.

Then how to quantify the distance between two probability distributions?

Let $\mathscr{H}$ be a *separating* class[2] of real bounded measurable functions and we define, for two probability measures μ, ν, that $d_{\mathscr{H}}(\mu,\nu) := \sup_{h\in\mathscr{H}} \left| \int_{\mathbb{R}} h\,d\nu - \int_{\mathbb{R}} h\,d\mu \right|$. It is trivial that $d_{\mathscr{H}}$ is a metric on the set of probability distributions on $\mathbb{R}$ and we also write $d_{\mathscr{H}}(X,Y) = d_{\mathscr{H}}(\mu,\nu)$ if $X \sim \mu, Y \sim \nu$. For example, $\mathscr{H}_1 := \{\mathbf{I}_A : A \in \mathscr{B}(\mathbb{R})\}$, $\mathscr{H}_2 := \{\mathbf{I}_{(-\infty,x]} : x \in \mathbb{R}\}$ and $\mathscr{H}_3 := \{f : \|f\|_\infty + \|f'\|_\infty \leq 1\}$ are three separating classes; $d_{\mathscr{H}_1}, d_{\mathscr{H}_2}, d_{\mathscr{H}_3}$ are known as the *total-variation distance* d_{TV}, the *Kolmogorov distance* d_{Kol} and the *Fortet-Mourier distance* d_{FM} respectively. It is known from [35, Section 11.3] that d_{FM} induces the same topology as the weak convergence of probability measures. It is clear that if μ_n, μ are probability measures on $\mathbb{R}$, then the following implications hold in view of the Portemanteau's theorem:

$$\lim_{n\to+\infty} d_{\mathrm{TV}}(\mu_n,\mu) = 0 \quad\Longrightarrow\quad \lim_{n\to+\infty} d_{\mathrm{Kol}}(\mu_n,\mu) = 0 \quad\Longrightarrow\quad \lim_{n\to+\infty} d_{\mathrm{FM}}(\mu_n,\mu) = 0\,.$$

Another important distance that we will use is the *Wasserstein distance* d_{W}, defined by $d_{\mathrm{W}}(\mu,\nu) := \sup\{\int_{\mathbb{R}} h\,d\nu - \int_{\mathbb{R}} h\,d\mu \;:\; \|h'\|_\infty \leq 1\}$ for two probability measures μ and ν on $\mathbb{R}$. Note d_{FM} is sometimes called the "bounded Wasserstein distance".

[1]We say F has at most *polynomial growth at infinity*, if there are some universal constants $C > 0$ and $d \in N^*$ such that $|F(x)| \leq C + C|x|^d$ for each $x \in \mathbb{R}$. The proof of Lemma 2.1.1 is omitted here, for it is exactly the same as in previous discussion.

[2]that is, for any two different probability measures μ, ν on $\mathbb{R}$, there exists $f \in \mathscr{H}$ such that $\int_{\mathbb{R}} f\,d\nu \neq \int_{\mathbb{R}} f\,d\mu$.

Ingeniously, Charles Stein introduced the following equation:

$$f'(x) - xf(x) = h(x) - \mathbb{E}[h(N)]\,, \tag{2.1.1}$$

where $N \sim \mathcal{N}(0,1)$ and $h : \mathbb{R} \to \mathbb{R}$ is a measurable function satisfying $h(N) \in L^1(\mathbb{P})$. The equation (2.1.1) is known as the *Stein's equation with unknown h*. We call f a solution to (2.1.1), if it is *absolutely continuous* and one version of f' satisfies (2.1.1) everywhere.

Let us now stop to digest the ingenuity of (2.1.1): suppose that we can solve (2.1.1) and get nice properties of f in terms of those of h, then we replace the dummy variable x in (2.1.1) by the random variable Z, so that after taking expectation on both sides, we get

$$\mathbb{E}[f'(Z) - Zf(Z)] = \mathbb{E}[h(Z) - h(N)]\,. \tag{2.1.2}$$

That is to say, in order to get uniform control of the right-hand side, one can instead try to uniformly control the left-hand side, which fits the heuristic after Stein's lemma. In essence, Stein's method replaces the complex-valued characteristic function by the above real characterizing equation, and many coupling methods have been developed so far to deal with the left-hand side of (2.1.2), see aforementioned references.

The following lemma gives an explicit form for solutions to Stein's equation (2.1.1).

Lemma 2.1.2. Let N, h be given as before, the solutions f to (2.1.1) are given by

$$f(x) = ce^{x^2/2} + e^{x^2/2}\int_{-\infty}^{x} \Big[h(y) - \mathbb{E}\big(h(N)\big)\Big]e^{-y^2/2}\,dy\,, \quad x \in \mathbb{R}\,,$$

where $c \in \mathbb{R}$. In particular, the function

$$f_h(x) := e^{x^2/2}\int_{-\infty}^{x} \Big[h(y) - \mathbb{E}\big(h(N)\big)\Big]e^{-y^2/2}\,dy\,, \quad x \in \mathbb{R} \tag{2.1.3}$$

is the unique solution to (2.1.1) verifying $\lim_{|x|\to+\infty} \exp(-x^2/2)f(x) = 0$. We call f_h the Stein's solution to (2.1.1).

Proof. Noticing that $e^{x^2/2}\dfrac{d}{dx}(f(x)e^{-x^2/2}) = f'(x) - xf(x) = h(x) - \mathbb{E}[h(N)]$, we have[3]

$$f(x)e^{-x^2/2} - f(0) = \int_0^x \Big[h(y) - \mathbb{E}[h(N)]\Big]e^{-y^2/2}\,dy\,.$$

By the dominated convergence theorem and

$$\int_{\mathbb{R}} \Big(h(y) - \mathbb{E}[h(N)]\Big)\, e^{-y^2/2}\,dy = 0\,, \tag{2.1.4}$$

we have $c := \lim_{x\to-\infty} e^{-x^2/2}f(x) = -\lim_{x\to+\infty} e^{-x^2/2}f(x)$. Thus, for any $x \in \mathbb{R}$,

$$f(x)e^{-x^2/2} - c = \int_{-\infty}^{x} \Big[h(y) - \mathbb{E}[h(N)]\Big]e^{-y^2/2}\,dy\,,$$

which gives us the desired form for the solutions to (2.1.1). And the rest is trivial. □

[3] We follow the convention $\int_0^x = -\int_x^0$ if $x < 0$

Now starting with the expression in (2.1.3), we study the properties of the Stein's solution f_h when h is 1-Lipschitz function or h is merely bounded measurable:

(i) Suppose that $h : \mathbb{R} \to [0,1]$ is measurable, then it follows from Stein's equation that

$$|f_h'(x)| \leq |h(x) - \mathbb{E}[h(N)]| + |xf_h(x)| \leq 1 + |xf_h(x)|$$

and from (2.1.4) that $\int_{-\infty}^x (h(y) - \mathbb{E}[h(N)])e^{-y^2/2}dy = -\int_x^\infty (h(y) - \mathbb{E}[h(N)])e^{-y^2/2}dy$. Thus, $|f_h(x)| \leq e^{x^2/2}\int_{|x|}^\infty e^{-y^2/2}\,dy \leq \sqrt{\pi/2}$, and $|xf_h(x)| \leq |x| \cdot e^{x^2/2} \cdot \int_{|x|}^\infty e^{-y^2/2}\,dy \leq 1$. So we can conclude that f_h is 2-Lipschitz and uniformly bounded by $\sqrt{\pi/2}$.

(ii) Suppose that $h : \mathbb{R} \to \mathbb{R}$ is K-Lipschitz for some $K \in (0, +\infty)$, then $\|f_h\|_\infty \leq 2K$, $\|f_h'\|_\infty \leq \sqrt{2/\pi}K$ and $\|f_h''\|_\infty \leq 2K$. The proof for these bounds involves a careful analysis, here we just give a very sketchy one and refer interested readers to [19, Lemma 2.4]: first denote the standard Gaussian density, distributional functions by ϕ, Φ respectively, one starts with $h(x) - \mathbb{E}[h(N)] = \int_\mathbb{R}[h(x) - h(u)]\phi(u)\,du$, then it follows from an application of Fubini's theorem that $h(x) - \mathbb{E}[h(N)] = \int_{-\infty}^x h'(t)\Phi(t)\,dt - \int_x^\infty h'(t)[1 - \Phi(t)]\,dt$, together with which (2.1.3) implies

$$\begin{aligned} f_h(\mathrm{w}) &= e^{\mathrm{w}^2/2}\int_{-\infty}^{\mathrm{w}}\left(\int_{-\infty}^x h'(t)\Phi(t)\,dt - \int_x^\infty h'(t)[1-\Phi(t)]\,dt\right)e^{-x^2/2}\,dx \\ &= -\sqrt{2\pi}e^{\mathrm{w}^2/2}(1-\Phi(\mathrm{w}))\int_{-\infty}^{\mathrm{w}} h'(t)\Phi(t)\,dt - \sqrt{2\pi}e^{\mathrm{w}^2/2}\Phi(\mathrm{w})\int_{\mathrm{w}}^\infty h'(t)[1-\Phi(t)]\,dt \end{aligned}$$

and

$$\begin{aligned} f_h'(\mathrm{w}) = \mathrm{w}f_h(\mathrm{w}) + h(\mathrm{w}) - \mathbb{E}[h(N)] =& \{1 - \sqrt{2\pi}\mathrm{w}e^{\mathrm{w}^2/2}[1-\Phi(\mathrm{w})]\}\int_{-\infty}^{\mathrm{w}} h'(t)\Phi(t)\,dt \\ &- \{1 + \sqrt{2\pi}\mathrm{w}e^{\mathrm{w}^2/2}\Phi(\mathrm{w})\}\int_{\mathrm{w}}^\infty h'(t)[1-\Phi(t)]\,dt\,. \end{aligned}$$

So the first two bounds follow from some standard Gaussian computations. Similar computations can be done for $\|f_h''\|_\infty$ once we derive $f_h''(\mathrm{w}) = (1+\mathrm{w}^2)f_h(\mathrm{w}) + \mathrm{w}(h(\mathrm{w}) - \mathbb{E}[h(N)]) + h'(\mathrm{w})$ using (2.1.1).

(iii) Combining the above two points, we can also assert that if $h \in C_b^\infty(\mathbb{R})$ satisfies $0 \leq h \leq 1$, then $\|f_h\|_\infty \leq \sqrt{\pi/2}$, $\|f_h'\|_\infty \leq 2$ and $\|f_h''\|_\infty \leq 2\|h'\|_\infty < +\infty$.

Now we give Stein's bounds based on the above discussions.

Proposition 2.1.1. Let F be a real integrable random variable and $N \sim \mathcal{N}(0,1)$, then

(1) $d_{\mathrm{TV}}(F,N) \leq \sup\left|\mathbb{E}[\varphi'(F) - F\varphi(F)]\right|$, where the supremum runs over all $\varphi \in C^2(\mathbb{R})$ with $\|\varphi\|_\infty \leq \sqrt{\pi/2}$, $\|\varphi'\|_\infty \leq 2$ and $\|\varphi''\|_\infty < +\infty$.

(2) $d_{\mathrm{W}}(F,N) \leq \sup\left|\mathbb{E}[\varphi'(F) - F\varphi(F)]\right|$, where the supremum runs over all $\varphi \in C^2(\mathbb{R})$ with $\|\varphi\|_\infty \leq 2$, $\|\varphi'\|_\infty \leq \sqrt{2/\pi}$ and $\|\varphi''\|_\infty \leq 2$.

Proof. It follows first from *Lusin's theorem*[4] that

$$\sup_{\substack{h \text{ measurable} \\ 0\le h\le 1}} \left|\mathbb{E}\big(h(N)-h(F)\big)\right| = \sup_{\substack{h\in C(\mathbb{R}) \\ 0\le h\le 1}} \left|\mathbb{E}\big(h(N)-h(F)\big)\right|.$$

And for $h : \mathbb{R} \to [0,1]$ continuous, by standard mollification procedure, one can construct $h_\varepsilon \in C_b^\infty(\mathbb{R})$ such that $0 \le h_\varepsilon \le 1$ and $h_\varepsilon \to h$ pointwise, as $\varepsilon \downarrow 0$. So

$$d_{\mathrm{TV}}(F,N) = \sup_{\substack{h\in C_b^\infty(\mathbb{R}) \\ 0\le h\le 1}} \left|\mathbb{E}\big(h(N)-h(F)\big)\right| \le \sup_{\substack{\|\varphi\|_\infty\le\sqrt{\pi/2} \\ \|\varphi'\|_\infty\le 2 \\ \|\varphi''\|_\infty<+\infty}} \left|\mathbb{E}[\varphi'(F)-F\varphi(F)]\right|$$

where the last inequality follows from point (iii) on last page and (2.1.2), this proves (1) while (2) follows from point (ii) on last page and (2.1.2). □

Let us now look at an easy example.

Example 2.1.1. Let $\mathbf{Y} = (Y_i, i \in \mathbb{N})$ be i.i.d symmetric Bernoulli random variables, then $F_n := n^{-1/2}(Y_1 + \ldots + Y_n)$ converges in law to the standard Gaussian, according to the classic central limit theorem. Now fix $\varphi \in C^2(\mathbb{R})$ with $\|\varphi\|_\infty \le 2$, $\|\varphi'\|_\infty \le \sqrt{2/\pi}$ and $\|\varphi''\|_\infty \le 2$, then we have

$$\begin{aligned}
\mathbb{E}[F_n\varphi(F_n)] &= \sqrt{n}\mathbb{E}[Y_1\varphi(F_n)] \quad \text{using symmetry} \\
&= \frac{\sqrt{n}}{2}\mathbb{E}\Big[\varphi\big(\frac{1+Y_2+\ldots+Y_n}{\sqrt{n}}\big) - \varphi\big(\frac{-1+Y_2+\ldots+Y_n}{\sqrt{n}}\big)\Big] \quad \text{using independence} \\
&= \frac{\sqrt{n}}{2}\mathbb{E}\Big[\varphi\big(\frac{1+Y_2+\ldots+Y_n}{\sqrt{n}}\big) - \varphi(F_n) + \varphi(F_n) - \varphi\big(\frac{-1+Y_2+\ldots+Y_n}{\sqrt{n}}\big)\Big] \\
&= \mathbb{E}[\varphi'(F_n)] + \frac{\sqrt{n}}{2}R_n \quad \text{with } |R_n| \le 2\|\varphi''\|_\infty n^{-1}
\end{aligned}$$

where the last line follows from the usual Taylor expansion. Thus, it follows from Proposition 2.1.1 that[5] $d_{\mathrm{W}}(F_n, \mathcal{N}(0,1)) \le 2n^{-1/2}$.

In the above toy-example, independence and symmetry within the structure contribute to the easy proof. Nevertheless, the strategy of combining an application of Taylor expansion with Stein's equation is usually effective in practice, see *e.g.* our Proposition 2.1.2. Stein's ideas have contributed to many excellent results and solutions to numerous problems, such as local dependence, minimal spanning trees, concentration inequalities and many others, see *e.g.* Barbour and Chen's Volume [12] and Chatterjee's ICM survey [20].

Since Stein's introduction of his method, there have been many successful uses of the multivariate version, see *e.g.* [9, 10, 38] that initiated Barbour's *generator approach*. Roughly

[4] Let $\mu = (\mathbb{P}\circ F^{-1} + \mathbb{P}\circ N^{-1})/2$, then for h measurable with values in $[0,1]$, by appropriate application of Lusin's Theorem (*e.g.* see the *Red Rudin*) one can find a sequence of continuous functions h_n such that $0 \le h_n \le 1$ and $h_n \to h$ μ-a.s., thus $h_n \to h$ $\mathbb{P}\circ F^{-1}$-a.s. and $\mathbb{P}\circ N^{-1}$–a.s..

[5] By explicit calculation, $\mathbb{E}[F_n^4] - 3 = -2n^{-1}$, so we obtain the fourth moment bound $d_{\mathrm{W}}(F_n, \mathcal{N}(0,1)) \le \sqrt{2(3-\mathbb{E}[F_n^4])}$.

speaking, the generator approach uses the properties of Markov processes to deduce and solve Stein's equation associated to the invariant measure of those Markov processes. Note that the standard Gaussian distribution on $\mathbb{R}^k$ ($k \in \mathbb{N}$) is the invariant distribution of the k-dimensional Ornstein-Uhlenbeck process.

Let's first introduce some further notation, before we present the multivariate Stein's method.

Notation C. For $x = (x_1, \dots, x_d)^T$, $y = (y_1, \dots, y_d)^T \in \mathbb{R}^d$, we denote by $\|x\|$ and $\langle x, y\rangle$ the *Euclidean norm* of x and the *scalar product* of x, y respectively; and for a matrix $A \in \mathbb{R}^{d\times d}$, we denote by $\|A\|_{\mathrm{op}}$ the *operator norm* induced by the Euclidean norm, i.e., $\|A\|_{\mathrm{op}} := \sup\{\|Ax\| : \|x\| = 1\}$. More generally, for a k-multilinear form $\psi : (\mathbb{R}^d)^k \to \mathbb{R}$, $k \in \mathbb{N}$, we define the *operator norm*

$$\|\psi\|_{\mathrm{op}} := \sup\Big\{|\psi(u_1, \dots, u_k)| \,:\, u_j \in \mathbb{R}^d,\ \|u_j\| = 1,\ j = 1, \dots, k\Big\}\,.$$

Recall that for a function $h : \mathbb{R}^d \to \mathbb{R}$, its (minimal) Lipschitz constant $M_1(h)$ is given by

$$M_1(h) := \sup_{x\neq y} \frac{|h(x) - h(y)|}{\|x - y\|} \in [0, \infty]\,.$$

If h is differentiable, then $M_1(h) = \sup_{x\in\mathbb{R}^d} \|Dh(x)\|_{\mathrm{op}}$. For random vectors X, Y in $\mathbb{R}^d$, we define their Wasserstein distance by

$$d_{\mathrm{W}}(X, Y) := \sup\Big\{\mathbb{E}[h(X) - h(Y)] \,:\, M_1(h) \le 1\Big\}.$$

For $k \in \mathbb{N}$ and a $(k-1)$-times differentiable function $h : \mathbb{R}^d \to \mathbb{R}$, we set

$$M_k(h) := \sup_{x\neq y} \frac{\|D^{k-1}h(x) - D^{k-1}h(y)\|_{\mathrm{op}}}{\|x - y\|}\,,$$

viewing the $(k-1)$-th derivative $D^{k-1}h$ of h at any point x as a $(k-1)$-multilinear form. Then, if h is k-times differentiable, we have $M_k(h) = \sup\{\|D^k h(x)\|_{\mathrm{op}} : x \in \mathbb{R}^d\}$.

Recall that, for two matrices $A, B \in \mathbb{R}^{d\times d}$, their *Hilbert-Schmidt inner product* is defined by

$$\langle A, B\rangle_{\mathrm{HS}} := \mathrm{Tr}(AB^T) = \mathrm{Tr}(BA^T) = \mathrm{Tr}(B^T A) = \sum_{i,j=1}^d A_{i,j}B_{i,j}\,.$$

Thus, $\langle\cdot,\cdot\rangle_{\mathrm{HS}}$ is just the standard inner product on $\mathbb{R}^{d\times d} \cong \mathbb{R}^{d^2}$. The corresponding *Hilbert-Schmidt norm* will be denoted by $\|\cdot\|_{\mathrm{HS}}$. With this notion at hand, following [23] and [58], for $k = 2$, we define $\widetilde{M}_2(h) := \sup\{\|\mathrm{Hess}\, h(x)\|_{\mathrm{HS}} : x \in \mathbb{R}^d\}$, where $\mathrm{Hess}\, h$ is the *Hessian matrix* associated with h. Note for a symmetric matrix $A \in \mathbb{R}^{d\times d}$ with eigenvalues $\lambda_1(A) \le \dots \le \lambda_d(A)$, one has $\|A\|_{\mathrm{HS}}^2 = \lambda_1(A)^2 + \dots + \lambda_d(A)^2 \le d\|A\|_{\mathrm{op}}^2$. From this, it follows immediately that $\widetilde{M}_2(h) \le \sqrt{d}\, M_2(h)$. For later use, we define here another two distributional metrics: for random vectors X, Y in $\mathbb{R}^d$,

$$d_2(X, Y) := \sup\Big\{\mathbb{E}[h(X) - h(Y)] \,:\, M_1(h) \le 1, M_2(h) \le 1\Big\}$$

and

$$d_3(X, Y) := \sup\Big\{\mathbb{E}[h(X) - h(Y)] \,:\, M_1(h) \le 1, M_2(h) \le 1, M_3(h) \le 1\Big\}.$$

With the above notation, let us first present the multivariate Stein's lemma.

Lemma 2.1.3. Fix an integer $d \geq 2$. Let C be a non-negative definite $d \times d$ matrix. Given a random vector $N = (N_1, \dots, N_d)^T$, then $N \sim \mathcal{N}(0, C)$ if and only if

$$\mathbb{E}[\langle N, \nabla f(N)\rangle_{\mathbb{R}^d}] = \mathbb{E}[\langle C, \operatorname{Hess} f(N)\rangle_{\mathrm{HS}}] \tag{2.1.5}$$

for every C^2 function $f : \mathbb{R}^d \to \mathbb{R}$ with bounded first and second derivatives.

Proof. If $N \sim \mathcal{N}(0, C)$, then the equation (2.1.5) is a simple consequence of the integration by parts formula. Now assume (2.1.5) and fix an arbitrary vector $\Lambda = (\lambda_1, \dots, \lambda_d)^T \in \mathbb{R}^d$, we define the function $f(x) = \sin(\langle \Lambda, x\rangle)$, which is of class C^2 with bounded derivatives. Then (2.1.5) implies that

$$\mathbb{E}[\langle \Lambda, N\rangle \cos(\langle \Lambda, N\rangle)] = -\langle C, \Lambda\Lambda^T\rangle_{\mathrm{HS}}\mathbb{E}[\sin(\langle \Lambda, N\rangle)]$$

and similarly we get also

$$\mathbb{E}[\langle \Lambda, N\rangle \sin(\langle \Lambda, N\rangle)] = \langle C, \Lambda\Lambda^T\rangle_{\mathrm{HS}}\mathbb{E}[\cos(\langle \Lambda, N\rangle)]\,.$$

With the same arguments as in Section 2.1.1, we can build an ODE for $G(b) = \mathbb{E}[e^{ib\langle \Lambda, N\rangle}]$, that is,

$$G'(b) = -b\langle C, \Lambda\Lambda^T\rangle_{\mathrm{HS}} G(b) \quad \text{subject to } G(0) = 1.$$

Solving this ODE gives us $G(b) = \exp\big(-\langle C, \Lambda\Lambda^T\rangle_{\mathrm{HS}} b^2/2\big)$. It follows that $\langle \Lambda, N\rangle$ is a centered Gaussian random variable with variance $\langle C, \Lambda\Lambda^T\rangle_{\mathrm{HS}}$, thus N is a centered Gaussian vector with covariance matrix C. □

Starting from the multivariate Stein's lemma, one can build the multivariate Stein's equation

$$\langle x, \nabla f(x)\rangle_{\mathbb{R}^d} - \langle \operatorname{Hess} f(x), C\rangle_{\mathrm{HS}} = h(x) - \mathbb{E}[h(Z)]\,, \quad Z \sim \mathcal{N}(0, C)\,. \tag{2.1.6}$$

If $h \in C^2(\mathbb{R}^d)$ verifies[6] $M_1(h) + M_2(h) < +\infty$, then the function

$$f_h(x) := \int_0^1 \frac{1}{2t}\mathbb{E}[h(\sqrt{t}x + \sqrt{1-t}Z) - h(Z)]\,dt \tag{2.1.7}$$

belongs to $C^2(\mathbb{R}^d)$ and solves the Stein's equation (2.1.6); and here we collect several useful facts on Stein's solution:

★ for $r = 1, 2, 3$,

$$M_r(f_h) \leq r^{-1} M_r(h) \quad \text{and} \quad \widetilde{M}_2(f_h) \leq \frac{1}{2}\widetilde{M}_2(h)\,.$$

In particular, if C is positive definite, then

$$\widetilde{M}_2(f_h) \leq \sqrt{2/\pi}\,\|C^{-1/2}\|_{\mathrm{op}} M_1(h) \quad \text{and} \quad M_3(f_h) \leq \frac{\sqrt{2\pi}}{4}\|C^{-1/2}\|_{\mathrm{op}} M_2(h)\,.$$

For the detailed proofs, one can refer to Section 3.1 of Döbler's dissertation [28].

[6] To guarantee that $f_h(x)$ is well defined, it is enough to assume h is Lipschitz continuous, see [28, Lemma 3.1.1]. If in addition C is positive definite, with h Lipschitz, it holds true that f_h given in (2.1.7) belongs to $C^2(\mathbb{R}^d)$ and satisfies $\widetilde{M}_2(f_h) \leq \sqrt{d} M_1(h)\|C^{-1}\|_{\mathrm{op}}\|C\|_{\mathrm{op}}^{1/2}$, see [67, Proposition 4.3.2].

2.1.2 Stein's method of exchangeable pairs

The exchangeable pairs approach within Stein's method was first used in the paper [27] which, however, attributed the method to Charles Stein. Later, this technique was presented in a systematic way in Stein's monograph [101]. We recall that a pair (X, X') of random elements on a common probability space is said to be *exchangeable*, if (X, X') has the same distribution as (X', X). In the book [101, Lecture III], it is highlighted that a given real random variable W is close in distribution to a standard normal variable N, whenever one can construct an exchangeable pair (W, W') such that W' is close to W in *some* sense and such that the *linear regression property*

$$\mathbb{E}[W' - W \,\big|\, W] = -\lambda W \tag{2.1.8}$$

is satisfied for some small $\lambda > 0$ and $\mathbf{Var}(\frac{1}{2\lambda}\mathbb{E}[(W' - W)^2|W])$ is small. Note the relation (2.1.8) forces W to be centered.

Theorem 2.1.1 (C. Stein, 1986)**.** Let (W, W') be an exchangeable pair of random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbf{Var}(W) = 1$, (2.1.8) holds for some $\lambda > 0$ and $W \in L^3(\mathbb{P})$, then we have

$$d_{\mathrm{W}}(W, N) \leq \sqrt{2/\pi}\sqrt{\mathbf{Var}\left(\frac{1}{2\lambda}\mathbb{E}[(W - W')^2|W]\right)} + \frac{1}{2\lambda}\mathbb{E}[|W - W'|^3]\,,$$

where $N \sim \mathcal{N}(0, 1)$.

Proof. Let $f : \mathbb{R} \to \mathbb{R}$ belong to $C^2(\mathbb{R})$ with $\|f\|_\infty \leq 2$, $\|f'\|_\infty \leq \sqrt{2/\pi}$ and $\|f''\|_\infty \leq 2$, then it is clear that $f(W')W$ and $f(W)W$ are integrable. It follows from the exchangeability that $\mathbb{E}[(W - W')f(W)] = \mathbb{E}[(W' - W)f(W')]$, from which we get $\mathbb{E}[(W - W')[f(W) + f(W')]] = 0$. So

$$\begin{aligned}
0 &= \mathbb{E}[(W - W')[-f(W) + f(W')]] + 2\mathbb{E}[(W - W')f(W)] \\
&= \mathbb{E}[(W - W')[-f(W) + f(W')]] + 2\mathbb{E}[f(W)\mathbb{E}(W - W'|W)] \\
&= \mathbb{E}[(W - W')[-f(W) + f(W')]] + 2\lambda\mathbb{E}[f(W)W] \quad \text{by (2.1.8)}
\end{aligned}$$

that is, $\mathbb{E}[(W' - W)[f(W') - f(W)]] = 2\lambda\mathbb{E}[f(W)W]$; and particularly, $\mathbb{E}[(W' - W)^2] = 2\lambda$. Therefore,

$$\begin{aligned}
&2\lambda\mathbb{E}[f(W)W - f'(W)] = \mathbb{E}\Big((W' - W)[f(W') - f(W)] - 2\lambda f'(W)\Big) \\
&= \mathbb{E}\Big(f'(W)[(W' - W)^2 - 2\lambda]\Big) + \mathbb{E}\left[(W' - W)^3\int_0^1 (1 - t)f''(W + t(W' - W))\,dt\right] \\
&= \mathbb{E}\Big[f'(W)\Big(\mathbb{E}[(W' - W)^2|W] - 2\lambda\Big)\Big] + \mathbb{E}\left[(W' - W)^3\int_0^1 (1 - t)f''(W + t(W' - W))\,dt\right],
\end{aligned}$$

which together with Proposition 2.1.1 gives us the desired bound. □

In recent years, the method of exchangeable pairs has been generalized for other distributions and multi-dimensional settings in many papers like [93, 94, 100, 90, 23, 21, 22, 29, 36], to name a few.

Moreover, the articles [57, 23, 58, 33] develop versions of the exchangeable pairs method suitable for situations, in which one can construct a continuous family of exchangeable pairs. By their continuity assumptions, these papers succeed in reducing the order of smoothness of test functions and hence in obtaining bounds in more sophisticated probabilistic distances. For instance, the bounds from [57] are expressed in terms of the total-variation distance. It is this framework of exchangeable pairs that is most closely related to the variant of the method developed in the present thesis. In contrast to the quoted papers, however, our abstract results on exchangeable pairs do not make such strong continuity assumptions and hence, allow us to deal with the inherent discreteness of the Poisson space and Rademacher functionals, which does not even allow for convergence in the total-variation distance in general.

The following two propositions are generalizations of Meckes' theorems [57, 58].

Proposition 2.1.2 (C. Döbler, A. Vidotto and G. Zheng, 2017)**.** Let F and a family of real random variables $(F_t)_{t\geq 0}$ be defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $F_t \overset{law}{=} F$ for every $t \geq 0$. Assume that $F \in L^4(\Omega, \mathcal{G}, \mathbb{P})$ for some σ-algebra $\mathcal{G} \subset \mathcal{F}$ and that in $L^1(\mathbb{P})$,

(a) $\lim_{t\downarrow 0} \frac{1}{t}\, \mathbb{E}[F_t - F|\mathcal{G}] = -\lambda F$ for some $\lambda > 0$,

(b) $\lim_{t\downarrow 0} \frac{1}{t}\, \mathbb{E}[(F_t - F)^2|\mathcal{G}] = (2\lambda + S)\mathbf{Var}(F)$ for some centered[7] random variable S;

(c) and $\lim_{t\downarrow 0} \frac{1}{t}\, \mathbb{E}[(F_t - F)^4] = \rho(F)\mathbf{Var}(F)^2$ for some $\rho(F) \geq 0$.

Then, with $N \sim \mathcal{N}(0, \mathbf{Var}(F))$, we have

$$d_{\mathrm{W}}(F, N) \leq \frac{\sqrt{\mathbf{Var}(F)}}{\lambda\sqrt{2\pi}}\mathbb{E}[|S|] + \frac{\sqrt{2\mathbf{Var}(F)}}{3\sqrt{\lambda}}\sqrt{\rho(F)}\,.$$

If $\rho(F) = 0$, we have $d_{\mathrm{TV}}(F, N) \leq \lambda^{-1}\mathbb{E}[|S|]$.

Proof. Assume first that $\mathbf{Var}(F) = 1$ and fix an arbitrary $g \in C^2(\mathbb{R})$ with $\|g\|_\infty \leq c_1$, $\|g'\|_\infty \leq c_2$ and $\|g''\|_\infty \leq c_3$ for some $c_1, c_2, c_3 \in (0, +\infty)$. Let $G : \mathbb{R} \to \mathbb{R}$ be a differentiable function such that $G' = g$. Then due to $F_t \overset{law}{=} F$ and $F \in L^4(\mathbb{P})$, one has $0 = \mathbb{E}[G(F_t) - G(F)] = \mathbb{E}[g(F)(F_t - F) + \frac{1}{2}g'(F)(F_t - F)^2] + \mathbb{E}[R_g]$ with $|R_g| \leq \frac{1}{6}\|g''\|_\infty\,|F_t - F|^3$. It follows that

$$0 = \mathbb{E}\left[g(F) \times \frac{1}{t}\,\mathbb{E}[F_t - F|\mathcal{G}]\right] + \frac{1}{2}\,\mathbb{E}\left[g'(F) \times \frac{1}{t}\,\mathbb{E}[(F_t - F)^2|\mathcal{G}]\right] + \frac{1}{t}\,\mathbb{E}[R_g].$$

By assumption (c) and as $t \downarrow 0$,

$$\left|\frac{\mathbb{E}[R_g]}{t}\right| \leq \frac{\|g''\|_\infty \mathbb{E}[|F_t - F|^3]}{6t} \leq \frac{\|g''\|_\infty}{6}\sqrt{\frac{1}{t}\mathbb{E}[(F_t - F)^2]}\sqrt{\frac{1}{t}\mathbb{E}[(F_t - F)^4]} \to \frac{\|g''\|_\infty\sqrt{2\lambda\rho(F)}}{6}\,.$$

[7] $\mathbb{E}[S] = 0$ is not restrictive at all. In fact, in our future applications, the convergence in (a) often takes place in $L^2(\mathbb{P})$, so the exchangeability together with condition (a) would imply $t^{-1}\mathbb{E}[(F_t - F)^2] = -2t^{-1}\,\mathbb{E}[(F_t - F)F] = -2\mathbb{E}[Ft^{-1}\mathbb{E}(F_t - F|F)] \to 2\lambda\mathbf{Var}(F)$, as $t \downarrow 0$.

Therefore as $t \downarrow 0$, assumptions (a) and (b) imply that[8]

$$0 = \lambda\mathbb{E}[g'(F) - Fg(F)] + \frac{1}{2}\,\mathbb{E}[g'(F)S] + \lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}[R_g]\,, \tag{2.1.9}$$

with $\left|\lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}[R_g]\right| \leq \sqrt{2\lambda}\,\sqrt{\rho(F)}\,\|g''\|_\infty/6$. Plugging this into the Stein's equation, we deduce from Proposition 2.1.1 that

$$\begin{aligned} d_{\mathrm{W}}(F,N) &\leq \sup_{\substack{\|g'\|_\infty \leq \sqrt{2/\pi} \\ \|g''\|_\infty\leq 2}} \Big|\mathbb{E}[g'(F) - Fg(F)]\Big| \\ &\leq \sup_{\substack{\|g'\|_\infty \leq \sqrt{2/\pi} \\ \|g''\|_\infty\leq 2}} \left(\frac{\|g'\|_\infty}{2\lambda}\mathbb{E}[|S|] + \left|\frac{1}{\lambda}\lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}[R_g]\right|\right) \\ &\leq \frac{1}{\lambda\sqrt{2\pi}}\mathbb{E}[|S|] + \frac{\sqrt{2}}{3\sqrt{\lambda}}\sqrt{\rho(F)}\,. \end{aligned}$$

Let us now consider the case where $\rho(F) = 0$, then the limit in (2.1.9) is zero and

$$\mathbb{E}\big[-g'(F) + Fg(F)\big] = \frac{1}{2\lambda}\,\mathbb{E}[g'(F)S]\,,$$

because $\|g''\|_\infty < +\infty$. Thus, it follows from Proposition 2.1.1 that $d_{\mathrm{TV}}(F,N) \leq \lambda^{-1}\mathbb{E}[|S|]$.

The general case follows from the facts that $d_{\mathrm{W}}(F,N) = \sigma\, d_{\mathrm{W}}(F/\sigma, N/\sigma)$ and $d_{\mathrm{TV}}(F,N) = d_{\mathrm{TV}}(F/\sigma, N/\sigma)$ for $\sigma > 0$. □

Remark 2.1.1. The case where $\rho(F) = 0$ in the above proposition is studied in E. Meckes' dissertation [57], in which Meckes further developed an earlier idea in Stein's technical report [102]. She called it an infinitesimal version of exchangeable pairs. Her use of a family of exchangeable pairs enables her to obtain nice bound in total-variation distance, which is beyond the reach of the usual exchangeable pair method, see *e.g.* Proposition 2.1.1. Note that in Meckes' formulation, she required the exchangeability of (F, F_t), which is not our case. Our consideration is motivated by Röllin's short note [94].

As one will see shortly, the constant $\rho(F)$ in Proposition 2.1.2 is often nonzero in our applications on Poisson Wiener chaos and Rademacher chaos; and in these situations, we will look at the infinitesimal version of exchangeable pairs as an efficient way of approximation.

The following result is a multidimensional generalization of Proposition 2.1.2, proved in **[P4]**.

Proposition 2.1.3 (C. Döbler, A. Vidotto and G. Zheng, 2017)**.** For each $t > 0$, let (F, F_t) be an exchangeable pair of centered d-dimensional random vectors defined on a common probability space. Let $\mathscr{G}$ be a σ-algebra that contains $\sigma\{F\}$. Assume that $\Lambda \in \mathbb{R}^{d\times d}$ is an invertible deterministic matrix and Σ is a symmetric, non-negative definite deterministic matrix such that

[8]The equation (2.1.9) shall be understood as follows: the limit $\lim_{t\downarrow 0} t^{-1}\mathbb{E}[R_g]$ exists and is equal to $-\lambda\mathbb{E}[g'(F) - Fg(F)] - \frac{1}{2}\,\mathbb{E}[g'(F)S]$, bounded by $\sqrt{2\lambda}\,\sqrt{\rho(F)}\,\|g''\|_\infty/6$.

(a) $\lim_{t\downarrow 0} \frac{1}{t}\,\mathbb{E}[F_t - F|\mathscr{G}] = -\Lambda F$ in $L^1(\Omega)$;

(b) $\lim_{t\downarrow 0} \frac{1}{t}\,\mathbb{E}[(F_t - F)(F_t - F)^T|\mathscr{G}] = 2\Lambda\Sigma + S$ in $L^1(\Omega; \|\cdot\|_{\mathrm{HS}})$ for some random matrix S with centered entries;

(c) for each $i \in \{1, \dots, d\}$, there exists some real number $\rho_i(F)$ such that

$$\lim_{t\downarrow 0} \frac{1}{t}\,\mathbb{E}[(F_{i,t} - F_i)^4] = \rho_i(F)$$

where $F_{i,t}$ (resp. F_i) stands for the ith coordinate of F_t (resp. F).

Then, we have, with $Z \sim \mathcal{N}(0, \Sigma)$,

$$d_3(F, Z) \le \frac{\|\Lambda^{-1}\|_{\mathrm{op}}\sqrt{d}}{4}\mathbb{E}[\|S\|_{\mathrm{HS}}] + \frac{\sqrt{d}\|\Lambda^{-1}\|_{\mathrm{op}}}{18}\sqrt{\sum_{i=1}^{d} 2\Lambda_{i,i}\Sigma_{i,i}}\sqrt{\sum_{i=1}^{d}\rho_i(F)}\,.$$

If in addition, Σ is positive definite, then

$$d_2(F, Z) \le \frac{\|\Lambda^{-1}\|_{\mathrm{op}}\|\Sigma^{-1/2}\|_{\mathrm{op}}}{\sqrt{2\pi}}\mathbb{E}[\|S\|_{\mathrm{HS}}] + \frac{\sqrt{2\pi}\|\Lambda^{-1}\|_{\mathrm{op}}\|\Sigma^{-1/2}\|_{\mathrm{op}}}{24}\sqrt{\sum_{i=1}^{d} 2\Lambda_{i,i}\Sigma_{i,i}}\sqrt{\sum_{i=1}^{d}\rho_i(F)}\,.$$

Remark 2.1.2. The particular case where $\rho_i(F) = 0$ for each $i \in [d]$ corresponds to Theorem 3 in Meckes' paper [58]. In this case, we have

$$d_2(F, Z) \le \frac{\|\Lambda^{-1}\|_{\mathrm{op}}\sqrt{d}}{4}\mathbb{E}[\|S\|_{\mathrm{HS}}]\,,$$

and if additionally Σ is positive definite, we have

$$d_{\mathrm{W}}(F, Z) \le \frac{\|\Lambda^{-1}\|_{\mathrm{op}}\|\Sigma^{-1/2}\|_{\mathrm{op}}}{\sqrt{2\pi}}\mathbb{E}[\|S\|_{\mathrm{HS}}]\,,$$

see also Proposition 9.1 in **[P3]**.

Proof of Proposition 2.1.3. By the same argument as in the proof of Theorem 3 in [58], we can assume $g \in C^\infty(\mathbb{R}^d)$ with $M_1(g) + M_2(g) + M_3(g) < +\infty$ and define

$$f(x) = \int_0^1 \frac{1}{2t}\Big(\mathbb{E}[g(\sqrt{t}\,x + \sqrt{1-t}\,N)] - \mathbb{E}[g(N)]\Big)\,dt\ ,$$

which is a solution to the following *Stein's equation*

$$\langle x, \nabla f(x)\rangle - \langle \mathrm{Hess}\, f(x), \Sigma\rangle_{\mathrm{HS}} = g(x) - \mathbb{E}[g(N)]\ . \tag{2.1.10}$$

It is known from Section 2.1.1 that for $r = 1, 2, 3$,

$$M_r(f) \le \frac{M_r(g)}{r}$$

and $\widetilde{M}_2(f) \le \frac{1}{2}\widetilde{M}_2(g)$. In particular, if Σ is positive definite, then $\widetilde{M}_2(f) \le \sqrt{2/\pi}\,\|\Sigma^{-1/2}\|_{\mathrm{op}}\, M_1(g)$ and $M_3(f) \le \sqrt{2\pi}\,\|\Sigma^{-1/2}\|_{\mathrm{op}}\, M_2(g)/4$.

Again, it follows from the same arguments as in [58] that

$$\begin{aligned} 0 = \frac{1}{t}\,\mathbb{E}\Big[\ \frac{1}{2}\Big\langle \mathrm{Hess}\, f(X), \Lambda^{-1}(F_t - F)(F_t - F)^T\Big\rangle_{\mathrm{HS}}\Big] \\ + \frac{1}{t}\mathbb{E}\Big[\,\langle \Lambda^{-1}(F_t - F), \nabla f(F)\rangle\Big] + \frac{1}{2t}\mathbb{E}[R], \end{aligned} \tag{2.1.11}$$

where R is the error in the Taylor approximation satisfying

$$|R| \le \frac{1}{3}\|\Lambda^{-1}\|_{\mathrm{op}}\|F_t - F\|^3\beta \le \frac{\sqrt{d}}{3}\|\Lambda^{-1}\|_{\mathrm{op}}\beta\sqrt{\sum_{i=1}^d (F_{i,t} - F_i)^2}\sqrt{\sum_{i=1}^d (F_{i,t} - F_i)^4}\ ,$$

where $\beta := \min\{M_3(g)/3,\ \sqrt{2\pi}\|\Sigma^{-1/2}\|_{\mathrm{op}}\, M_2(g)/4\}$, and the last inequality follows from the elementary inequality $\|x - y\|^2 \le \sqrt{d}\big(\sum_{i=1}^d (x_i - y_i)^4\big)^{1/2}$ for $x, y \in \mathbb{R}^d$.

Notice meanwhile that the assumptions (a) and (b) imply that the limit $t^{-1}\mathbb{E}[R]$, as $t \downarrow 0$, is well defined and

$$\begin{aligned} -\lim_{t\downarrow 0}\frac{1}{2t}\mathbb{E}[R] &= \mathbb{E}\Big[\langle \mathrm{Hess}\, f(F), \Sigma\rangle_{\mathrm{HS}} - \langle F, \nabla f(F)\rangle\Big] + \frac{1}{2}\mathbb{E}\Big[\langle \mathrm{Hess}\, f(F), \Lambda^{-1}S\rangle_{\mathrm{HS}}\Big] \\ &= \mathbb{E}[g(N) - g(F)] + \frac{1}{2}\mathbb{E}\Big[\langle \mathrm{Hess}\, f(F), \Lambda^{-1}S\rangle_{\mathrm{HS}}\Big]\ , \end{aligned}$$

where the last equality comes from the definition of Stein's equation. Moreover, by assumption (c) and the above inequality, we have

$$\begin{aligned} \Big|\lim_{t\to 0}\frac{1}{t}\,\mathbb{E}[R]\Big| &\le \frac{\sqrt{d}}{3}\|\Lambda^{-1}\|_{\mathrm{op}}\,\beta\,\sqrt{\lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}\sum_{i=1}^d (F_{i,t} - F_i)^2}\sqrt{\lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}\sum_{i=1}^d (F_{i,t} - F_i)^4} \\ &= \frac{\sqrt{d}}{3}\|\Lambda^{-1}\|_{\mathrm{op}}\,\beta\,\sqrt{\sum_{i=1}^d 2\Lambda_{i,i}\Sigma_{i,i}}\sqrt{\sum_{i=1}^d \rho_i(F)}, \end{aligned} \tag{2.1.12}$$

where the last equality follows from assumptions (b) and (c). To conclude our proof, it suffices to notice that

$$\Big|\mathbb{E}[\langle \mathrm{Hess}\, f(F), \Lambda^{-1}S\rangle_{\mathrm{HS}}]\Big| \le \min\left\{\frac{1}{2}\widetilde{M}_2(g),\ \sqrt{\frac{2}{\pi}}\,\|\Sigma^{-1/2}\|_{\mathrm{op}}\, M_1(g)\right\}\|\Lambda^{-1}\|_{\mathrm{op}}\,\mathbb{E}[\|S\|_{\mathrm{HS}}]\ .$$

□

2.2 Exchangeable pairs on chaoses

One goal of this section is to explain the following result: for each $t \in \mathbb{R}_+$, $\mathfrak{X} \in \{W, \eta, \mathbf{X}\}$, one can construct another random object $\mathfrak{X}^t$ such that $(\mathfrak{X}, \mathfrak{X}^t)$ is exchangeable and for any $F = \mathfrak{f}(\mathfrak{X}) \in L^2(\Omega, \sigma\{\mathfrak{X}\}, \mathbb{P})$, one has

$$P_t F = \mathbb{E}[f(\mathfrak{X}^t)|\sigma\{\mathfrak{X}\}]\,, \tag{2.2.1}$$

which is known as the Mehler's formula for the Ornstein-Uhlenbeck semigroup $(P_t, t \in \mathbb{R}_+)$. And the other goal is to connect these exchangeable pairs to the carré du champ, see Proposition 2.2.1.

Construction of W^t: this is indeed trivial, we put $W^t = e^{-t}W + \sqrt{1-e^{-2t}}W'$, with W' an *independent* copy of W. One can easily verify that W^t is a Gaussian random measure with the same intensity measure μ, and for the exchangeability of (W^t, W), it suffices to consider their finite-dimensional distributions, that is, let $A_1, \dots, A_m \in \mathscr{Z}_\mu$, then the following two Gaussian vectors

$$(W^t(A_1), \dots, W^t(A_m)) \quad \text{and} \quad (W(A_1), \dots, W(A_m))$$

are clearly exchangeable in view of their (joint) covariance matrix. Now let us verify (2.2.1): note first that by density argument, it is enough to assume $F = I_p^W(f)$ for some $f \in L_s^2(\mu^p)$ of the form (1.2.2), that is,

$$I_p^W(f) = \sum_{i_1,\dots,i_p=1}^m \beta_{i_1,\dots,i_p} \prod_{j=1}^p W(A_{i_j})\,,$$

where $m \in \mathbb{N}$, $A_1, \dots, A_m \in \mathscr{Z}_\mu$ are *pairwise disjoint*, and the coefficients $\beta_{i_1,\dots,i_p}$ are symmetric and vanish whenever any two of the indices $i_1, \dots, i_p$ are equal. We define

$$F_t := I_p^{W^t}(f) = \sum_{i_1,\dots,i_p=1}^m \beta_{i_1,\dots,i_p} \prod_{j=1}^p W^t(A_{i_j})$$

and by standard computation, $\mathbb{E}[F_t|\sigma\{W\}] = e^{-pt}F$. This finishes our verification of (2.2.1) in the case where $\mathfrak{X} = W$.

The above discussion gives us a natural exchangeable pair coupling on Gaussian Wiener chaos, that is, given $f \in L_s^2(\mu^p)$, $I_p^W(f)$ and $I_p^{W^t}(f)$ are exchangeable.

Construction of $\mathbf{X}^t$: let $\mathbf{X}'$ be an independent copy of $\mathbf{X}$ and $\Theta = (\theta_k, k \in \mathbb{N})$ be a sequence of *i.i.d.* standard exponential random variables such that $\mathbf{X}$, $\mathbf{X}'$ and Θ are independent. For each $t \in [0, +\infty)$, we define

$$X_k^t := X_k \mathbf{I}_{(\theta_k \geq t)} + X_k' \mathbf{I}_{(\theta_k < t)}\ .$$

It has been pointed out in [45] that $\mathbf{X}^t$ has the same distribution as $\mathbf{X}$, see also Remark 3.4 in [70] for the symmetric case. Assuming the exchangeability for now and writing $F = \mathfrak{f}(\mathbf{X})$ for some representative $\mathfrak{f} : \{\pm 1\}^{\mathbb{N}} \to \mathbb{R}^d$, we can set $F_t = \mathfrak{f}(\mathbf{X}^t)$. It is easy to see that the exchangeability can be passed to (F, F_t) now. If $F = (Q_{p_1}(f_1; \mathbf{Y}), \dots, Q_{p_d}(f_d; \mathbf{Y}))$, then we can write $F_t = (Q_{p_1}(f_1; \mathbf{Y}^t), \dots, Q_{p_d}(f_d; \mathbf{Y}^t))$ with $\mathbf{Y}^t$ the normalised version of $\mathbf{X}^t$ in the sense of (1.3.3).

Now let us verify the exchangeability. Note first that $\mathbf{X}^t$ is a sequence of independent Rademacher random variables for each $t \in [0, +\infty)$. For each $k \in \mathbb{N}$, it is easy to check that

$$\mathbb{P}(X_k^t = -1, X_k = 1) = \mathbb{P}(X_k^t = 1, X_k = -1) = (1 - e^{-t})p_k q_k \ .$$

This gives us the exchangeability of (X_k, X_k^t) for each $k \in \mathbb{N}$. Let $\mathbf{a} = (a_i, i \in \mathbb{N}), \mathbf{b} = (b_i, i \in \mathbb{N}) \in \{\pm 1\}^{\mathbb{N}}$, then using the independence within those two sequences $\mathbf{X}, \mathbf{X}^t$, we obtain

$$\begin{aligned}
\mathbb{P}(\mathbf{X} = \mathbf{a}, \mathbf{X}^t = \mathbf{b}) &= \prod_{k \in \mathbb{N}} \mathbb{P}(X_k = a_k, X_k^t = b_k) \\
&= \prod_{k \in \mathbb{N}} \mathbb{P}(X_k = b_k, X_k^t = a_k) \quad \text{by exchangeability of } X_k, X_k^t \\
&= \mathbb{P}(\mathbf{X} = \mathbf{b}, \mathbf{X}^t = \mathbf{a}) \ .
\end{aligned}$$

This proves the exchangeability of $\mathbf{X}, \mathbf{X}^t$. The rest follows from a standard approximation argument: it is clear that after truncation, (with $[N] := \{1, \dots, N\}$)

$$(Q_{p_1}(f_1 \mathbf{I}_{[N]^{p_1}}; \mathbf{Y}), \dots, Q_{p_d}(f_d \mathbf{I}_{[N]^{p_d}}; \mathbf{Y})) \text{ and } (Q_{p_1}(f_1 \mathbf{I}_{[N]^{p_1}}; \mathbf{Y}^t), \dots, Q_{p_d}(f_d \mathbf{I}_{[N]^{p_d}}; \mathbf{Y}^t))$$

form an exchangeable pair; letting $N \to +\infty$ and keeping in mind that the exchangeability is preserved in limit, we get the desired result. The verification of Mehler formula (2.2.1) can also be done by truncation argument, see for instance, [45, Proposition 3.1]

Construction of η^t: it is much more delicate. Recall from Remark 1.2.1 that our Poisson random measure can be represented as follows:

$$\eta = \sum_{n=1}^{\kappa} \delta_{X_n} \ , \tag{2.2.2}$$

where X_n, $n \geq 1$ are random variables with values in $\mathcal{Z}$ and κ is a $\mathbb{N}_0 \cup \{+\infty\}$-valued random variable. Now let Q be a standard exponential measure on $\mathbb{R}_+$ and let $(Y_n)_{n \in \mathbb{N}}$ be a sequence of *i.i.d.* random variables with distribution Q, independent of (κ, X_n). Then the *marked* point process ξ, given by

$$\xi := \sum_{n=1}^{\kappa} \delta_{(X_n, Y_n)} \ ,$$

is a Poisson point process with intensity measure $\mu \otimes Q$. For each $t \in \mathbb{R}_+$, we define

$$\eta_{e^{-t}}(A) := \xi(A \times [t, +\infty)) \, , \; A \in \mathscr{Z} \, ,$$

which is called the e^{-t}-thinning of η: it is obtained by removing the atoms (X_n) in η independently of each other with probability $1 - e^{-t}$. Moreover, $\eta_{e^{-t}}$ and $\eta - \eta_{e^{-t}}$ are two independent[9] Poisson point processes with intensity measure $e^{-t}\mu$, $(1 - e^{-t})\mu$ respectively. One can refer to [49, Chapter 5] for more details.

For any fixed $t \geq 0$, let $\eta'_{1-e^{-t}}$ be a Poisson point process on $\mathcal{Z}$ with control $(1 - e^{-t})\mu$ such that it is independent of $(\eta, \eta_{e^{-t}})$. And we put $\eta^t = \eta_{e^{-t}} + \eta'_{1-e^{-t}}$.

[9]The independence stems from the fact that they charge disjoint sets.

It is easy to verify that (η, η^t) is an exchangeable pair of Poisson point processes.

Indeed, it suffices to notice that $\eta = \eta_{e^{-t}} + \eta - \eta_{e^{-t}}$ and that $\eta - \eta_{e^{-t}}$, $\eta'_{1-e^{-t}}$ have the same law, and are both independent of $\eta_{e^{-t}}$. Now let $\mathfrak{f} : \mathcal{S}_\sigma \to \mathbb{R}$ be $\mathscr{S}_\sigma$-measurable, then for any Borel subsets A_1, A_2 of $\mathbb{R}$,

$$\begin{aligned}\mathbb{P}(\mathfrak{f}(\eta) \in A_1\,,\ \mathfrak{f}(\eta^t) \in A_2) &= \mathbb{P}(\mathfrak{f}(\eta_{e^{-t}} + \eta - \eta_{e^{-t}}) \in A_1\,,\ \mathfrak{f}(\eta_{e^{-t}} + \eta'_{1-e^{-t}}) \in A_2)\\ &= \mathbb{P}(\mathfrak{f}(\eta_{e^{-t}} + \eta'_{1-e^{-t}}) \in A_1\,,\ \mathfrak{f}(\eta_{e^{-t}} + \eta - \eta_{e^{-t}}) \in A_2)\\ &= \mathbb{P}(\mathfrak{f}(\eta^t) \in A_1\,,\ \mathfrak{f}(\eta) \in A_2)\ .\end{aligned}$$

This implies the exchangeability of (η, η^t).

Now let us verify the Mehler formula (2.2.1): in fact, we can go through the same density arguments as in the Gaussian setting, and we only need to establish the following claim: for $A \in \mathscr{Z}_\mu$,

$$\mathbb{E}[\widehat{\eta^t}(A)|\sigma\{\eta\}] = e^{-t}\,\widehat{\eta}(A)\,, \tag{2.2.3}$$

where $\widehat{\eta^t} := \eta^t - \mu$ denotes the compensate Poisson random measure. It simply follows from the construction of η^t that

$$\begin{aligned}\mathbb{E}[\widehat{\eta^t}(A)|\sigma\{\eta\}] &= \mathbb{E}[\widehat{\eta_{e^{-t}}}(A) + \widehat{\eta'_{1-e^{-t}}}(A)|\sigma\{\eta\}] = \mathbb{E}[\widehat{\eta_{e^{-t}}}(A)|\sigma\{\eta\}] \quad \text{by independence}\\ &= \mathbb{E}[\eta_{e^{-t}}(A)|\sigma\{\eta\}] - e^{-t}\mu(A) = \mathbb{E}\Big[\xi(A\times[t,+\infty))|\sigma\{\eta\}\Big] - e^{-t}\mu(A)\\ &= \mathbb{E}\left[\sum_{n=1}^{\kappa} \delta_{(X_n,Y_n)}(A\times[t,+\infty))|\sigma\{\eta\}\right] - e^{-t}\mu(A)\\ &= \mathbb{E}\left[\sum_{n=1}^{\kappa} \delta_{X_n}(A)\mathbf{I}_{\{Y_n\ge t\}}|\sigma\{\eta\}\right] - e^{-t}\mu(A)\\ &= e^{-t}\,\mathbb{E}\left[\sum_{n=1}^{\kappa} \delta_{X_n}(A)|\sigma\{\eta\}\right] - e^{-t}\mu(A) \quad \text{by independence}\\ &= e^{-t}\,\widehat{\eta}(A)\,,\end{aligned}$$

which gives us (2.2.3).

The above discussion also gives us a natural exchangeable pair coupling on Poisson Wiener chaos, that is, given $f \in L^2_s(\mu^p)$, $I^\eta_p(f)$ and $I^{\eta^t}_p(f)$ are exchangeable.

It is clear to us/readers now: we have constructed exchangeable pairs on three chaoses, and we shall apply our plug-in results (*i.e.* Propositions 2.1.2, 2.1.3).

Let us hold on for a while and recall that the random variable F in Proposition 2.1.2 is required to be in $L^4(\mathbb{P})$. If F is in a Gaussian Wiener chaos, then automatically, F has finite moment of any order, which, however, is not the case on the Poisson space and in the general Rademacher setting. And moreover, the set of finitely many Poisson Wiener chaoses (or Rademacher chaoses) is not stable under multiplication. As a consequence, $L(FG)$ may be ill-defined for general $F, G \in \mathbb{C}^\eta_p \cup \mathbb{C}^\eta_q$ $(p, q \in \mathbb{N})$. Same remark applies in the general Rademacher setting.

The following lemma points out the situations, in which one can have the product of two chaoses in finitely many chaoses.

Lemma 2.2.1. (i) Given $p, q \in \mathbb{N}$ and $f \in L^2_s(\mu^p)$, $g \in L^2_s(\mu^q)$, if $F = I^\eta_p(f)$, $G = I^\eta_q(g)$ belong to $L^4(\mathbb{P})$, then $FG \in L^2(\mathbb{P})$ has the finite chaos expansion:

$$FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q-1} J_k(FG) + I^\eta_{p+q}(f\widetilde{\otimes}g)\,.$$

(ii) Given $p, q \in \mathbb{N}$ and $f \in \mathfrak{H}_0^{\odot p}$, $g \in \mathfrak{H}_0^{\odot q}$, if $F = Q_p(f)$, $G = Q_q(g)$ belong to $L^4(\mathbb{P})$, then $FG \in L^2(\mathbb{P})$ has the finite chaos expansion:

$$FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q-1} J_k(FG) + Q_{p+q}(f\widetilde{\otimes}g\mathbf{I}_{\Delta_{p+q}})\,. \tag{2.2.4}$$

In particular, if $Q_1(h)$ belongs to $L^4(\mathbb{P})$ for some $h \in \mathfrak{H}$, then

$$Q_1(h)^2 = \|h\|^2_{\mathfrak{H}} + Q_1(w) + Q_2(h\widetilde{\otimes}h\mathbf{I}_{\Delta_2}) \quad \text{with } w(k) = \frac{h(k)^2(q_k - p_k)}{\sqrt{p_k q_k}}, k \in \mathbb{N}.$$

Proof. (i) See [32, Lemma 2.4] for a proof. (ii) See [30, Lemma 2.3] or [**P5**, Lemma 2.1] for a proof. Here we only provide an alternative proof for (ii): one can first truncate the kernels, that is, put $f_n = f\mathbf{I}_{[n]^p}$, $g_n = g\mathbf{I}_{[n]^q}$ and $F_n = Q_p(f_n)$, $G_n = Q_q(g_n)$, then it is easy to see

$$F_nG_n = \mathbb{E}[F_nG_n] + \sum_{k=1}^{p+q-1} J_k(F_nG_n) + Q_{p+q}(f_n\widetilde{\otimes}g_n\mathbf{I}_{\Delta_{p+q}})\,, \tag{2.2.5}$$

while it follows from martingale convergence theorem that $\mathbb{E}[F|\sigma\{X_1, \dots, X_n\}] = F_n$ converges in $L^4(\mathbb{P})$ to F and in the same way, G_n converges in $L^4(\mathbb{P})$ to G, thus F_nG_n converges in $L^2(\mathbb{P})$ to FG, as $n \to +\infty$. Note also that J_k is bounded linear operator on $L^2(\Omega, \sigma\{\mathbf{X}\}, \mathbb{P})$. Thus, the equality (2.2.4) follows from (2.2.5) by passing n to infinity. $\square$

Remark 2.2.1. Given $\mathfrak{X} \in \{W, \eta, \mathbf{Y}\}$ and $p, q \in \mathbb{N}$, if $F \in \mathbb{C}^{\mathfrak{X}}_p$, $G \in \mathbb{C}^{\mathfrak{X}}_q$ have finite fourth moments, then by product formula (1.2.9) (in the Gaussian setting) and Lemma 2.2.1, we conclude that $\Gamma(F, G)$ is well defined and belongs to $L^2(\mathbb{P})$. Moreover, it has the following nice chaos expansion:

$$\Gamma(F, G) = \frac{p+q}{2}\mathbb{E}[FG] + \sum_{k=1}^{p+q-1} \frac{p+q-k}{2} J_k(FG)\,, \tag{2.2.6}$$

and as a consequence of the orthogonality properties, one deduces that

$$\mathbf{Var}(\Gamma(F, G)) \le \frac{(p+q-1)^2}{4} \sum_{k=1}^{p+q-1} \mathbf{Var}(J_k(FG)) \le \max\{p^2, q^2\} \sum_{k=1}^{p+q-1} \mathbf{Var}(J_k(FG))\,. \tag{2.2.7}$$

We will close this chapter with the following crucial result, which links the exchangeable pairs to carré du champ via *regression*.

Proposition 2.2.1. Fix $p, q \in \mathbb{N}$ and let (F, F_t) and (G, G_t) be two exchangeable pairs of random variables in any of the following three cases:

- $F = I_p^W(f)$, $F_t = I_p^{W^t}(f)$ and $G = I_q^W(g)$, $G_t = I_q^{W^t}(g)$ for some $f \in L_s^2(\mu^p)$, $g \in L_s^2(\mu^q)$;
- $F = I_p^\eta(f)$, $F_t = I_p^{\eta^t}(f)$ and $G = I_q^\eta(g)$, $G_t = I_q^{\eta^t}(g)$ for some $f \in L_s^2(\mu^p)$, $g \in L_s^2(\mu^q)$;
- $F = Q_p(f; \mathbf{Y})$, $F_t = Q_p(f; \mathbf{Y}^t)$ and $G = Q_q(g; \mathbf{Y})$, $G_t = Q_q(g; \mathbf{Y}^t)$ for some $f \in \mathfrak{H}_0^{\odot p}$, $g \in \mathfrak{H}_0^{\odot q}$.

If $F, G \in L^4(\mathbb{P})$, then, with $\mathfrak{X} \in \{W, \eta, \mathbf{Y}\}$ corresponding to any of the above cases,

(a) $\lim_{t \downarrow 0} \frac{1}{t} \mathbb{E}[F_t - F | \sigma\{\mathfrak{X}\}] = LF = -pF$ in $L^4(\mathbb{P})$.

(b) $\lim_{t \downarrow 0} \frac{1}{t} \mathbb{E}[(F_t - F)(G_t - G) | \sigma\{\mathfrak{X}\}] = 2\Gamma(F, G)$, with the convergence in $L^2(\mathbb{P})$.

(c) $\lim_{t \downarrow 0} \frac{1}{t} \mathbb{E}[(F_t - F)^4] = -4p\, \mathbb{E}[F^4] + 12\, \mathbb{E}[F^2 \Gamma(F)] \geq 0$.

Proof. By the Mehler formula (2.2.1), we have

$$\frac{1}{t} \mathbb{E}[F_t - F | \sigma\{\mathfrak{X}\}] = \frac{P_t(F) - F}{t} = \frac{e^{-pt} - 1}{t} F,$$

converges in $L^4(\mathbb{P})$ to $-pF = LF$, as $t \downarrow 0$. By Lemma 2.2.1, FG has a finite chaos expansion of the form:

- when $\mathfrak{X} = W$, $\quad FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q} I_k^W(f_k)$ for some $f_k \in L_s^2(\mu^k)$, $k = 1, \dots, p+q$;
- when $\mathfrak{X} = \eta$, $\quad FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q} I_k^\eta(g_k)$ for some $g_k \in L_s^2(\mu^k)$, $k = 1, \dots, p+q$;
- when $\mathfrak{X} = \mathbf{Y}$, $\quad FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q} Q_k(h_k; \mathbf{Y})$ for some $h_k \in \mathfrak{H}_0^{\odot k}$, $k = 1, \dots, p+q$.

Therefore, accordingly, $F_t G_t$ can be expressed as follows:

- when $\mathfrak{X} = W$, $\quad F_t G_t = \mathbb{E}[FG] + \sum_{k=1}^{p+q} I_k^{W^t}(f_k)$;
- when $\mathfrak{X} = \eta$, $\quad F_t G_t = \mathbb{E}[FG] + \sum_{k=1}^{p+q} I_k^{\eta^t}(g_k)$;
- when $\mathfrak{X} = \mathbf{Y}$, $\quad F_t G_t = \mathbb{E}[FG] + \sum_{k=1}^{p+q} Q_k(h_k; \mathbf{Y}^t)$.

It follows immediately that

$$\frac{1}{t}\mathbb{E}[F_tG_t - FG|\sigma\{\mathfrak{X}\}] = \begin{cases} \sum_{k=1}^{p+q} \frac{1}{t}\mathbb{E}[I_k^{W^t}(f_k) - I_k^W(f_k)|\sigma\{W\}] & \text{when } \mathfrak{X} = W\,; \\ \sum_{k=1}^{p+q} \frac{1}{t}\mathbb{E}[I_k^{\eta^t}(g_k) - I_k^{\eta}(g_k)|\sigma\{\eta\}] & \text{when } \mathfrak{X} = \eta\,; \\ \sum_{k=1}^{p+q} \frac{1}{t}\mathbb{E}[Q_k(h_k;\mathbf{Y}^t) - Q_k(h_k;\mathbf{Y})|\sigma\{\mathbf{Y}\}] & \text{when } \mathfrak{X} = \mathbf{Y}\,. \end{cases}$$

Thus, in $L^2(\mathbb{P})$ and as $t \downarrow 0$,

$$\frac{1}{t}\mathbb{E}[F_tG_t - FG|\sigma\{\mathfrak{X}\}] \longrightarrow \sum_{k=1}^{p+q} -k\, J_k(FG) = L(FG)$$

Hence, we infer that in $L^2(\mathbb{P})$ and as $t \downarrow 0$,

$$\begin{aligned} \frac{1}{t}\mathbb{E}[(F_t - F)(G_t - G)|\sigma\{\mathfrak{X}\}] &= \frac{1}{t}\mathbb{E}[F_tG_t - FG|\sigma\{\mathfrak{X}\}] - F\frac{\mathbb{E}[G_t - G|\sigma\{\mathfrak{X}\}]}{t} - G\frac{\mathbb{E}[F_t - F|\sigma\{\mathfrak{X}\}]}{t} \\ &\to L(FG) - FLG - GLF = 2\,\Gamma(F,G)\,. \end{aligned}$$

Since the pair (F, F_t) is exchangeable, we can write

$$\begin{aligned} \mathbb{E}[(F_t - F)^4] &= \mathbb{E}[F_t^4 + F^4 - 4F_t^3F - 4F^3F_t + 6F_t^2F^2] \\ &= 2\mathbb{E}[F^4] - 8\mathbb{E}[F^3F_t] + 6\mathbb{E}[F^2F_t^2] \qquad \text{(by exchangeability of } (F,F_t)) \\ &= 4\mathbb{E}[F^3(F_t - F)] + 6\mathbb{E}[F^2(F_t - F)^2] \quad \text{(after rearrangement)} \\ &= 4\mathbb{E}[F^3\mathbb{E}[F_t - F|\sigma\{\mathfrak{X}\}]] + 6\mathbb{E}[F^2\mathbb{E}[(F_t - F)^2|\sigma\{\mathfrak{X}\}]]. \end{aligned}$$

so (c) follows immediately from (a),(b) and the fact that $F \in L^4(\mathbb{P})$. □

Remark 2.2.2. (1) In fact, the statement (a) implies (b) and (c). In the paper [**P3**] written with Ivan Nourdin, we made a *novel* observation that in the Gaussian setting, once we have an infinitesimal version of exchangeable pairs of multiple Wiener-Itô integrals verifying the asymptotic linear regression (a), the quadratic regression (b) and the fourth order regression (c) follow immediately. In the same paper, we presented another construction of exchangeable pairs via the *Gibbs sampling procedure*; see [**P3**, Section 4] for more details.

(2) As we can see from the above proof, this implication also holds true on the Poisson space and in the Rademacher setting. However, there is one noticeable difference, that is, in the Gaussian setting, due to the diffusion property, $-4p\,\mathbb{E}[F^4] + 12\,\mathbb{E}[F^2\Gamma(F,F)] = 0$; while this quantity is believed to be strictly positive in the discrete settings, otherwise, F could be a random variable with density, in view of Proposition 2.1.2, see also Remark 3.1.1.

Chapter 3

Fourth moment phenomena via exchangeable pairs

Abstract

This chapter presents detailed proofs of results stated in Section 1.4. More precisely, in Section 3.1 we present unified proofs for FMTs on the Gaussian and Poisson spaces. The second section is devoted to the extension of our strategy to the Rademacher setting, while the last section covers a collection of universality results about homogeneous sums.

3.1 FMTs on Gaussian and Poisson space

3.1.1 Main results

Let us first consider the univariate case and state one of the main results in this chapter.

Theorem 3.1.1. (i) Fix $p \in \mathbb{N}$ and $F \in \mathbb{C}_p^\eta$ such that $F \in L^4(\mathbb{P})$ and $\sigma := \sqrt{\mathbf{Var}(F)} > 0$. Then

$$d_{\mathrm{W}}(F, \mathcal{N}(0,\sigma^2)) \leq \left(\frac{2p-1}{p\sigma\sqrt{2\pi}} + \frac{2}{3\sigma}\sqrt{\frac{4p-3}{p}} \right) \sqrt{\mathbb{E}[F^4] - 3\sigma^4}\,.$$

(ii) Fix an integer $p \geq 2$ and $F \in \mathbb{C}_p^W$ such that $\sigma := \sqrt{\mathbf{Var}(F)} > 0$, then

$$d_{\mathrm{TV}}(F, \mathcal{N}(0,\sigma^2)) \leq \frac{2p-1}{\sigma^2 p} \sqrt{\mathbb{E}[F^4] - 3\sigma^4}$$

and

$$d_{\mathrm{W}}(F, \mathcal{N}(0,\sigma^2)) \leq \frac{2p-1}{p\sigma\sqrt{2\pi}} \sqrt{\mathbb{E}[F^4] - 3\sigma^4}\,.$$

The coefficient in the above total-variation bound can be improved to $\frac{2}{\sigma^2}\sqrt{\frac{p-1}{3p}}$, see Theorem 5.2.6 in [67].

Proof of Theorem 3.1.1. First we can write $F = I_p^{\mathfrak{X}}(f)$ for some $f \in L_s^2(\mu^p)$, where $\mathfrak{X} = W$ or $\mathfrak{X} = \eta$ depending on whether $F \in \mathbb{C}_p^W$ or $F \in \mathbb{C}_p^\eta$. Now we define $F_t = I_p^{\mathfrak{X}^t}(f)$ and we deduce from Proposition 2.2.1 that as $t \downarrow 0$,

- $\frac{1}{t}\mathbb{E}[F_t - F|\sigma\{\mathfrak{X}\}] \to -pF$ in $L^4(\mathbb{P})$,
- $\frac{1}{t}\mathbb{E}[(F_t - F)^2|\sigma\{\mathfrak{X}\}] \to 2\Gamma(F)$ in $L^2(\mathbb{P})$,
- $\frac{1}{t}\mathbb{E}[(F_t - F)^4] \to 4p\Big(\frac{3}{p}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4]\Big)$.

It follows from Proposition 2.1.2, with $S = \frac{2}{\sigma^2}\Gamma(F) - 2p$ and $\rho(F) = \frac{4p}{\sigma^4}\Big(\frac{3}{p}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4]\Big)$ that

$$d_{\mathrm{W}}(F, \mathcal{N}(0,\sigma^2)) \le \frac{2}{\sigma p\sqrt{2\pi}}\mathbb{E}[|\Gamma(F) - p\sigma^2|] + \frac{2\sqrt{2}}{3\sigma}\sqrt{\frac{3}{p}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4]}\,.$$

Since $\mathbb{E}[\Gamma(F)] = p\sigma^2$, by Cauchy-Schwarz we have $\mathbb{E}[|\Gamma(F) - p\sigma^2|] \le \sqrt{\mathbf{Var}(\Gamma(F))}$. Therefore, we get

$$d_{\mathrm{W}}(F, \mathcal{N}(0,\sigma^2)) \le \frac{\sqrt{2/\pi}}{\sigma}\sqrt{\mathbf{Var}(p^{-1}\Gamma(F))} + \frac{2\sqrt{2}}{3\sigma}\sqrt{\frac{3}{p}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4]}\,.$$

If in addition, $F \in \mathbb{C}_p^W$, we have $\frac{3}{p}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4] = 0$ and the above bound reduces to

$$d_{\mathrm{W}}(F, \mathcal{N}(0,\sigma^2)) \le \frac{\sqrt{2/\pi}}{\sigma}\sqrt{\mathbf{Var}(p^{-1}\Gamma(F))}\,,$$

and we can also deduce from Proposition 2.1.2 that

$$d_{\mathrm{TV}}(F, \mathcal{N}(0,\sigma^2)) \le \frac{2}{p\sigma^2}\mathbb{E}[|\Gamma(F) - p\sigma^2|] \le \frac{2}{\sigma^2}\sqrt{\mathbf{Var}(p^{-1}\Gamma(F))}\,.$$

It remains to estimate the two quantities: $\mathbf{Var}(p^{-1}\Gamma(F))$ and $\frac{3}{p}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4]$. Then our theorem follows from the following Lemma 3.1.1. □

The following result was contained in Lemmas 3.1 and 3.2 of [32] for the Poisson space. Its proof verbatim works on the Gaussian space.

Lemma 3.1.1. Given $F \in L^4(\mathbb{P}) \cap \mathbb{C}_p^{\mathfrak{X}}$ with $p \in \mathbb{N}$ and $\mathfrak{X} \in \{W, \eta\}$, we have

$$\mathbf{Var}(\Gamma(F)) \le \frac{(2p-1)^2}{4}\big(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2\big)\,, \tag{3.1.1}$$

and

$$0 \le \frac{3}{p}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4] \le \frac{4p-3}{2p}\big(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2\big)\,. \tag{3.1.2}$$

Remark 3.1.1. (i) The first inequality in (3.1.2) was proved in [32] under the assumption (**A**). Indeed, the authors of [32] were able to use Mecke's equation to prove that under the assumption (**A**),

$$\frac{3}{p}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4] = \frac{1}{2p}\int_Z \mathbb{E}[|D_z^+ F|^4]\,\mu(dz) \geq 0\,.$$

However, this inequality is trivial to us, in view of the following limit

$$\lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}[(F_t - F)^4] = 4p\Big(\frac{3}{p}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4]\Big)\,.$$

(ii) Due to the discrete structure of the Poisson space, the quantity

$$\frac{3}{p}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4]$$

is believed to be *strictly* positive for nonzero random variable $F \in \mathbb{C}_p^\eta$. Unfortunately, we do not have a rigorous proof. Here is some heuristic argument: if we have the equality in (3.1.2), by our proof of Theorem 3.1.1, we would obtain

$$d_{\mathrm{TV}}(F, \mathcal{N}(0, \mathbf{Var}(F))) \leq \frac{2}{\mathbf{Var}(F)}\sqrt{\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2}\,,$$

which is *unlikely* true. For instance, we can consider the normalized partial sums of i.i.d standard Poisson random variables: their fourth cumulants will vanish asymptotically, while their total-variation distance to the standard Gaussian is always one.

Remark 3.1.2. (i) To appreciate more the simplicity of our strategy, we sketch Döbler and Peccati's proof in [32] here: they started with the pathwise representation of Malliavin derivative D and carré du champ Γ, and along the lines[1] of discrete Malliavin-Stein approach, they were able to obtain (see Proposition 4.1 in [32])

$$d_{\mathrm{W}}(F, \mathcal{N}(0,1)) \leq \sqrt{\mathbf{Var}(q^{-1}\Gamma(F))} + \frac{1}{\sqrt{q}}\left(\int_Z \mathbb{E}[|D_z^+ F|^4]\,\mu(dz)\right)^{1/2}$$

where $F \in \mathbb{C}_q^\eta$ has unit variance and satisfies some integrability conditions; as we already mentioned in point (i) of Remark 3.1.1, under their assumption (**A**), the second term in the bound can be re-expressed as a multiple of $3p^{-1}\mathbb{E}[F^2\Gamma(F)] - \mathbb{E}[F^4]$. The rest of their proof consists of a straightforward application of Lemma 3.1.1. It is still surprising to ourselves that our use of exchangeable pairs miraculously helps us to avoid the assumption (**A**).

(ii) Note that Döbler and Peccati [32] also provided a Berry-Esséen bound with same order as the Wasserstein bound, under a local version of their assumption (**A**). We are not sure if our strategy can prove it under the weakest possible assumption of finite fourth moment.

The following lemma provides us with some generalization of Lemma 3.1.1 that will be useful for our multivariate results. The points (ii) and (iii) are motivated by Proposition 3.6 in [17].

[1] One may want to compare the bound with that in (1.3.8).

Lemma 3.1.2. Fix $\mathfrak{X} \in \{W, \eta\}$ and integers $p, q \geq 1$. Let $F \in \mathbb{C}_p^{\mathfrak{X}}$ and $G \in \mathbb{C}_q^{\mathfrak{X}}$ have *finite* fourth moments.

(i) The following inequality generalizes (3.1.1):

$$\mathbf{Var}(\Gamma(F,G)) \leq \frac{(p+q-1)^2}{4}\Big(\mathbb{E}[F^2G^2] - 2\,\mathbb{E}[FG]^2 - \mathbf{Var}(F)\mathbf{Var}(G)\Big)\,. \tag{3.1.3}$$

(ii) If $p < q$, then

$$0 \leq \mathbf{Cov}(F^2, G^2) = \mathbb{E}[F^2G^2] - \mathbf{Var}(F)\mathbf{Var}(G) \leq \sqrt{\mathbb{E}[F^4]}\sqrt{\mathbb{E}[G^4] - 3\mathbb{E}[G^2]^2}\,. \tag{3.1.4}$$

(iii) If $p = q$, then

$$0 \leq \mathbf{Cov}(F^2, G^2) - 2\,\mathbb{E}[FG]^2 \leq 2\sqrt{(\mathbb{E}[F^4] - 3\,\mathbb{E}[F^2]^2)(\mathbb{E}[G^4] - 3\mathbb{E}[G^2]^2)}\,. \tag{3.1.5}$$

The proof of this lemma is postponed to Section 3.1.2.

Now let us state the multivariate fourth-moment bound.

Theorem 3.1.2 (C. Döbler, A. Vidotto and G. Zheng, 2017)**.** Fix integers $d \geq 2$ and $1 \leq q_1 \leq \ldots \leq q_d$. Consider $F = (F_1, \ldots, F_d)^T$ with $F_j \in \mathbb{C}_{q_j}^{\eta} \cap L^4(\mathbb{P})$ for each $j \in [d]$ and assume Σ is the covariance matrix of F. Then, with $N \sim \mathcal{N}(0, \Sigma)$, we have

$$\begin{aligned} d_3(F,N) \leq &\left(\frac{(2q_d - 1)\sqrt{2d}}{4q_1} + \frac{2q_d\sqrt{d\mathrm{Tr}(\Sigma)}}{9q_1}\right)\sum_{i=1}^d \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2} \\ &+ \frac{(2q_d - 1)\sqrt{2d}}{4q_1}\left(\sum_{i=1}^{d-1}\mathbb{E}[F_i^4]^{1/4}\right)\sum_{j=2}^d (\mathbb{E}[F_j^4] - 3\mathbb{E}[F_j^2]^2)^{1/4}\,. \end{aligned} \tag{3.1.6}$$

If in addition, Σ is positive definite, then we have

$$\begin{aligned} d_2(F,N) \leq &\left(\frac{(2q_d - 1)\|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1\sqrt{\pi}} + \frac{q_d\sqrt{2\pi}\|\Sigma^{-1/2}\|_{\mathrm{op}}\sqrt{\mathrm{Tr}(\Sigma)}}{6q_1}\right)\sum_{i=1}^d \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2} \\ &+ \frac{(2q_d - 1)\|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1\sqrt{\pi}}\left(\sum_{i=1}^{d-1}\mathbb{E}[F_i^4]^{1/4}\right)\sum_{j=2}^d (\mathbb{E}[F_j^4] - 3\mathbb{E}[F_j^2]^2)^{1/4}\,. \end{aligned} \tag{3.1.7}$$

Proof. Assume that $F = (F_1, \ldots, F_d)^T = (I_{q_1}^{\eta}(f_1), \ldots, I_{q_d}^{\eta}(f_d))^T$ with $f_j \in L_s^2(\mu^{q_j})$ for each $j \in [d]$. For each $t \in \mathbb{R}_+$, we define $F_t = (F_{1,t}, \ldots, F_{d,t})^T = (I_{q_1}^{\eta^t}(f_1), \ldots, I_{q_d}^{\eta^t}(f_d))^T$. Then, it is clear that (F_t, F) is an exchangeable pair and by Proposition 2.2.1, we deduce

$$\mathbb{E}\left[\frac{1}{t}(F_{i,t} - F_i)(F_{j,t} - F_j) - 2\Gamma(F_i, F_j)\,\Big|\,\sigma\{\eta\}\right] \xrightarrow[\text{as } \varepsilon\downarrow 0]{\text{in } L^2(\mathbb{P})} 0\,.$$

Therefore, as $t \downarrow 0$ and in $L^1(\mathbb{P})$, we have

$$\left\| \frac{1}{t}\mathbb{E}[(F_t - F)(F_t - F)^T|\sigma\{\eta\}] - (2\Gamma(F_i, F_j))_{1\le i,j\le d} \right\|_{\mathrm{HS}}^2$$
$$= \sum_{i,j=1}^{d} \left(\mathbb{E}\left[\frac{1}{t}(F_{i,t} - F_i)(F_{j,t} - F_j) - 2\Gamma(F_i, F_j) \,\middle|\, \sigma\{\eta\} \right] \right)^2 \to 0 \ .$$

It is easy to see that for each $j \in [d]$,

$$\lim_{t\downarrow 0} \frac{1}{t}\, \mathbb{E}[F_{j,t} - F_j|\sigma\{\eta\}] = -q_j F_j \quad \text{in } L^4(\mathbb{P}),$$

from which we deduce that as $t \downarrow 0$ and in $L^2(\mathbb{P})$, we have

$$\left\| \ \frac{1}{t}\mathbb{E}[F_t - F|\sigma\{\eta\}] - \Lambda F \right\|_{\mathbb{R}^d}^2 = \sum_{j=1}^{d} \left(\mathbb{E}\left[\frac{F_{j,t} - F_j}{t} + q_j F_j \,\middle|\, \sigma\{\eta\} \right] \right)^2 \to 0 \ ,$$

with $\Lambda = \mathrm{diag}(q_1, \ldots, q_d)$ in such a way that $\|\Lambda^{-1}\|_{\mathrm{op}} = 1/q_1$.

It is also clear that, for each $i \in [d]$,

$$\begin{aligned} \rho_i(F) := \lim_{t\downarrow 0} \frac{1}{t}\, \mathbb{E}[(F_{i,t} - F_i)^4] &= -4q_i\, \mathbb{E}[F_i^4] + 12\mathbb{E}[F_i^2\Gamma(F_i, F_i)] \\ &\le 2(4q_i - 3)\Big(\mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2\Big) \quad \text{by (3.1.2).} \end{aligned}$$

Now define $S_{i,j} := 2\Gamma(F_i, F_j) - 2q_i\, \Sigma_{i,j}$ for $i, j \in [d]$, and observe in particular that $S_{i,j}$ is centered. Thus,

$$\begin{aligned} \sqrt{\sum_{i=1}^{d} 2\Lambda_{i,i}\Sigma_{i,i}} \sqrt{\sum_{i=1}^{d} \rho_i(F)} &\le \sqrt{\sum_{i=1}^{d} 2q_i\Sigma_{i,i}} \sqrt{\sum_{i=1}^{d} 2(4q_i - 3)\Big(\mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2\Big)} \\ &\le \sqrt{4q_d(4q_d - 3)\mathrm{Tr}(\Sigma)} \sqrt{\sum_{i=1}^{d} \Big(\mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2\Big)} \\ &\le 4q_d \sqrt{\mathrm{Tr}(\Sigma)} \sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2}\,, \end{aligned} \tag{3.1.8}$$

where the last inequality follows from the elementary fact that $\sqrt{a_1 + \ldots + a_d} \le \sqrt{a_1} + \ldots + \sqrt{a_d}$ for any nonnegative reals $a_1, \ldots, a_d$.

Now we consider $\mathbb{E}[\|S\|_{\mathrm{HS}}]$:

$$\mathbb{E}[\|S\|_{\mathrm{HS}}] \le \left(\sum_{i,j=1}^{d} \mathbb{E}[S_{i,j}^2] \right)^{1/2} = 2\left(\sum_{i,j=1}^{d} \mathbf{Var}(\Gamma(F_i, F_j)) \right)^{1/2} . \tag{3.1.9}$$

It follows from (3.1.3) that

$$\sum_{i,j=1}^{d}\mathbf{Var}(\Gamma(F_i,F_j)) \le \sum_{i,j=1}^{d}\frac{(q_i+q_j-1)^2}{4}\Big(\mathbb{E}[F_i^2F_j^2]-2\mathbb{E}[F_iF_j]^2-\mathbf{Var}(F_i)\mathbf{Var}(F_j)\Big)$$
$$\le \frac{(2q_d-1)^2}{4}\sum_{i,j=1}^{d}\Big(\mathbb{E}[F_i^2F_j^2]-2\mathbb{E}[F_iF_j]^2-\mathbf{Var}(F_i)\mathbf{Var}(F_j)\Big)$$
$$= \frac{(2q_d-1)^2}{4}\mathbb{E}[\|F\|^4-\|N\|^4]\,, \tag{3.1.10}$$

where the last equality is a consequence of the fact that (see *e.g.* (4.2) in [72])

$$\mathbb{E}[\|N\|^4]=\sum_{i,j=1}^{d}(\Sigma_{i,i}\Sigma_{j,j}+2\Sigma_{i,j}^2)\,.$$

It follows from Proposition 2.1.3 that

$$d_3(F,N)\le \frac{(2q_d-1)\sqrt{d}}{4q_1}\sqrt{\mathbb{E}[\|F\|^4-\|N\|^4]}+\frac{2q_d\sqrt{d}\mathrm{Tr}(\Sigma)}{9q_1}\sum_{i=1}^{d}\sqrt{\mathbb{E}[F_i^4]-3\mathbb{E}[F_i^2]^2} \tag{3.1.11}$$

and if in addition Σ is positive definite, we have

$$d_2(F,N)\le \frac{(2q_d-1)\|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1\sqrt{2\pi}}\sqrt{\mathbb{E}[\|F\|^4-\|N\|^4]}$$
$$+\frac{q_d\|\Sigma^{-1/2}\|_{\mathrm{op}}\sqrt{2\pi}\mathrm{Tr}(\Sigma)}{6q_1}\sum_{i=1}^{d}\sqrt{\mathbb{E}[F_i^4]-3\mathbb{E}[F_i^2]^2}\,. \tag{3.1.12}$$

Hence we can conclude our proof by evoking the following Lemma 3.1.3. □

Remark 3.1.3. It follows from Remark 2.1.2 that given $q_1\le\ldots\le q_d$, if $F=(F_1,\ldots,F_d)^T$ satisfies that $F_j\in\mathbb{C}^W_{q_j}$ for each $j\in[d]$ and that Σ is the covariance matrix of F, then with $N\sim\mathcal{N}(0,\Sigma)$, we have

$$d_2(F,N)\le\frac{\|\Lambda^{-1}\|_{\mathrm{op}}\sqrt{d}}{4}\mathbb{E}[\|S\|_{\mathrm{HS}}]\le\frac{(2q_d-1)\sqrt{d}}{4q_1}\sqrt{\mathbb{E}[\|F\|^4-\|N\|^4]}\,;$$

if in addition Σ is positive definite, we have

$$d_{\mathrm{W}}(F,N)\le\frac{\|\Lambda^{-1}\|_{\mathrm{op}}\|\Sigma^{-1/2}\|_{\mathrm{op}}}{\sqrt{2\pi}}\mathbb{E}[\|S\|_{\mathrm{HS}}]\le\frac{(2q_d-1)\|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1\sqrt{2\pi}}\sqrt{\mathbb{E}[\|F\|^4-\|N\|^4]}\,.$$

See also Theorem 4.3 in [72].

One can see from (3.1.10) that $\mathbb{E}[\|F\|^4-\|N\|^4]\ge 0$, and the following result provides the upper bound for $\mathbb{E}[\|F\|^4-\|N\|^4]$.

Lemma 3.1.3. Let F, N be given as in Theorem 3.1.2, then

$$\mathbb{E}[\|F\|^4 - \|N\|^4] \leq 2\left(\sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\right)^2 + 2\left(\sum_{i=1}^{d-1} \sqrt{\mathbb{E}[F_i^4]}\right)\sum_{j=2}^{d} \sqrt{\mathbb{E}[F_j^4] - 3\,\mathbb{E}[F_j^2]^2}\ .$$

In particular, if $q_1 = \ldots = q_d$, one has,

$$\mathbb{E}[\|F\|^4 - \|N\|^4] \leq 2\left(\sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\right)^2 .$$

Proof. Let us first consider the particular case where $q_1 = \ldots = q_d$. One obtains from Lemma 3.1.2 that

$$\begin{aligned}\mathbb{E}[\|F\|^4 - \|N\|^4] &= \sum_{i,j=1}^{d} \Big(\mathbb{E}[F_i^2F_j^2] - 2\mathbb{E}[F_iF_j]^2 - \mathbf{Var}(F_i)\mathbf{Var}(F_j)\Big) \\ &\leq 2\sum_{i,j=1}^{d} \sqrt{(\mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2)(\mathbb{E}[F_j^4] - 3\mathbb{E}[F_j^2]^2)} = 2\left(\sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\right)^2 .\end{aligned}$$

In the general case where $q_1 \leq \ldots \leq q_d$, Lemma 3.1.2 implies

$$\begin{aligned}&\mathbb{E}[\|F\|^4 - \|N\|^4] \\ &= \sum_{i,j=1}^{d} \mathbf{I}_{(q_i=q_j)}\Big(\mathbf{Cov}(F_i^2, F_j^2) - 2\mathbb{E}[F_iF_j]^2\Big) + 2\sum_{1\leq i<j\leq d} \mathbf{I}_{(q_i<q_j)}\mathbf{Cov}(F_i^2, F_j^2) \qquad (3.1.13) \\ &\leq 2\left(\sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\right)^2 + 2\sum_{1\leq i<j\leq d} \sqrt{\mathbb{E}[F_i^4]}\sqrt{(\mathbb{E}[F_j^4] - 3\mathbb{E}[F_j^2]^2)}\ .\end{aligned}$$

One can rewrite $\sum_{1\leq i<j\leq d}$ as $\sum_{j=2}^{d}\sum_{i=1}^{j-1}$ and then the desired result follows. □

Remark 3.1.4. (i) With the notation and assumptions given as in Theorem 3.1.2, if in addition $q_1 = q_d$, that is, all the component of the random vector F belong to the same Poisson Wiener chaos, then we can obtain better bounds, namely:

$$d_3(F, N) \leq \left(\frac{2q_d - 1)\sqrt{2d}}{4q_1} + \frac{2q_d\sqrt{d}\mathrm{Tr}(\Sigma)}{9q_1}\right)\sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\,.$$

If in addition Σ is positive definite, we have

$$d_2(F, N) \leq \left(\frac{(2q_d - 1)\|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1\sqrt{\pi}} + \frac{q_d\|\Sigma^{-1/2}\|_{\mathrm{op}}\sqrt{2\pi\mathrm{Tr}(\Sigma)}}{6q_1}\right)\sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\,.$$

(ii) Following the assumptions in Remark 3.1.3 and assuming in addition that $q_1 = q_d$, we get

$$d_2(F, N) \leq \frac{(2q_d - 1)\sqrt{2d}}{4q_1} \sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\,;$$

if in addition Σ is positive definite, we have

$$d_{\mathrm{W}}(F, N) \leq \frac{(2q_d - 1)\|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1\sqrt{\pi}} \sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\,;$$

for the general case where $q_1 \leq q_2 \leq \ldots \leq q_d$, one can also obtain fourth moment bounds in d_2 distance and Wasserstein distance. We refer interested readers to the paper [60] by Noreddine and Nourdin and in particular, one may want to compare this remark and Remark 3.1.3 with Theorem 1.5 therein.

We conclude this section with the Peccati-Tudor theorem on the Poisson space, which is an easy corollary of our Theorem 3.1.2.

Corollary 3.1.1 (C. Döbler, A. Vidotto and G. Zheng, 2017)**.** Fix $d \in \mathbb{N}$ and $q_1, \ldots, q_d \in \mathbb{N}$ and suppose that, for each $n \in \mathbb{N}$, $F^{(n)} := (F_1^{(n)}, \ldots, F_d^{(n)})^T$ is a random vector such that each $F_k^{(n)}$ belongs to the q_k-th Poisson Wiener chaos. Moreover, assume that Σ is a fixed nonnegative definite matrix and that $N = (N_1, \ldots, N_d)^T$ is a centered Gaussian vector with covariance matrix Σ. Assume that the following two conditions hold true:

(i) The covariance matrix of $F^{(n)}$ converges to Σ as $n \to \infty$.

(ii) For each $1 \leq k \leq d$ it holds that $\lim_{n\to\infty} \mathbb{E}[(F_k^{(n)})^4] = 3\Sigma_{k,k}^2$.

Then, as $n \to \infty$, the random vector $F^{(n)}$ converges in distribution to N.

Remark 3.1.5. The condition (ii) in Corollary 3.1.1 can be replaced by the following ***equivalent*** condition:

$$\lim_{n\to+\infty} \mathbb{E}[\|F^{(n)}\|^4] = \mathbb{E}[\|N\|^4]\,.$$

The equivalence is not trivial: assume (ii), then the above limit follows from Lemma 3.1.3; for the converse direction, we assume (i) and $\lim_{n\to+\infty} \mathbb{E}[\|F^{(n)}\|^4] = \mathbb{E}[\|N\|^4]$, then by (3.1.10), we have $\mathbf{Var}(\Gamma(F_i^{(n)})) \to 0$ for each $i \in [d]$. As we will see in Remark 3.1.6-(2),

$$\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 \leq \frac{6}{p}\mathbf{Var}(\Gamma(F))$$

for any $F \in \mathbb{C}_p^{\eta}$, then (ii) follows immediately.

3.1.2 Proof of Lemma 3.1.2 and Remarks

In this section, we first provide the proof of Lemma 3.1.2. The following result from [72, Lemma 2.2] will be helfpul.

Lemma 3.1.4 (I. Nourdin and J. Rosiński, 2014)**.** Given $p, q \in \mathbb{N}$, $f \in L^2_s(\mu^p)$ and $g \in L^2_s(\mu^q)$, we have

$$(p+q)!\, \|f \widetilde{\otimes} g\|^2 = p!q! \sum_{r=0}^{p\wedge q} \binom{p}{r}\binom{q}{r} \|f \otimes_r g\|^2 \geq p!q!\, \|f\|^2 \|g\|^2 + \mathbf{I}_{\{p=q\}} p!q!\, \langle f, g\rangle^2 \,,$$

and in the case of $p = q$, one has

$$(2p)! \langle f \widetilde{\otimes} f, g \widetilde{\otimes} g\rangle = 2p!^2 \langle f, g\rangle^2 + \sum_{r=1}^{p-1} p!^2 \binom{p}{r}^2 \langle f \otimes_r g, g \otimes_r f\rangle \;.$$

Here we follow the convention that $\sum_{r=1}^{0} = 0$.

Proof of Lemma 3.1.2. Without loss of generality, we assume $F = I^\eta_p(f)$ and $G = I^\eta_q(g)$ for some $f \in L^2_s(\mu^p)$ and $g \in L^2_s(\mu^q)$. The proof on the Gaussian space works verbatim as follows.

First it follows from Lemma 2.2.1-(i) and Remark 2.2.1 that $J_{p+q}(FG) = I^\eta_{p+q}(f \widetilde{\otimes} g)$ and

$$2\,\Gamma(F, G) = (p+q)\mathbb{E}[FG] + \sum_{k=1}^{p+q-1} (p+q-k)\, J_k(FG)\,. \tag{3.1.14}$$

It follows from Remark 2.2.1 that $\mathbf{Var}(\Gamma(F, G)) \leq \frac{(p+q-1)^2}{4} \sum_{k=1}^{p+q-1} \mathbf{Var}(J_k(FG))$. Similarly, as $FG \in L^2(\mathbb{P})$, we have $FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q} J_k(FG)$ so that

$$\begin{aligned} \mathbb{E}[F^2G^2] &= \mathbb{E}[FG]^2 + \sum_{k=1}^{p+q-1} \mathbf{Var}(J_k(FG)) + \mathbf{Var}(J_{p+q}(FG)) \\ &= \mathbb{E}[FG]^2 + \sum_{k=1}^{p+q-1} \mathbf{Var}(J_k(FG)) + (p+q)! \|f \widetilde{\otimes} g\|^2 \;. \end{aligned} \tag{3.1.15}$$

It follows from Lemma 3.1.4 that $(p+q)!\|f \widetilde{\otimes} g\|^2 \geq \mathbf{Var}(F)\mathbf{Var}(G) + \mathbb{E}[FG]^2$. Hence

$$\begin{aligned} \mathbf{Var}(\Gamma(F, G)) &\leq \frac{(p+q-1)^2}{4} \sum_{k=1}^{p+q-1} \mathbf{Var}(J_k(FG)) \\ &= \frac{(p+q-1)^2}{4} \Big(\mathbb{E}[F^2G^2] - \mathbb{E}[FG]^2 - (p+q)!\|f \widetilde{\otimes} g\|^2 \Big) \\ &\leq \frac{(p+q-1)^2}{4} \Big(\mathbb{E}[F^2G^2] - 2\,\mathbb{E}[FG]^2 - \mathbf{Var}(F)\mathbf{Var}(G) \Big) \;. \end{aligned} \tag{3.1.16}$$

In particular, Lemma 3.1.4, applied to $p = q$ and $f = g$, gives us

$$(2p)!\|f\widetilde{\otimes}f\|^2 = 2p!^2\|f\|^4 + p!^2\sum_{r=1}^{p-1}\binom{p}{r}^2\|f\otimes_r f\|^2\ ,$$

therefore implying

$$\begin{aligned}\mathbf{Var}(\Gamma(F)) &\leq \frac{(2p-1)^2}{4}\left(\mathbb{E}[F^4]-\mathbb{E}[F^2]^2-(2p)!\|f\widetilde{\otimes}f\|^2\right)\\ &= \frac{(2p-1)^2}{4}\left(\mathbb{E}[F^4]-3\,\mathbb{E}[F^2]^2-p!^2\sum_{r=1}^{p-1}\binom{p}{r}^2\|f\otimes_r f\|^2\right)\ .\end{aligned}$$

This proves (3.1.1) and

$$p!^2\sum_{r=1}^{p-1}\binom{p}{r}^2\|f\otimes_r f\|^2 \leq \mathbb{E}[F^4]-3\,\mathbb{E}[F^2]^2\ . \tag{3.1.17}$$

It is also clear from (3.1.16) that

$$\sum_{k=1}^{2p-1}\mathbf{Var}(J_k(F^2)) \leq \mathbb{E}[F^4]-3\,\mathbb{E}[F^2]^2\,. \tag{3.1.18}$$

By now, we have proved (i) of Lemma 3.1.2. Now let us prove (ii) and (iii).

To see $\mathbf{Cov}(F^2,G^2)\geq 0$, it is enough to rewrite $\mathbf{Cov}(F^2,G^2)$ using (3.1.15) and Lemma 3.1.4:

$$\mathbf{Cov}(F^2,G^2) = 2\mathbb{E}[FG]^2+\sum_{k=1}^{p+q-1}\mathbf{Var}(J_k(FG))+p!q!\sum_{r=1}^{p\wedge q}\binom{p}{r}\binom{q}{r}\|f\otimes_r g\|^2\,. \tag{3.1.19}$$

Now we turn to the upper bounds for $\mathbf{Cov}(F^2,G^2)$. First by Lemma 2.2.1-(i), we have $J_{2p}(F^2) = I^\eta_{2p}(f\widetilde{\otimes}f)$ and $J_{2q}(G^2) = I^\eta_{2q}(g\widetilde{\otimes}g)$. Moreover, one has

$$\begin{aligned}\mathbb{E}[F^2G^2] = \mathbb{E}\left[F^2\sum_{k=0}^{2q}J_k(G^2)\right] &= \mathbb{E}[F^2J_0(G^2)]+\mathbb{E}[F^2J_{2q}(G^2)]+\mathbb{E}\left[F^2\sum_{k=1}^{2q-1}J_k(G^2)\right]\\ &= \mathbf{Var}(F)\mathbf{Var}(G)+\mathbb{E}[F^2\,J_{2q}(G^2)]+\mathbb{E}\left[F^2\sum_{k=1}^{2q-1}J_k(G^2)\right]\,.\end{aligned}$$

If $p < q$, then $\mathbb{E}[F^2\,J_{2q}(G^2)] = 0$, so that

$$\mathbf{Cov}(F^2,G^2) = \mathbb{E}\left[F^2\sum_{k=1}^{2q-1}J_k(G^2)\right] \leq \sqrt{\mathbb{E}[F^4]}\sqrt{\sum_{k=1}^{2q-1}\mathbf{Var}(J_k(G^2))}\ ,$$

where the above inequality follows from Cauchy-Schwarz inequality and isometry property. The desired result (3.1.4) follows from (3.1.18).

Now we consider the case where $p = q$,

$$\begin{aligned}\mathbb{E}\left[F^2 \sum_{k=1}^{2p-1} J_k(G^2)\right] &= \sum_{k=1}^{2p-1} \mathbb{E}[J_k(F^2)J_k(G^2)] \\ &\le \sqrt{\sum_{k=1}^{2p-1} \mathbf{Var}(J_k(F^2))}\ \sqrt{\sum_{k=1}^{2p-1} \mathbf{Var}(J_k(G^2))} \quad \text{(by Cauchy-Schwarz)} \\ &\le \sqrt{(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2)(\mathbb{E}[G^4] - 3\mathbb{E}[G^2]^2)} \quad \text{due to (3.1.18).}\end{aligned}$$

By orthogonality property, we have

$$\mathbb{E}[J_{2p}(F^2)J_{2p}(G^2)] = (2p)!\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\rangle = 2p!^2\langle f,g\rangle^2 + \sum_{r=1}^{p-1} p!^2\binom{p}{r}^2 \langle f\otimes_r g, g\otimes_r f\rangle,$$

where the last equality follows from Lemma 3.1.4. As a consequence, one has

$$\mathbb{E}[F^2 J_{2p}(G^2)] - 2\mathbb{E}[FG]^2 = \sum_{r=1}^{p-1} p!^2\binom{p}{r}^2 \langle f\otimes_r g, g\otimes_r f\rangle \le \sum_{r=1}^{p-1} p!^2\binom{p}{r}^2 \left\|f\otimes_r g\right\|^2$$

by Cauchy-Schwarz. Note that, by definition of contractions and Fubini theorem, we have $\|f\otimes_r g\|^2 = \langle f\otimes_{p-r} f, g\otimes_{p-r} g\rangle$ for each $r = 1,\ldots,p-1$. Thus,

$$\begin{aligned}\sum_{r=1}^{p-1} p!^2\binom{p}{r}^2 \left\|f\otimes_r g\right\|^2 &= \sum_{r=1}^{p-1} p!^2\binom{p}{r}^2 \langle f\otimes_{p-r} f, g\otimes_{p-r} g\rangle = \sum_{r=1}^{p-1} p!^2\binom{p}{r}^2 \langle f\otimes_r f, g\otimes_r g\rangle \\ &\le \sum_{r=1}^{p-1} p!^2\binom{p}{r}^2 \|f\otimes_r f\| \times \|g\otimes_r g\|_2 \quad \text{(by Cauchy-Schwarz)} \\ &\le \sqrt{\sum_{r=1}^{p-1} p!^2\binom{p}{r}^2 \|f\otimes_r f\|^2}\sqrt{\sum_{r=1}^{p-1} p!^2\binom{p}{r}^2 \|g\otimes_r g\|^2} \quad \text{(by Cauchy-Schwarz)} \\ &\le \sqrt{\mathbb{E}[F^4] - 3\,\mathbb{E}[F^2]^2}\ \sqrt{\mathbb{E}[G^4] - 3\,\mathbb{E}[G^2]^2} \quad \text{due to (3.1.17).}\end{aligned}$$

Hence, we obtain

$$\begin{aligned}\mathbf{Cov}(F^2, G^2) - 2\,\mathbb{E}[FG]^2 &= \mathbb{E}[F^2 J_{2p}(G^2)] - 2\,\mathbb{E}[FG]^2 + \mathbb{E}\left[F^2 \sum_{k=1}^{2p-1} J_k(G^2)\right] \\ &\le 2\sqrt{\mathbb{E}[F^4] - 3\,\mathbb{E}[F^2]^2}\ \sqrt{\mathbb{E}[G^4] - 3\,\mathbb{E}[G^2]^2}\ .\end{aligned}$$

The proof of Lemma 3.1.2 is completed now. □

For the sake of completeness, let us provide a quick proof of Lemma 3.1.1.

Proof of Lemma 3.1.1 . Note the inequality (3.1.1) in Lemma 3.1.1 is recovered by taking $F = G$ in (3.1.3) and the second inequality in (3.1.2) can be proved in similar lines: we write $F^2 = \mathbb{E}[F^2] + \sum_{k=1}^{2p} J_k(F^2)$ and by (3.1.14)

$$\Gamma(F) = p\,\mathbb{E}[F^2] + \frac{1}{2}\sum_{k=1}^{2p-1}(2p-k)J_k(F^2)\ . \tag{3.1.20}$$

So by orthogonality and (3.1.18), we have

$$\begin{aligned}-\mathbb{E}[F^4] + \frac{3}{p}\mathbb{E}[F^2\,\Gamma(F)] &= -\mathbb{E}[F^4] + 3\mathbb{E}[F^2]^2 + \frac{3}{2p}\sum_{k=1}^{2p-1}(2p-k)\,\mathbf{Var}(J_k(F^2))\\ &\le -\mathbb{E}[F^4] + 3\mathbb{E}[F^2]^2 + \frac{3(2p-1)}{2p}(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2) = \frac{4p-3}{2p}\kappa_4(F)\ ,\end{aligned}$$

where $\kappa_4(F) := \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2$. This gives us the second inequality in (3.1.2). □

There are several interesting relations in the above proof. We will summarize them in the following remark.

Remark 3.1.6. (1) Let $F \in \mathbb{C}_p^\eta$ have nonzero variance, then it holds true that $\mathbb{E}[F^4] > 3\mathbb{E}[F^2]^2$. Indeed, we can always assume $F \in L^4(\mathbb{P})$. If $p = 1$, $F = I_1^\eta(f)$ for some $f \in L^2(\mu)$, then by product formula (see *e.g.* Proposition 6.1 in [47]), one has $\mathbb{E}[I_1^\eta(f)^4] = 3\,\|f\|_2^4 + \int_{\mathcal{Z}} f(z)^4\,d\mu > 3\mathbb{E}[F^2]^2$. For $p \ge 2$, $F = I_p^\eta(f)$ for some $f \in L_s^2(\mu^p)$, then according to (3.1.17), $\mathbb{E}[F^4] = 3\mathbb{E}[F^2]^2$ would imply $\|f \otimes_1 f\|_2 = 0$, which would further imply by standard arguments that $f = 0$ μ-almost everywhere, which is a contradiction to the fact that F is nonzero. If $F \in \mathbb{C}_p^W$ with $p \ge 2$ and $\mathbf{Var}(F) > 0$, the same argument will give us $\mathbb{E}[F^4] > 3\mathbb{E}[F^2]^2$. Note that random variables in the first Gaussian Wiener chaos are centered Gaussian distributed.

(2) Let $F \in \mathbb{C}_p^{\mathfrak{X}} \cap L^4(\mathbb{P})$, one has $p(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2) \le 6\mathbf{Var}(\Gamma(F,F))$, which shall be compared with (3.1.1). In fact, it follows first from (3.1.2) that $\mathbb{E}[F^4]-3\mathbb{E}[F^2]^2 \le 3\mathbb{E}[F^2(p^{-1}\Gamma(F)-\mathbb{E}[F^2])]$, and by (3.1.20) and orthogonality property, we have

$$\mathbb{E}[F^2(\Gamma(F) - p\mathbb{E}[F^2])] = \sum_{k=1}^{2p-1}\frac{2p-k}{2}\mathbf{Var}(J_k(F^2)) \le \sum_{k=1}^{2p-1}\frac{(2p-k)^2}{2}\mathbf{Var}(J_k(F^2)) = 2\mathbf{Var}(\Gamma(F))\,,$$

hence $p(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2) \le 6\mathbf{Var}(\Gamma(F))$. This inequality also explains the equivalence in Remark 3.1.5. Moreover, this echoes Nualart and Ortiz-Latorre's work[2] [75]: if $F_n \in \mathbb{C}_p^\eta$ has unit variance for each n, then the following implication holds:

$$\Gamma(F_n) \xrightarrow{n\to+\infty} p \text{ in } L^2(\mathbb{P}) \quad\Longrightarrow\quad F_n \xrightarrow[n\to+\infty]{law} \mathcal{N}(0,1)\,.$$

(3) Let F, N be given as in Theorem 3.1.2, then from (3.1.10) it follows that $\mathbb{E}[\|F\|^4] \ge \mathbb{E}[\|N\|^4]$. Moreover, if one of the components F_j in F has nonzero variance, it follows from the above two points and again (3.1.10) that $\mathbb{E}[\|F\|^4] > \mathbb{E}[\|N\|^4]$.

[2]Recall that on the Gaussian space, $\Gamma(F_n) = \|DF_n\|^2$ if $F_n \in \mathbb{C}_p^W$.

(4) Let the assumptions in Proposition 1.4.1 prevail. It follows from (3.1.17) that

$$p!^2 \sum_{r=1}^{p-1} \binom{p}{r}^2 \|f_n \otimes_r f_n\|^2 \leq \mathbb{E}[I_p^\eta(f_n)^4] - 3\,\mathbb{E}[I_p^\eta(f_n)^2]^2 \ .$$

If $\mathbb{E}[I_p^\eta(f_n)^4] \to 3$ as $n \to +\infty$, then $\|f_n \otimes_r f_n\| \to 0$ for each $r \in \{1, \dots, p-1\}$. Therefore by Nualart and Peccati's FMT, $\mathbb{E}[I_p^W(f_n)^4] \to 3$ and moreover by Nourdin-Peccati bound, $d_{\mathrm{TV}}(I_p^W(f_n), N) \to 0$, as $n \to +\infty$. This completes the proof of our transfer principle, *i.e.* Proposition 1.4.1.

(5) Fix $F = I_p^W(f)$ and $G = I_q^W(g)$ for some $f \in L_s^2(\mu^p)$ and $g \in L_s^2(\mu^q)$, with $p, q \in \mathbb{N}$. The equality (3.1.19) still holds true with the same proof. So it holds true that $\mathbf{Cov}(F^2, G^2) \geq 0$ with equality only when $\|f \otimes_1 g\| = 0$. So one can easily see that Rosiński and Samorodnitsky's observation (Proposition 1.2.1) follows from Üstünel and Zakai's result (Theorem 1.2.2).

(6) Fix $F = I_p^\eta(f) \in L^4(\mathbb{P})$ and $G = I_q^\eta(g) \in L^4(\mathbb{P})$ for some $f \in L_s^2(\mu^p)$ and $g \in L_s^2(\mu^q)$, with $p, q \in \mathbb{N}$. The equality (3.1.19) gives us the following equivalence:

$$\mathbf{Cov}(F^2, G^2) = 0 \quad \Leftrightarrow \quad \|f \otimes_1 g\| = 0 \text{ and } FG = I_{p+q}^\eta(f \widetilde{\otimes} g).$$

This equivalence seems new to us. However, we have no idea to go further and get the equivalent statement for $F \perp G$ in terms of contractions. It is believed that $F \perp G$ if and only if $f \star_1^0 g = 0$, while the proof has been *elusive* for a very long time. See [89] and references therein for more information in this direction.

3.2 Extension to the Rademacher setting

In this section, we adapt our strategy to the Rademacher setting. We will present the main results in Section 3.2.1 and defer the proofs of auxiliary lemmas in Section 3.2.2. The essential difference from the proofs in previous section is that in Rademacher setting, we need to be very careful about the off-diagonal part of the kernels, as shown in the following lemmas.

Lemma 3.2.1. Given $F = Q_p(f)$ with $f \in \mathfrak{H}_0^{\odot p}$ and $G = Q_q(g)$ with $g \in \mathfrak{H}_0^{\odot q}$, we assume that $F, G \in L^4(\mathbb{P})$. Then we have the following estimates:

$$\sum_{k=1}^{p+q-1} \mathbf{Var}(J_k(FG)) \leq \mathbb{E}[F^2G^2] - 2\mathbb{E}[FG]^2 - \mathbf{Var}(F)\mathbf{Var}(G) + (p+q)!\left\| f \widetilde{\otimes} g \mathbf{I}_{\triangle_{p+q}^c} \right\|^2 , \tag{3.2.1}$$

and in particular,

$$\max\left\{ \sum_{k=1}^{2p-1} \mathbf{Var}(J_k(F^2)), p!^2 \sum_{r=1}^{p-1} \binom{p}{r}^2 \|f \otimes_r f\|^2 \right\} \leq \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + (2p)!\|f \widetilde{\otimes} f \mathbf{I}_{\triangle_{2p}^c}\|^2 , \tag{3.2.2}$$

with

$$\left\| f \widetilde{\otimes} g \mathbf{I}_{\triangle_{p+q}^c} \right\|^2 \leq \sum_{r=1}^{p \wedge q} r! \binom{p}{r}\binom{q}{r} \min\left\{ \|f\|^2 \mathcal{M}(g), \|g\|^2 \mathcal{M}(f) \right\}. \tag{3.2.3}$$

As a convention, we put $\sum_{r=1}^{0} = 0$. The proof of this lemma will be presented in the end of this section.

Lemma 3.2.2. Given $F = Q_p(f) \in L^4(\mathbb{P})$ and $G = Q_q(g) \in L^4(\mathbb{P})$ for some $f \in \mathfrak{H}_0^{\odot p}$ and $g \in \mathfrak{H}_0^{\odot q}$. If $p = q$, then

$$\left|\mathbf{Cov}(F^2, G^2) - 2\mathbb{E}[FG]^2\right| \le 2\sqrt{(\kappa_4(F) + \gamma_q \mathcal{M}(f)\mathbb{E}[F^2])(\kappa_4(G) + \gamma_q \mathcal{M}(g)\mathbb{E}[G^2])} \tag{3.2.4}$$

$$+ \min\left\{ \|f\|^2 \sqrt{(2q)!\gamma_q \mathbb{E}[G^2]\mathcal{M}(g)}, \|g\|^2 \sqrt{(2q)!\gamma_q \mathbb{E}[F^2]\mathcal{M}(f)} \right\}. \tag{3.2.5}$$

If $p < q$, then

$$\left|\mathbf{Cov}(F^2, G^2) - 2\mathbb{E}[FG]^2\right| \le \sqrt{\mathbb{E}[F^4]}\sqrt{\kappa_4(G) + \gamma_q \mathbb{E}[G^2]\mathcal{M}(g)}, \tag{3.2.6}$$

where $\gamma_q := \dfrac{(2q)!}{q!}\displaystyle\sum_{r=1}^{q} r!\binom{q}{r}^2$.

3.2.1 Fourth moment-influence theorems

Let us first prove Döbler and Krokowski's univariate fourth moment-influence theorem as an illustration of our elementary strategy. The multivariate case is much more complicated and for the convenience of readers, we recall the univariate result in the following.

Theorem 3.2.1 (C. Döbler and K. Krokowski, 2017)**.** Fix $p \in \mathbb{N}$ and $f \in \mathfrak{H}_0^{\odot p}$ satisfying $p!\|f\|^2 = 1$. Let Z be a standard Gaussian and $F = Q_p(f; \mathbf{Y}) \in L^4(\mathbb{P})$, then,

$$d_{\mathrm{W}}(F, Z) \le C_1 \sqrt{\left|\mathbb{E}[F^4] - 3\right|} + C_2 \sqrt{\mathcal{M}(f)},$$

where $C_1 = \sqrt{2/\pi} + \dfrac{4}{3}$ and $C_2 = (\sqrt{2/\pi} + \dfrac{2\sqrt{6}}{3})\sqrt{\dfrac{(2p)!}{p!}\displaystyle\sum_{r=1}^{p} r!\binom{p}{r}^2}$.

Proof. Now given $F = Q_p(f; \mathbf{Y}) \in L^4(\mathbb{P})$ (with $\mathbb{E}[F^2] = 1$), we can get by using (2.2.7) and (3.2.2) that

$$\begin{aligned}\mathbf{Var}\Big(p^{-1}\Gamma(F)\Big) &\le \sum_{k=1}^{2p-1} \mathbf{Var}\Big(J_k(F^2)\Big) \\ &\le \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + (2p)!\left\|f \widetilde{\otimes} f \mathbf{I}_{\Delta_{2p}^c}\right\|^2 \\ &\le \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + \gamma_p \mathbb{E}[F^2]\mathcal{M}(f), \end{aligned} \tag{3.2.7}$$

with $\gamma_p := \dfrac{(2p)!}{p!}\displaystyle\sum_{r=1}^{p} r!\binom{p}{r}^2$.

Using the chaos expansion of F^2 and $\Gamma(F)$ as well as the orthogonality property, we have

$$
\begin{aligned}
0 \leq 3\mathbb{E}[F^2\Gamma(F)] - p\mathbb{E}[F^4] &= 3\mathbb{E}\Big[F^2(\Gamma(F) - p)\Big] - p\Big(\mathbb{E}[F^4] - 3\Big) \\
&= 3\mathbb{E}\left[\left(\sum_{k=0}^{2p} J_k(F^2)\right)\left(\sum_{k=1}^{2p-1} \frac{2p-k}{2} J_k(F^2)\right)\right] - p\Big(\mathbb{E}[F^4] - 3\Big) \\
&\leq 3p \sum_{k=1}^{2p-1} \mathbf{Var}\Big(J_k(F^2)\Big) - p\Big(\mathbb{E}[F^4] - 3\Big).
\end{aligned}
$$

It follows from (3.2.7) that

$$
0 \leq 3\mathbb{E}[F^2\Gamma(F)] - p\mathbb{E}[F^4] \leq 2p\Big(\mathbb{E}[F^4] - 3\Big) + 3p\gamma_p\mathcal{M}(f)\,. \tag{3.2.8}
$$

Now define $F_t = Q_p(f; \mathbf{Y}^t)$ for each $t \in [0, +\infty)$, then by Proposition 2.2.1, (F_t, F) is an exchangeable pair satisfying the conditions in Proposition 2.1.2 with $\mathscr{G} = \sigma\{\mathbf{X}\}$, $\lambda = p$, $S = 2\Gamma(F) - 2p$ and $\rho(F) = -4p\,\mathbb{E}[F^4] + 12\,\mathbb{E}[F^2\Gamma(F,F)]$. Therefore,

$$
\begin{aligned}
d_{\mathrm{W}}(F, N) &\leq \frac{1}{p\sqrt{2\pi}}\mathbb{E}[|2\Gamma(F) - 2p|] + \frac{\sqrt{2p}}{3p}\sqrt{-4p\,\mathbb{E}[F^4] + 12\,\mathbb{E}[F^2\Gamma(F)]} \\
&\leq \frac{2}{\sqrt{2\pi}}\sqrt{\mathbf{Var}\Big(p^{-1}\Gamma(F)\Big)} + \frac{\sqrt{2p}}{3p}\sqrt{-4p\,\mathbb{E}[F^4] + 12\,\mathbb{E}[F^2\Gamma(F)]}\,,
\end{aligned}
$$

then the desired bound follows from (3.2.7) and (3.2.8). □

Remark 3.2.1. Here is an interesting observation about our proof: given $f \in \mathfrak{H}_0^{\odot p}$ satisfying $p!\|f\|^2_{\mathfrak{H}^{\otimes p}} = 1$, if the fourth moment of $F = Q_p(f)$ is less than or equal to 3, then the above proof will imply that $d_{\mathrm{W}}(F, N) \leq C_2\sqrt{\mathcal{M}(f)}$.

The following is an easy corollary of the above fourth-moment-influence bound.

Corollary 3.2.1. Fix $p \in \mathbb{N}$. Let $\{f_n, n \in \mathbb{N}\} \subset \mathfrak{H}_0^{\odot p}$ satisfy $p!\|f_n\|^2 = 1$ for every $n \in \mathbb{N}$. Set $F_n = Q_p(f_n)$.

(i) Assume that $\mathbb{E}[F_n^4] \to 3$ and $\mathcal{M}(f_n) \to 0$, as $n \to +\infty$. Then F_n converges in law to a standard Gaussian, as $n \to +\infty$.

(ii) Assume[3] that $\mathbb{E}[F_n^4] \leq 3$ for each $n \in \mathbb{N}$ and $\mathcal{M}(f_k) \to 0$, as $k \to +\infty$. Then F_n converges in law to a standard Gaussian, as $n \to +\infty$.

Note that in the statement (ii), we do not require the convergence of fourth moments.

[3]This assumption is reasonable, see *e.g.* footnote 5 in Chapter 2.

Remark 3.2.2. (1) For F in the first Rademacher chaos, one can directly prove Theorem 3.2.1 without using the exchangeable pairs. Indeed, if $F = Q_1(h) \in L^4(\mathbb{P})$ for some $h \in \mathfrak{H}$ with $\|h\|_{\mathfrak{H}} = 1$ and $Z \sim \mathcal{N}(0,1)$, then by [**P1**, Theorem 3.1],

$$d_{\mathrm{W}}(F,Z) \leq \sqrt{\sum_{k=1}^{\infty} \frac{1}{p_k q_k} h(k)^4}\,.$$

By Lemma 2.2.1, $F^2 = 1 + Q_1(w) + Q_2(h \otimes h\mathbf{I}_{\Delta_2})$ with $w(k) = \frac{h(k)^2(q_k - p_k)}{\sqrt{p_k q_k}}$, $k \in \mathbb{N}$. This implies

$$\begin{aligned}
\mathbb{E}[F^4] &= 1 + \sum_{k=1}^{\infty} h(k)^4 \frac{(q_k - p_k)^2}{p_k q_k} + 2\|h \otimes h\|^2_{\mathfrak{H}^{\otimes 2}} - 2\|h \otimes h\mathbf{I}_{\Delta_2^c}\|^2_{\mathfrak{H}^{\otimes 2}} \\
&= 3 + \sum_{k=1}^{\infty} h(k)^4 \frac{(q_k - p_k)^2}{p_k q_k} - 2\sum_{k=1}^{\infty} h(k)^4 = 3 + \sum_{k=1}^{\infty} h(k)^4 \frac{q_k^2 + p_k^2}{p_k q_k} - 4\sum_{k=1}^{\infty} h(k)^4\,.
\end{aligned}$$

Noticing $p_k^2 + q_k^2 \geq 1/2$ for each $k \in \mathbb{N}$, we have

$$\frac{1}{2}\sum_{k=1}^{\infty} \frac{1}{p_k q_k} h(k)^4 \leq 4\sum_{k=1}^{\infty} h(k)^4 + \mathbb{E}[F^4] - 3 \leq 4\mathcal{M}(h) + \mathbb{E}[F^4] - 3\,.$$

Hence, $d_{\mathrm{W}}(F,Z) \leq \sqrt{2}\sqrt{\left|\mathbb{E}[F^4] - 3\right|} + 2\sqrt{2}\sqrt{\mathcal{M}(h)}$. Moreover, using the so-called second-order Poincaré inequality in [45, Theorem 4.1], we can have the Berry-Esséen bound

$$d_{\mathrm{Kol}}(F,Z) \leq 2\sqrt{\sum_{k=1}^{\infty} \frac{1}{p_k q_k} h(k)^4} \leq 3\sqrt{\left|\mathbb{E}[F^4] - 3\right|} + 6\sqrt{\mathcal{M}(h)}\,.$$

(2) Continuing the discussion in previous point and assuming $p_k = p = 1 - q = 1 - q_k$ for each k, we have

$$\mathbb{E}[F^4] - 3 = \frac{p^2 + q^4 - 4pq}{pq} \sum_{k=1}^{\infty} h(k)^4\,. \tag{3.2.9}$$

If $p \in (0,1) \setminus \{\frac{1}{2} \pm \frac{1}{2\sqrt{3}}\}$, then we have the exact fourth moment bounds:

$$d_{\mathrm{W}}(F,Z) \leq \sqrt{\frac{1}{pq}\sum_{k=1}^{\infty} h(k)^4} \leq \left(\frac{\mathbb{E}[F^4] - 3}{p^2 + q^2 - 4pq}\right)^{1/2} \text{ and } d_{\mathrm{Kol}}(F,Z) \leq 2\left(\frac{\mathbb{E}[F^4] - 3}{p^2 + q^2 - 4pq}\right)^{1/2},$$

see also Corollary 1.4 in [30].

Now we present the main result in this section, namely a Peccati-Tudor type result in the Rademacher setting.

Theorem 3.2.2. Fix integers $d \geq 2$ and $1 \leq q_1 \leq \ldots \leq q_d$. Assume that $F = (F_1, \ldots, F_d)$ verifies that $F_j = Q_{q_j}(f_j) \in L^4(\mathbb{P})$ with $f_j \in \mathfrak{H}_0^{\odot q_j}$ for each $j \in [d]$. Let Σ be the covariance matrix of F and $Z \sim \mathcal{N}(0, \Sigma)$. Then we have,

$$d_3(F, Z) \leq C_1 \sum_{i=1}^{d} \sqrt{|\kappa_4(F_i)|} + C_2 \sum_{j=1}^{d} \sqrt{\mathcal{M}(f_j)} + C_3 \sum_{i=2}^{d} |\kappa_4(F_i)|^{1/4} + C_4 \sum_{j=2}^{d} \mathcal{M}(f_j)^{1/4}\,;$$

in particular, when $q_1 = q_d$, we have $d_3(F, Z) \leq C_1 \sum_{i=1}^{d} \sqrt{|\kappa_4(F_i)|} + C_2 \sum_{j=1}^{d} \sqrt{\mathcal{M}(f_j)}\,.$

The constants C_1, C_2 and C_3 are given as follows:

$$C_1 := \frac{2q_d\sqrt{d}}{q_1}\Big(\frac{\sqrt{2}}{4} + \frac{\sqrt{\mathrm{Tr}(\Sigma)}}{9}\Big)\,, \quad C_2 := \frac{q_d}{q_1}\Big(\sqrt{d\gamma_{q_d}(2q_d)!\mathrm{Tr}(\Sigma)} + \frac{\sqrt{6d\gamma_{q_d}\mathrm{Tr}(\Sigma)}}{9}\Big)$$

$$C_3 := \frac{q_d\sqrt{2d}}{2q_1}\sum_{i=1}^{d-1}\mathbb{E}[F_i^4]^{1/4} \quad \text{and} \quad C_4 := \frac{q_d\sqrt{2d}}{2q_1}\big(\gamma_{q_d}\mathrm{Tr}(\Sigma)\big)^{1/4}\sum_{i=1}^{d-1}\mathbb{E}[F_i^4]^{1/4}\,,$$

with γ_p defined as in (3.2.7) for each $p \in \mathbb{N}$.

The following result follows immediately from the above bound.

Corollary 3.2.2. Fix integers $d \geq 2$ and $1 \leq q_1 \leq \ldots \leq q_d$, and consider the sequence of random vectors

$$F^{(n)} = (F_1^{(n)}, \ldots, F_d^{(n)})^T := (Q_{q_1}(f_{1,n}), \ldots, Q_{q_d}(f_{d,n}))^T$$

with kernels $f_{j,n}$ in $\mathfrak{H}_0^{\odot q_j}$ for each $n \in \mathbb{N}, j \in [d]$. Assume that the covariance matrix Σ_n of $F^{(n)}$ converges in Hilbert-Schmidt norm to a nonnegative definite symmetric matrix Σ, as $n \to +\infty$. Suppose that the following condition holds:

$$\lim_{n\to+\infty} [\mathcal{M}(f_{1,n}) + \ldots + \mathcal{M}(f_{d,n})] = 0\,.$$

If for each $j \in [d]$, $\mathbb{E}[(F_j^{(n)})^4]$ converges to $3\Sigma_{j,j}^2$, as $n \to +\infty$, then $F^{(n)}$ converges in law to $Z \sim \mathcal{N}(0, \Sigma)$, as $n \to +\infty$.

Proof of Theorem 3.2.2. For each $t \in \mathbb{R}_+$, we define $F_t = (F_{1,t}, \ldots, F_{d,t})$ with $F_{j,t} = Q_{q_j}(f_j; \mathbf{Y}^t)$ for each $j \in [d]$. Then (F_t, F) is an exchangeable pair and it follows from Proposition 2.1.3 and Proposition 2.2.1 that[4]

$$d_3(F, Z) \leq \frac{\sqrt{d}}{4q_1}\mathbb{E}[\|S\|_{\mathrm{HS}}] + \frac{\sqrt{2dq_d\mathrm{Tr}(\Sigma)}}{18q_1}\sqrt{\sum_{i=1}^{d}\rho_i(F)}\,, \tag{3.2.10}$$

where the random matrix S is defined by $S_{i,j} = 2\Gamma(F_i, F_j) - 2q_j\Sigma_{i,j}$ and $\rho_i(F) = -4q_i\mathbb{E}[F_i^4] + 12\mathbb{E}[F_i^2\Gamma(F_i)]$ for each $i, j \in [d]$.

[4]The argument is identical to that in previous section, so we omit it here.

Let us first deal with the easy part in the bound (3.2.10): by (3.2.8), we have

$$\rho_i(F) \le 4\big(2q_i\kappa_4(F_i) + 3q_i\gamma_{q_i}\mathcal{M}(f_i)\big)\,,$$

where $\kappa_4(F_i) = \mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2$ is the fourth cumulant of F_i and the number γ_{q_i} is defined as in (3.2.7) for each $i \in [d]$. Therefore,

$$\sqrt{\sum_{i=1}^d \rho_i(F)} \le 2\sqrt{\sum_{i\in[d]} 2q_i\kappa_4(F_i) + 3q_i\gamma_{q_i}\mathcal{M}(f_i)} \tag{3.2.11}$$

$$\le 2\sqrt{2q_d}\left(\sum_{i\in[d]}\sqrt{|\kappa_4(F_i)|}\right) + 2\sqrt{3q_d\gamma_{q_d}}\sum_{i\in[d]}\sqrt{\mathcal{M}(f_i)}. \tag{3.2.12}$$

Now let us consider the more complicated part, that is, to estimate $\mathbb{E}[\|S\|_{\mathrm{HS}}]$.

It follows from Lemma 3.2.1 and (2.2.7) that

$$\begin{aligned}
\mathbb{E}[\|S\|_{\mathrm{HS}}] &= 2\left(\sum_{i,j=1}^d \mathbf{Var}(\Gamma(F_i,F_j))\right)^{1/2}\\
&\le 2q_d\left(\sum_{i,j=1}^d\sum_{k=1}^{q_i+q_j-1}\mathbf{Var}(J_k(F_iF_j))\right)^{1/2}\\
&\le 2q_d\left(\sum_{i,j=1}^d\Big\{\mathbb{E}[F_i^2F_j^2] - 2\mathbb{E}[F_iF_j]^2 - \mathbf{Var}(F_i)\mathbf{Var}(F_j) + (q_i+q_j)!\|f_i\widetilde{\otimes} f_j\mathbf{I}_{\triangle^c_{q_i+q_j}}\|^2\Big\}\right)^{1/2}\\
&= 2q_d\left(\sum_{i,j=1}^d\Big|\mathbf{Cov}(F_i^2,F_j^2) - 2\mathbb{E}[F_iF_j]^2\Big|\right)^{1/2} + 2q_d\left(\sum_{i,j=1}^d(q_i+q_j)!\|f_i\widetilde{\otimes} f_j\mathbf{I}_{\triangle^c_{q_i+q_j}}\|^2\right)^{1/2}.
\end{aligned}$$

By using (3.2.3), we have

$$\|f_i\widetilde{\otimes} f_j\mathbf{I}_{\triangle^c_{q_i+q_j}}\|^2 \le \gamma_{q_d}\min\{\|f_i\|^2\mathcal{M}(f_j), \|f_j\|^2\mathcal{M}(f_i)\}$$

and thus,

$$\begin{aligned}
2q_d\left(\sum_{i,j=1}^d(q_i+q_j)!\|f_i\widetilde{\otimes} f_j\mathbf{I}_{\triangle^c_{q_i+q_j}}\|^2\right)^{1/2} &\le 2q_d\sqrt{\gamma_{q_d}(2q_d)!}\left(\sum_{i,j=1}^d\min\{\|f_i\|^2\mathcal{M}(f_j), \|f_j\|^2\mathcal{M}(f_i)\}\right)^{1/2}\\
&\le 2q_d\sqrt{\gamma_{q_d}(2q_d)!}\left(\sum_{i=1}^d\|f_i\|^2\sum_{j=1}^d\mathcal{M}(f_j)\right)^{1/2}\\
&\le 2q_d\sqrt{\gamma_{q_d}(2q_d)!\mathrm{Tr}(\Sigma)}\sum_{j=1}^d\sqrt{\mathcal{M}(f_j)}\,.
\end{aligned}$$

It remains to apply Lemma 3.2.2:

$$\sum_{i,j=1}^{d}\left|\mathbf{Cov}(F_i^2,F_j^2)-2\mathbb{E}[F_iF_j]^2\right|$$

$$\leq \sum_{i,j\in[d]}\mathbf{I}_{\{q_i=q_j\}}2\sqrt{(\kappa_4(F_i)+\gamma_{q_i}\mathcal{M}(f_i)\mathbb{E}[F_i^2])(\kappa_4(F_j)+\gamma_{q_j}\mathcal{M}(f_j)\mathbb{E}[F_j^2])} \tag{3.2.13}$$

$$+2\sum_{1\leq i<j\leq d}\mathbf{I}_{\{q_i<q_j\}}\sqrt{\mathbb{E}[F_i^4]}\sqrt{\kappa_4(F_j)+\gamma_{q_j}\mathcal{M}(f_j)\mathbb{E}[F_j^2]} \tag{3.2.14}$$

$$\leq 2\left(\sum_{j\in[d]}\sqrt{\kappa_4(F_j)+\gamma_{q_j}\mathcal{M}(f_j)\mathbb{E}[F_j^2]}\right)^2$$

$$+2\left(\sum_{i=1}^{d-1}\sqrt{\mathbb{E}[F_i^4]}\right)\sum_{j=2}^{d}\sqrt{\kappa_4(F_j)+\gamma_{q_j}\mathcal{M}(f_j)\mathbb{E}[F_j^2]}\,.$$

Therefore, we have

$$\begin{aligned}\mathbb{E}[\|S\|_{\mathrm{HS}}] &\leq 2\sqrt{2}q_d\sum_{j\in[d]}\sqrt{\kappa_4(F_j)+\gamma_{q_j}\mathcal{M}(f_j)\Sigma_{j,j}}\\ &\quad+2\sqrt{2}q_d\left(\sum_{i=1}^{d-1}\|F_i\|_{L^4(\mathbb{P})}\right)\sum_{j=2}^{d}\left[\kappa_4(F_j)+\gamma_{q_j}\mathcal{M}(f_j)\Sigma_{j,j}\right]^{1/4}\\ &\quad+2q_d\sqrt{\gamma_{q_d}(2q_d)!\mathrm{Tr}(\Sigma)}\sum_{j=1}^{d}\sqrt{\mathcal{M}(f_j)}\\ &\leq 2\sqrt{2}q_d\sum_{j\in[d]}\sqrt{|\kappa_4(F_j)|}+4q_d\sqrt{\gamma_{q_d}(2q_d)!\mathrm{Tr}(\Sigma)}\sum_{j\in[d]}\sqrt{\mathcal{M}(f_j)}\\ &\quad+2\sqrt{2}q_d\left(\sum_{i=1}^{d-1}\|F_i\|_{L^4(\mathbb{P})}\right)\left\{\sum_{j=2}^{d}|\kappa_4(F_j)|^{1/4}+(\gamma_{q_d}\mathrm{Tr}(\Sigma))^{1/4}\sum_{j=2}^{d}\mathcal{M}(f_j)^{1/4}\right\}.\end{aligned}$$

In particular, when $q_1=q_d$, the term in (3.2.14) vanishes, so that we can obtain a slightly neater estimate:

$$\mathbb{E}[\|S\|_{\mathrm{HS}}]\leq 2\sqrt{2}q_d\sum_{j\in[d]}\sqrt{|\kappa_4(F_j)|}+4q_d\sqrt{\gamma_{q_d}(2q_d)!\mathrm{Tr}(\Sigma)}\sum_{j\in[d]}\sqrt{\mathcal{M}(f_j)}.$$

Hence our proof is completed by putting all the estimates together into (3.2.10). □

Remark 3.2.3. (i) Let the assumptions of Theorem 3.2.2 prevail and we assume in addition that Σ is positive definite, then it follows from Proposition 2.1.3 that

$$d_2(F,Z)\leq\frac{\|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1\sqrt{2\pi}}\mathbb{E}[\|S\|_{\mathrm{HS}}]+\frac{\sqrt{\pi}\|\Sigma^{-1/2}\|_{\mathrm{op}}}{12q_1}\sqrt{q_d\mathrm{Tr}(\Sigma)}\sqrt{\sum_{i=1}^{d}\rho_i(F)}\,.$$

Therefore, using the estimates from the above proof, we can obtain

$$d_2(F,Z) \le C_1' \sum_{i=1}^{d} \sqrt{|\kappa_4(F_i)|} + C_2' \sum_{j=1}^{d} \sqrt{\mathcal{M}(f_j)} + C_3' \sum_{i=2}^{d} |\kappa_4(F_i)|^{1/4} + C_4' \sum_{j=2}^{d} \mathcal{M}(f_j)^{1/4} \,;$$

in particular, when $q_1 = q_d$, we have

$$d_2(F,Z) \le C_1' \sum_{i=1}^{d} \sqrt{|\kappa_4(F_i)|} + C_2' \sum_{j=1}^{d} \sqrt{\mathcal{M}(f_j)}$$

where

$$C_1' := \frac{q_d \|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1}\Big(2 + \sqrt{\mathrm{Tr}(\Sigma)}\Big)\,, \qquad C_2' := \frac{3q_d \|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1} \sqrt{\gamma_{q_d}(2q_d)!\mathrm{Tr}(\Sigma)}$$

$$C_3' := \frac{q_d \|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1} \sum_{i=1}^{d-1} \mathbb{E}[F_i^4]^{1/4} \quad \text{and} \quad C_4' := \frac{q_d \|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1} \big(\gamma_{q_d} \mathrm{Tr}(\Sigma)\big)^{1/4} \sum_{i=1}^{d-1} \mathbb{E}[F_i^4]^{1/4}\,.$$

(ii) One can see from (3.2.11), (3.2.14) and (3.2.13) that if the fourth cumulant of each component F_j is non-positive, then maximal influences alone could control the distance to normality. So one can formulate a multivariate version of point (ii) in Corollary 3.2.1.

3.2.2 Proofs of Lemma 3.2.1 and Lemma 3.2.2

Proof of Lemma 3.2.1. It follows from Lemma 2.2.1-(ii) that

$$FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q-1} J_k(FG) + Q_{p+q}\Big(f \widetilde{\otimes} g \mathbf{I}_{\triangle_{p+q}}\Big)\,,$$

therefore, by orthogonality property, one has

$$\begin{aligned}
\mathbb{E}[F^2G^2] &= \mathbb{E}[FG]^2 + \sum_{k=1}^{p+q-1} \mathbf{Var}(J_k(FG)) + (p+q)!\, \Big\| f \widetilde{\otimes} g \mathbf{I}_{\triangle_{p+q}} \Big\|^2 \\
&= \mathbb{E}[FG]^2 + \sum_{k=1}^{p+q-1} \mathbf{Var}(J_k(FG)) + (p+q)!\, \Big\| f \widetilde{\otimes} g \Big\|^2 - (p+q)! \Big\| f \widetilde{\otimes} g \mathbf{I}_{\triangle^c_{p+q}} \Big\|^2\,,
\end{aligned}$$

thus, the inequality (3.2.1) follows from (3.1.4). It also follows from (3.1.4) that

$$\sum_{k=1}^{p+q-1} \mathbf{Var}(J_k(F^2)) = \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 - p!^2 \sum_{r=1}^{p-1} \binom{p}{r}^2 \Big\| f \otimes_r f \Big\|^2 + (2p)! \Big\| f \widetilde{\otimes} f \mathbf{I}_{\triangle^c_{2p}} \Big\|^2 \,, \tag{3.2.15}$$

which implies (3.2.2).

It remains to prove (3.2.3) and we'll use the same arguments as in the proof of [30, Lemma 3.3]:

$$\|f\widetilde{\otimes}g\mathbf{I}_{\triangle^c_{p+q}}\|^2 \leq \|f\otimes g\mathbf{I}_{\triangle^c_{p+q}}\|^2 = \sum_{(i_1,\dots,i_p,j_1,\dots,j_q)\in\triangle^c_{p+q}} f(i_1,\dots,i_p)^2 g(j_1,\dots,j_q)^2$$

$$(3.2.16)\qquad = \sum_{r=1}^{p\wedge q} r!\binom{p}{r}\binom{q}{r} \sum_{\substack{(i_1,\dots,i_p)\in\triangle_p\\(j_1,\dots,j_q)\in\triangle_q\\ \mathrm{card}(\{i_1,\dots,i_p\}\cap\{j_1,\dots,j_q\})=r}} f(i_1,\dots,i_p)^2 g(j_1,\dots,j_q)^2\,,$$

where card(A) means the cardinality of the set A, and the combinatorial constant $r!\binom{p}{r}\binom{q}{r}$ is the number of ways one can build r pairs of identical indices out of $(i_1,\dots,i_p)\in\triangle_p$ and $(j_1,\dots,j_q)\in\triangle_q$. Therefore, it is enough to notice that for each $r\in\{1,\dots,p\wedge q\}$, the inner sum in (3.2.16) is bounded by

$$\sum_{\substack{(i_1,\dots,i_{p-r},k_1,\dots,k_r)\in\triangle_p\\(j_1,\dots,j_{q-r},k_1,\dots,k_r)\in\triangle_q}} f(i_1,\dots,i_{p-r},k_1,\dots,k_r)^2 g(j_1,\dots,j_{q-r},k_1,\dots,k_r)^2$$

$$\leq \sum_{\substack{(i_1,\dots,i_{p-1},k)\in\triangle_p\\(j_1,\dots,j_{q-1},k)\in\triangle_q}} f(i_1,\dots,i_{p-1},k)^2 g(j_1,\dots,j_{q-1},k)^2 \leq \min\left\{\|f\|^2\mathcal{M}(g),\|g\|^2\mathcal{M}(f)\right\}.$$

The proof of Lemma 3.2.1 is complete. □

Proof of Lemma 3.2.2. Assume $p\leq q$ and we begin with the following chaos expansion:

$$\mathbb{E}[F^2G^2] = \mathbb{E}\left(F^2\left\{\mathbb{E}[G^2] + \sum_{k=1}^{2q-1} J_k(G^2) + J_{2q}(G^2)\right\}\right)$$

$$= \mathbf{Var}(F)\mathbf{Var}(G) + \mathbb{E}\left(F^2\sum_{k=1}^{2q-1} J_k(G^2)\right) + \mathbf{I}_{(p=q)}\mathbb{E}[J_{2q}(F^2)J_{2q}(G^2)]\,.$$

If $p<q$, then $\mathbb{E}[FG]=0$ and

$$\left|\mathbf{Cov}(F^2,G^2)\right| \leq \sqrt{\mathbb{E}[F^4]}\sqrt{\sum_{k=1}^{2q-1}\mathbf{Var}(J_k(G^2))} \leq \sqrt{\mathbb{E}[F^4]}\sqrt{\kappa_4(G)+\gamma_q\mathbb{E}[G^2]\mathcal{M}(g)}\,,$$

where the second inequality follows from (3.2.7) and the constant γ_q is given therein.
If $p=q$, then

$$\mathbb{E}[J_{2q}(F^2)J_{2q}(G^2)] = (2q)!\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\mathbf{I}_{\triangle_{2q}}\rangle = (2q)!\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\rangle - (2q)!\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\mathbf{I}_{\triangle^c_{2q}}\rangle$$

$$= 2q!^2\langle f,g\rangle^2 + \sum_{r=1}^{q-1} q!^2\binom{q}{r}^2\langle f\otimes_r g, g\otimes_r f\rangle - (2q)!\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\mathbf{I}_{\triangle^c_{2q}}\rangle\,,$$

where the last equality follows from Lemma 3.1.4.

Consequently, $\mathbf{Cov}(F^2, G^2) - 2\mathbb{E}[FG]^2$ is equal to

$$\mathbb{E}\left(F^2 \sum_{k=1}^{2q-1} J_k(G^2)\right) + \sum_{r=1}^{q-1} q!^2 \binom{q}{r}^2 \langle f \otimes_r g, g \otimes_r f\rangle - (2q)! \left\langle f\widetilde{\otimes} f, g\widetilde{\otimes} g \mathbf{I}_{\Delta_{2q}^c}\right\rangle. \tag{3.2.17}$$

The first term in (3.2.17) can be rewritten as $\mathbb{E}\left[\sum_{k=1}^{2q-1} J_k(F^2) J_k(G^2)\right]$, which can be bounded by

$$\sqrt{\sum_{k=1}^{2q-1} \mathbf{Var}(J_k(F^2))} \sqrt{\sum_{k=1}^{2q-1} \mathbf{Var}(J_k(F^2))} \leq \sqrt{\kappa_4(F) + \gamma_q \mathbb{E}[F^2] \mathcal{M}(f)} \sqrt{\kappa_4(G) + +\gamma_q \mathbb{E}[G^2] \mathcal{M}(g)}\ ;$$

and the second term in (3.2.17) can be bounded by

$$\sum_{r=1}^{q-1} q!^2 \binom{q}{r}^2 \|f \otimes_r g\|^2 \leq \sqrt{\sum_{r=1}^{q-1} q!^2 \binom{q}{r}^2 \|f \otimes_r f\|^2} \sqrt{\sum_{r=1}^{q-1} q!^2 \binom{q}{r}^2 \|g \otimes_r g\|^2} \tag{3.2.18}$$

$$\leq \sqrt{\kappa_4(F) + \gamma_q \mathcal{M}(f) \mathbb{E}[F^2]} \sqrt{\kappa_4(G) + \gamma_q \mathcal{M}(g) \mathbb{E}[G^2]}\ , \tag{3.2.19}$$

where (3.2.18) follows from the same argument as in Section 3.1.2 and (3.2.19) can be deduced from Lemma 3.2.1 and (3.2.7); finally, the third term in (3.2.17) can be bounded by $\|f\|^2 (2q)! \|g\widetilde{\otimes} g \mathbf{I}_{\Delta_{2q}^c}\| \leq \|f\|^2 \sqrt{(2q)! \gamma_q \mathbb{E}[G^2] \mathcal{M}(g)}$; and by symmetry, we also have

$$(2q)! \left|\langle f\widetilde{\otimes} f, g\widetilde{\otimes} g \mathbf{I}_{\Delta_{2q}^c}\rangle\right| = (2q)! \left|\langle g\widetilde{\otimes} g, f\widetilde{\otimes} f \mathbf{I}_{\Delta_{2p}^c}\rangle\right| \leq \|g\|^2 \sqrt{(2p)! \gamma_p \mathbb{E}[F^2] \mathcal{M}(f)}\,.$$

Now we can conclude our proof by combining the above estimates. □

3.3 Universality results on homogeneous sums

For the convenience of readers, we restate Theorem 1.4.3 here.

Theorem 1.4.3 Fix integers $d \geq 2$ and $q_d \geq \ldots \geq q_1 \geq 2$. For each $j \in [d]$ and $n \in \mathbb{N}$, let the kernels $f_{j,n} \in \mathfrak{H}_0^{\odot q_j}$ satisfy $\sup_{n,j} \|f_{j,n}\|^2 < +\infty$ and for $k, l \in [d]$

$$\lim_{n\to+\infty} \mathbf{I}_{(q_k = q_l)} q_k! \sum_{i_1,\ldots,i_{q_k} \in \mathbb{N}} f_{k,n}(i_1, \ldots, i_{q_k}) f_{l,n}(i_1, \ldots, i_{q_k}) = \Sigma_{k,l}\,,$$

where Σ is a symmetric nonnegative definite d by d matrix. Then the following statements are equivalent, as $n \to +\infty$:

(C_1) Given a sequence $\mathbf{G}$ of i.i.d. standard Gaussians,

$$\left(Q_{q_1}(f_{1,n}; \mathbf{G}), \ldots, Q_{q_d}(f_{d,n}; \mathbf{G})\right)^T$$

converges in law to $\mathcal{N}(0, \Sigma)$.

(C_2) Given a sequence $\mathbf{V}$ of i.i.d. random variables with $V_1 + 1 \sim \text{Poi}(1)$,

$$(Q_{q_1}(f_{1,n}; \mathbf{V}), \ldots, Q_{q_d}(f_{d,n}; \mathbf{V}))^T$$

converges in law to $\mathcal{N}(0, \Sigma)$.

(C_3) In the ***symmetric*** Rademacher setting,

$$(Q_{q_1}(f_{1,n}; \mathbf{Y}), \ldots, Q_{q_d}(f_{d,n}; \mathbf{Y}))^T$$

converges in law to $\mathcal{N}(0, \Sigma)$ and $\mathcal{M}(f_{1,n}) + \ldots + \mathcal{M}(f_{d,n}) \to 0$.

(C_4) For ***every*** sequence $\Xi = (\xi_i, i \in \mathbb{N})$ of independent centered random variables with unit variance and $\sup_{i\in\mathbb{N}} \mathbb{E}[|\xi_i|^3] < +\infty$,

$$(Q_{q_1}(f_{1,n}; \Xi), \ldots, Q_{q_d}(f_{d,n}; \Xi))^T$$

converges in law to $\mathcal{N}(0, \Sigma)$.

Remark 3.3.1. The homogeneous sums over independent Gaussian and Poisson random variables are members of Gaussian Wiener chaos and Poisson Wiener chaos respectively. One can realize them on a common probability space as follows. Taking $\mathcal{Z} = \mathbb{R}$ and μ the standard Lebesgue measure on $\mathbb{R}$, we set $g_i = \mathbf{I}_{[i,i+1)}$ for each $i \in \mathbb{N}$. Given $f \in \mathfrak{H}_0^{\odot d}$, the homogeneous sum $Q_d(f; \mathbf{G})$ is equal to $I_d^W(\widehat{f})$ in law, where

$$\widehat{f} := \sum_{i_1,\ldots,i_d \in \mathbb{N}} f(i_1, \ldots, i_d) g_{i_1} \otimes \cdots \otimes g_{i_d}\,, \tag{3.3.1}$$

and W is a Gaussian random measure on $\mathbb{R}$ with intensity measure μ. Similarly, we can write $Q_d(f; \mathbf{V}) = I_d^\eta(\widehat{f})$ in law, where η is a Poisson random measure on $\mathbb{R}$ with intensity measure μ. From now on, we will identity f with $\widehat{f}$ in case of no confusion.

Proof of Theorem 1.4.3. The implication $(C_1) \Rightarrow (C_4)$ has been proved in [7] and let us now recall the truncation argument therein. For each $j \in [d]$ and $n \in \mathbb{N}$, we can truncate the kernel $f_{j,n}$ as follows:

$$f_{j,n}^{\mathbf{t_n}} = f_{j,n} \mathbf{I}_{[N_n]^{q_j}}$$

where $N_n \in \mathbb{N}$ diverges to infinity such that $\| f_{j,n}^{\mathbf{t_n}} - f_{j,n} \| \le 1/n$. In this way, we have

$$\mathbb{E}\Big([Q_{q_j}(f_{j,n}; \mathbf{G}) - Q_{q_j}(f_{j,n}^{\mathbf{t_n}}; \mathbf{G})]^2 \Big) = \mathbb{E}\Big([Q_{q_j}(f_{j,n}; \Theta) - Q_{q_j}(f_{j,n}^{\mathbf{t_n}}; \Theta)]^2 \Big) \le \frac{q_j!}{n^2}\,. \tag{3.3.2}$$

Assume (C_1), then $Q_{q_j}(f_{j,n}^{\mathbf{t_n}}; \mathbf{G})$ converges in law to $\mathcal{N}(0, \Sigma_{j,j})$ for each $j \in [d]$. Then it follows from Theorem 1.3.2 that $(Q_{q_1}(f_{1,n}^{\mathbf{t_n}}; \Xi), \ldots, Q_{q_d}(f_{d,n}^{\mathbf{t_n}}; \Xi))^T$ converges in law to $\mathcal{N}(0, \Sigma)$, thus by (3.3.2),

$$(Q_{q_1}(f_{1,n}; \Xi), \ldots, Q_{q_d}(f_{d,n}; \Xi))^T$$

converges in law to $\mathcal{N}(0, \Sigma)$.

Now let us prove the implication $(C_2) \Rightarrow (C_4)$. Assume (C_2), then $(Q_{q_1}(f_{1,n}^{\mathbf{t_n}}; \mathbf{V}), \ldots, Q_{q_d}(f_{d,n}^{\mathbf{t_n}}; \mathbf{V}))^T$ converges in law to $\mathcal{N}(0, \Sigma)$. It follows from Theorem 1.3.4 that $\mathbb{E}[Q_{q_j}(f_{j,n}^{\mathbf{t_n}}; \mathbf{V})^4] \to 3\Sigma_{j,j}^2$ for each $j \in [d]$. Moreover, it follows from the relation (3.1.17) that

$$\|\widehat{f_{j,n}^{\mathbf{t_n}}} \otimes_k \widehat{f_{j,n}^{\mathbf{t_n}}}\| \to 0$$

for each $k = 1, \ldots, q_j - 1$ and every $j \in [d]$. Thus, by the Peccati-Tudor theorem, we have

$$(Q_{q_1}(f_{1,n}^{\mathbf{t_n}}; \mathbf{G}), \ldots, Q_{q_d}(f_{d,n}^{\mathbf{t_n}}; \mathbf{G}))^T$$

converges in law to $\mathcal{N}(0, \Sigma)$, so does $(Q_{q_1}(f_{1,n}; \Xi), \ldots, Q_{q_d}(f_{d,n}; \Xi))^T$.

Now let us prove the implication $(C_3) \Rightarrow (C_4)$. Assume (C_3): as a consequence of the product formula (2.2.4), $\{Q_{q_j}(f_{j,n}; \mathbf{Y})^4, n \in \mathbb{N}\}$ is uniformly integrable for each $j \in [d]$, so that

$$\mathbb{E}[Q_{q_j}(f_{j,n}; \mathbf{Y})^4] \to 3\Sigma_{j,j}^2 ,$$

thus, the relations (3.2.2) and (3.2.3) in Lemma 3.2.1 imply $\|f_{j,n} \otimes_k f_{j,n}\| \to 0$ for each $k = 1, \ldots, q_j - 1$ and every $j \in [d]$. Hence (C_1) holds true and so does (C_4).

To conclude our proof, it suffices to show the implication $(C_1) \Rightarrow (C_3)$. Suppose (C_1) holds true, then

$$(Q_{q_1}(f_{1,n}; \mathbf{Y}), \ldots, Q_{q_d}(f_{d,n}; \mathbf{Y}))^T$$

converges in distribution to $\mathcal{N}(0, \Sigma)$ by the equivalence $(C_1) \Leftrightarrow (C_4)$. By the FMT on a Gaussian space, (C_1) implies that $\|f_{j,n} \otimes_{q_j - 1} f_{j,n}\| \to 0$, as $n \to +\infty$. Recall from [70, Lemma 2.4] that $\mathcal{M}(f) \leq \|f \otimes_{d-1} f\|$ for each $f \in \mathfrak{H}_0^{\odot d}$, therefore $\mathcal{M}(f_{j,n}) \to 0$ for each $j \in [d]$. This proves the implication $(C_1) \Rightarrow (C_3)$. □

Bibliography

[1] H. Airault, P. Malliavin and F. G. Viens. Stokes formula on the Wiener space and n-dimensional Nourdin-Peccati analysis. *J. Funct. Anal.*, Vol. 258, Issue 5, (2010), 1763–1783.

[2] J.-M. Azaïs, F. Dalmao and J.R. León. CLT for the zeros of classical random trigonometric polynomials, *Ann. Inst. H. Poincaré Probab. Statist.* vol. 52, No. 2 (2016), 804-820.

[3] J.-M. Azaïs and J.R. León. CLT for crossing of random trigonometric polynomials, *Electron. J. Probab.* Vol. 18, No. 68, 1–17, 2013.

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

[5] E. Azmoodeh, D. Malicet, G. Mijoule and G. Poly. Generalization of the Nualart-Peccati criterion. *Ann. Probab.* Vol. 44, No. 2, 924–954, 2016.

[6] S. Bai and M. S. Taqqu. Multivariate limit theorems in the context of long-range dependence, *J. Time Series Anal.* Vol. 34, No. 6, 717–743, 2013.

[7] S. Bai and M. S. Taqqu. The universality of homogeneous polynomial forms and critical limits. *J. Theoret. Probab.*, 29(4):1710–1727, 2016.

[8] D. Bakry, I. Gentil and M. Ledoux. *Analysis and geometry of Markov diffusion operators*. Volume 348 of Grundlehren der Mathematischen [Fundamental Principles of Mathematical Sciences]. Springer, Cham, 2014.

[9] A. D. Barbour. Stein's method and Poisson process convergence. *J. Appl. Probab.* **25A**, 175–184, 1988.

[10] A. D. Barbour. Stein's method for diffusion approximations. *Probab. Th. Rel. Fields* **84**, 297–322, 1990.

[11] A. D. Barbour and L. H. Y. Chen. *An Introduction to Stein's Method.* Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore: Volume 4, 2004.

[12] A. D. Barbour and L. H. Y. Chen. *Stein's Method and Applications.* Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore: Volume 5, 2004.

[13] P. Billingsley. *Probability and measures*, 3rd edition. Wiley series in probability and mathematical statistics, John Wiley & sons. 1995.

[14] S. G. Bobkov, G. P. Chistyakov and F. Götze. Rényi divergence and central limit theorem. preprint, 2016.

[15] S. Bourguin and G. Peccati. Semicircular limits on the free Poisson chaos: counterexamples to a transfer principle. *J. Funct. Anal.*, 267 (4), 963–997, 2014.

[16] R.H. Cameron and W. T. Martin. The orthogonal development of non-linear functionals in series of Fourier-Hermite functions. *Annals of Mathematics*, Second Series, Vol. 48, No. 2, pp. 385-392, 1947.

[17] S. Campese, I. Nourdin, G. Peccati and G. Poly. Multivariate Gaussian approximations on Markov chaoses. *Electron. Commun. Probab.*, Vol. 21 No. 48, 9, 2016.

[18] L. Chaumont and M. Yor. *Exercises in Probability, A Guided Tour from Measure Theory to Random Processes, via Conditioning*. Cambridge University Press, 2nd edition. 2012.

[19] L.H.Y. Chen, L. Goldstein and Q.M. Shao. *Normal approximation by Stein's method.* Probability and Its Applications, 2011 Springer-Verlag Berlin Heidelberg.

[20] S. Chatterjee. A short survey of Stein's method. *Proceedings of the ICM*, 2014.

[21] S. Chatterjee and Q.-M. Shao. Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model. *Ann. Appl. Probab.*, 21(2):464 - 483, 2011.

[22] S. Chatterjee, J. Fulman and A. Röllin. Exponential approximation by Stein's method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.*, 8:197 - 223, 2011.

[23] S. Chatterjee and E. Meckes. Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:257-283, 2008.

[24] L.H.Y. Chen and G. Poly. Stein's method, Malliavin calculus, Dirichlet forms and the fourth moment theorem. In: *Z-Q Chen, N. Jacob, M. Takeda and T. Uemura (eds.), Festschrift Masatoshi Fukushima*, Interdisciplinary Mathematical Sciences Vol. 17, pp. 107-130, World Scientific, 2015.

[25] F. Dalmao. Asymptotic variance and CLT for the number of zeros of Kostlan Shub Smale random polynomials. *Comptes Rendus Mathématique*, Volume 353, Issue 12 (2015), Pages 1141-1145

[26] P. de Jong. A central limit theorem for generalized multilinear forms. *J. Multivariate Anal.* **34** 275-289.

[27] P. Diaconis. The distribution of leading digits and uniform distribution mod 1. *Ann. Probability*, 5(1):72–81, 1977.

[28] C. Döbler. New developments in Stein's method with applications. *Ph.D dissertation*, Ruhr-Universität Bochum (2012). Available online at *https://tinyurl.com/y834hxx7*

[29] C. Döbler. Stein's method of exchangeable pairs for the Beta distribution and generalizations. *Electron. J. Probab.*, Vol. 20 no. 109, 1-34, 2015.

[30] C. Döbler and K. Krokowski. On the fourth moment condition for Rademacher chaos. To appear in *Ann. Inst. Henri Poincar Probab. Stat.*, 2017+

[31] C. Döbler and G. Peccati. Quantitative de Jong theorems in any dimension. *Electron. J. Probab.* **22** (2017), no.2, 1-35.

[32] C. Döbler and G. Peccati. The fourth moment theorem on the Poisson space. *Ann. Probab.* to appear (2017+)

[33] C. Döbler and M. Stolz. Stein's method and the multivariate CLT for traces of powers on the classical compact groups. *Electron. J. Probab.*, Vol. 16, no. 86, 2375-2405, 2011.

[34] C. Döbler, A. Vidotto and G. Zheng. Fourth moment theorems on the Poisson space in any dimension. preprint, 2017.

[35] R. M. Dudley. *Real Analysis and Probability.* 2nd edition, Cambridge University Press, 2010.

[36] P. Eichelsbacher and M. Löwe. Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.*, 15:no. 30, 962–988, 2010.

[37] L. Goldstein, I. Nourdin, and G. Peccati. Gaussian phase transitions and conic intrinsic volumes: Steining the Steiner formula. *Ann. Appl. Probab.* Vol. 27, No. 1, 1-47, 2017.

[38] F. Götez. On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19**(2), 724–739, 1991.

[39] D. Hug, G. Last, and M. Schulte. Second-order properties and central limit theorems for geometric functionals of Boolean models. *Ann. Appl. Probab.* Volume 26, Number 1 (2016), 73-135.

[40] K. Itô. Multiple Wiener Integral. *Journal of the Mathematical Society of Japan*. Vol. 3, No. 1, May, 1951.

[41] K. Itô. Spectral Type of the Shift Transformation of Differential Processes With Stationary Increments. *Transactions of the American Mathematical Society*, Vol. 81, No. 2, pp.253 - 263, 1956.

[42] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes.* Grundlehren der Mathematischen Wissenschaften, vol. 288, Springer-Verlag, Berlin, Heidelberg, New York, 1987

[43] K. Krokowski. Poisson approximation of Rademacher functionals by the Chen-Stein method and Malliavin calculus. *Commun. Stoch. Anal.* 11 (2017), no. 2, 195-222

[44] K. Krokowski, A. Reichenbachs and C. Thäle. Berry-Esseen bounds and multivariate limit theorems for functionals of Rademacher sequences. *Ann. Inst. Henri Poincaré Probab. Stat.* 52 (2016), no. 2, 763-803

[45] K. Krokowski, A. Reichenbachs and C. Thäle. Discrete Malliavin-Stein method: Berry-Esseen bounds for random graphs and percolation. *Ann. Probab.* 45 (2017), no. 2, 1071-1109

[46] K. Krokowski and C. Thäle. Multivariate central limit theorems for Rademacher functionals with applications, *Electron. J. Probab.* 22 (2017), no. 87, 1–30

[47] G. Last. Stochastic analysis for Poisson processes. In G. Peccati and M. Reitzner, editors, *Stochastic analysis for Poisson point processes*, Mathematics, Statistics, Finance and Economics, chapter 1, pages 1 – 36. Bocconi University Press and Springer, 2016.

[48] G. Last, G. Peccati, and M. Schulte. Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. *Probab. Theory Related Fields*, 165(3-4):667–723, 2016.

[49] G. Last and M. Penrose. *Lectures on the Poisson Process*. Cambridge University Press. 2017.

[50] A. Lechiheb, I. Nourdin, G. Zheng and E. Haouala. Convergence of random oscillatory integrals in the presence of long-range dependence and application to homogenization, *Probab. Math. Stat.* In press.

[51] M. Ledoux. Chaos of a Markov operator and the fourth moment condition. *Ann. Probab.* Vol. 40, No. 6, 2439–2459, 2012.

[52] D. Malicet, I. Nourdin, G. Peccati and G. Poly. Squared chaotic random variables: New moment inequalities with applications. *J. Funct. Anal.* 270 (2) 649–670, 2016.

[53] P. Malliavin. Stochastic calculus of variations and hypoelliptic operators. In: *Proc. Inter. Symp. on Stoch. Diff. Equations*, Kyoto 1976, Wiley 1978, 195-263.

[54] D. Marinucci and G. Peccati. Ergodicity and Gaussianity for Spherical Random Fields, *J. Math. Phys.* 51, (2010)

[55] D. Marinucci, G. Peccati, M. Rossi and I. Wigman. Non-Universality of Nodal Length Distribution for Arithmetic Random Waves, *Geom. Funct. Anal.* (2016).

[56] D. Marinucci and M. Rossi. Stein-Malliavin approximations for nonlinear functionals of random eigenfunctions on $\mathbf{S}^d$, *J. Funct. Anal.* 268 (8) (2015), 2379-2420.

[57] E. Meckes. An Infinitesimal Version of Stein's Method of Exchangeable Pairs. *Ph.D dissertation*, Stanford University (2006)

[58] E. Meckes. On Stein's method for multivariate normal approximation. *IMS collections, High dimensional Probability V*: The Luminy Volume, Vol. 5 (2009) 153-178.

[59] E. Mossel, R. O'Donnell and K. Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. of Math.* (2) (2010) 171 295-341

[60] S. Noreddine and I. Nourdin. On the Gaussian approximation of vector-valued multiple integrals. *J. Multivariate Anal.*, 102(6):1008–1017, 2011.

[61] I. Nourdin. Malliavin-Stein approach: a webpage maintained by Ivan Nourdin. `http://tinyurl.com/kvpdgcy`

[62] I. Nourdin. Lectures on Gaussian approximations with Malliavin calculus. *Séminaire de Probabilités* XLV, Lecture Notes in Math., Vol. 2078, Springer, Cham, 2013, pp. 3–89.

[63] I. Nourdin. Yet another proof of the Nualart-Peccati criterion. *Electron. Commun. Probab.* Vol. 16, No. 42, 467–481, 2011.

[64] I. Nourdin, D. Nualart and G. Peccati. Strong asymptotic independence on Wiener chaos. *Proc. Amer. Math. Soc.* 144, 875-886, 2016

[65] I. Nourdin and G. Peccati. Noncentral convergence of multiple integrals. *Ann. Probab.* Vol. 37, No. 4, 1412 –1426, 2009.

[66] I. Nourdin and G. Peccati. Stein's method on Wiener chaos, *Probab. Theory Relat. Fields* (2009), Vol. 145, Issue 1, p. 75-118.

[67] I. Nourdin and G. Peccati. Normal approximations with Malliavin calculus: from Stein's method to universality, *Cambridge tracts in Mathematics*, Vol. 192, 2012, Cambridge University Press.

[68] I. Nourdin, G. Peccati, G. Poly and R. Simone. Classical and free fourth moment theorems: universality and thresholds. *J. Theoret. Probab.*, 29(2):653–680, 2016.

[69] I. Nourdin, G. Peccati and A. Réveillac. Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Instit. H. Poincaré - Probab. Statist.* 46, no. 1 (2010), 45-58.

[70] I. Nourdin, G. Peccati and G. Reinert. Stein's Method and Stochastic Analysis of Rademacher Functionals. *Electron. J. Probab.* Vol. 15, No. 55, 1703 –1742, 2010

[71] I. Nourdin, G. Peccati and G. Reinert. Invariance principles for homogeneous sums: Universality of Gaussian Wiener chaos. *Ann. Probab.* Vol. 38, No. 5, 1947 – 1985, 2010.

[72] I. Nourdin and J. Rosiński. Asymptotic independence of multiple Wiener-Itô integrals and the resulting limit laws. *Ann. Probab.* 42, no. 2 (2014), 497-526

[73] I. Nourdin and G. Zheng. Exchangeable pairs on Wiener chaos. preprint, 2017

[74] D. Nualart. The Malliavin Calculus and Related Topics, second edition. *Probability and Its Applications*, Springer-Verlag Berlin Heidelberg (2006)

[75] D. Nualart and S. Ortiz-Latorre. Central limit theorems for multiple stochastic integrals and Malliavin calculus, *Stoch. Proces. Appl.* Vol. 118 (4), 2008, Pages 614–628

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

[77] D. Nualart and J. Vives. Anticipative calculus for the Poisson process based on the Fock space. *Séminaire de probabilités de Strasbourg*, Vol. 24, page 154 –165. 1990.

[78] R. O'Donnell. Analysis of Boolean functions. *Cambridge University Press*, 2014

[79] G. Peccati. Quantitative CLTs on a Gaussian space: a survey of recent developments. *ESAIM: Proceedings* 44: 61-78, 2014

[80] G. Peccati and M. Reitzner. *Stochastic analysis for Poisson point processes*, Mathematics, Statistics, Finance and Economics. Bocconi University Press and Springer, 2016.

[81] G. Peccati and M. Rossi. Quantitative limit theorems for local functionals of arithmetic random waves. *The Abel Symposium* 2016 – Springer (to appear)

[82] G. Peccati, J. L. Solé, M. S. Taqqu and F. Utzet. Stein's method and Normal approximation of Poisson functionals. *Ann. Probab.* Vol. 38, No. 2 (2010), 443–478.

[83] G. Peccati and M. S. Taqqu. Central limit theorems for double Poisson integrals. *Bernoulli*, **14** (3), 2008, 791–821.

[84] G. Peccati and C.A. Tudor. Gaussian limits for vector-valued multiple stochastic integrals. *Séminaire de Probabilités* XXXVIII, 2005

[85] G. Peccati and M. Yor. Hardy's inequality in $L^2([0,1])$ and principal values of Brownian local times. In *Asymptotic Methods in Stochastics* 75-87. Amer. Math. Soc., Providence, RI. 2004

[86] G. Peccati and C. Zheng. Universal Gaussian fluctuations on the discrete Poisson chaos. *Bernoulli*, 20(2):697-715, 2014

[87] J. Picard. Formules de dualités sur l'espace de Poisson. *Annales de l'I.H.P. Probabilités et statistiques* (1996) Volume: 32, Issue: 4, page 509-548

[88] N. Privault. Stochastic analysis of Bernoulli processes. *Probability Surveys* Vol. 5, 435–483, 2008.

[89] N. Privault. Independence of some multiple Poisson stochastic integrals with variable-sign kernels. In: Advances in Statistics, Probability and Actuarial Science: Volume 1, Stochastic Processes, Finance and Control – A Festschrift in Honor of Robert J. Elliott, pp. 73–86.

[90] G. Reinert and A. Röllin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.*, 37(6):2150–2173, 2009.

[91] M. Reitzner and M. Schulte. Central Limit Theorems for U-Statistics of Poisson Point Processes, *Ann. Probab.* Vol. 41, No. 6 (2013), 3879–3909

[92] Y. Rinott and V. Rotar. A multivariate CLT for local dependence with $n^{-1/2}\log n$ rate and applications to multivariate graph related statistics. *J. Multivariate Anal.* **56** (2), 333–350, 1996.

[93] Y. Rinott and V. Rotar. On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics. *Ann. Appl. Probab.*, 7(4):1080–1105, 1997.

[94] A. Röllin. A note on the exchangeability condition in Stein's method. *Statist. Probab. Lett.*, 78(13):1800–1806, 2008.

[95] J. Rosiński. Conference slides, page 3-4. Available at https://tinyurl.com/yamlwmjp, 2010.

[96] J. Rosiński and G. Samorodnitsky. Product formula, tails and independence of multiple stable integrals. In *Advances in Stochastic Inequalities* (Atlanta, GA, 1997). *Contemp. Math.* 234, 169 –194. Amer. Math. Soc., Providence, RI. 1999.

[97] M. Rossi. The Geometry of Spherical Random Fields. *Ph.D Thesis*, Univ. Rome Tor Vergata, 2016.

[98] M. Schulte. Normal approximation of Poisson functionals in Kolmogorov distance. *J. Theoret. Probab.*, 29(1):96–117, 2016.

[99] M. Schulte. A Central Limit Theorem for the Poisson-Voronoi Approximation. *Adv. Appl. Math.* 49, no. 3-5 (2012), 285-306

[100] Q.-M. Shao and Z.-G. Su. The Berry-Esseen bound for character ratios. *Proc. Amer. Math. Soc.*, 134(7):2153–2159 (electronic), 2006.

[101] C. Stein. Approximate computation of expectations. In *Institute of Mathematical Statistics Lecture Notes - Monograph Series*, volume 7. Institute of Mathematical Statistics, 1986.

[102] C. Stein. The accuracy of the normal approximation to the distribution of the traces of powers of random orthogonal matrices. *Technical Report* No. 470, Stanford University, Department of Statistics, 1995.

[103] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, *Berkeley Symp. on Math. Statist. and Prob. Proc. Sixth Berkeley Symp. on Math. Statist. and Prob., Vol. 2* (Univ. of Calif. Press, 1972), 583-602.

[104] D. Revuz and M. Yor. *Brownian motions and continuous martingales*. 3rd edition. Grundlehren der mathematischen Wissenschaften, Vol. 293. Springer-Verlag 2005.

[105] N. Ross. Fundamentals of Stein's method. *Probability Surveys.* Vol. 8, 210-293, 2011.

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

[107] M. Veraar. The stochastic Fubini theorem revisited. *Stochastics*, 84(4): 543 - 551, 2012.

[108] N. Wiener. The homogeneous chaos. *Amer. Journ. Math.* Vol. LV, No.4, 1938.

[109] G. Zheng. Normal approximation and almost sure central limit theorem for non-symmetric Rademacher functionals, *Stoch. Process. Appl.* Vol. 127, no.5, 1622-1636, 2017.

[110] G. Zheng. A Peccati-Tudor type theorem for Rademacher chaos. preprint, 2017.

Part II

This page is left blank.

Paper 1: Normal approximation and almost sure central limit theorem for non-symmetric Rademacher functionals

Guangqu Zheng*

Published in Stoch. Process. Appl. Volume 127, Issue 5, May 2017, Pages 1622-1636.

Abstract

In this work, we study the normal approximation and almost sure central limit theorems for some functionals of an independent sequence of Rademacher random variables. In particular, we provide a new chain rule that improves the one derived by Nourdin, Peccati and Reinert(2010) and then we deduce the bound on Wasserstein distance for normal approximation using the (discrete) Malliavin-Stein approach. Besides, we are able to give the almost sure central limit theorem for a sequence of random variables inside a fixed Rademacher chaos using the *Ibragimov-Lifshits criterion*.

1 Introduction

This work is devoted to the study of discrete Malliavin-Stein approach for two kinds of Rademacher functionals:

(S) $Y_k, k \in \mathbb{N}$ is a sequence of independent identically distributed (*i.i.d*) Rademacher random variables, *i.e.* $\mathbb{P}(Y_1 = -1) = \mathbb{P}(Y_1 = 1) = 1/2$. $F = f(Y_1, Y_2, \cdots)$, for some nice function f, is called a (symmetric) Rademacher functional over (Y_k).

(NS) $X_k, k \in \mathbb{N}$ is a sequence of independent non-symmetric, non-homogeneous Rademacher random variables, that is, $\mathbb{P}(X_k = 1) = p_k$, $\mathbb{P}(X_k = -1) = q_k$ for each $k \in \mathbb{N}$. Here $1 - q_k = p_k \in (0, 1)$ for each $k \in \mathbb{N}$. Of course this sequence reduces to the i.i.d. one when $p_k = q_k = 1/2$ for each k. $G = f(X_1, X_2, \cdots)$, for some nice function f, is called a (non-symmetric) Rademacher functional over (X_k). Usually, we consider the corresponding normalised sequence $(Y_k, k \in \mathbb{N})$ of X_k, that is, $Y_k := (X_k - p_k + q_k) \cdot (2\sqrt{p_k q_k})^{-1}$.

From now on, we write (S) for the symmetric setting, and (NS) for the non-symmetric, non-homogeneous setting.

Now let us explain several terms in the title. Malliavin-Stein method stands for the combination of two powerful tools in probability theory: Paul Malliavin's differential calculus and Charles

2010 *Mathematics Subject Classification*. Primary: 60F05, 60B12; Secondary: 47N30.

Key words and phrases. Rademacher functional ; Normal approximation ; Wasserstein distance ; Almost sure central limit theorem ; Malliavin-Stein approach

*Email: guangqu.zheng@uni.lu

Stein's method of normal approximation. This intersection originates from the seminal paper [12] by Nourdin and Peccati, who were able to associate a quantitative bound to the remarkable fourth moment theorem established by Nualart and Peccati [15] among many other things. For a comprehensive overview, one can refer to the website [11] and the recent monograph [13].

This method has found its extension to discrete settings: for the Poisson setting, see *e.g.* [16, 20]; for the Rademacher setting, the paper [14] by Nourdin, Peccati and Reinert was the first one to carry out the analysis of normal approximation for Rademacher functionals (possibly depending on infinitely many Rademacher variables) in the setting (S), and they were able to get a sufficient condition in terms of contractions for a central limit theorem (CLT) inside a fixed Rademacher chaos $\mathscr{C}_m$ (with $m \geq 2$), see Proposition 2.2 for the precise statement.

In the Rademacher setting, unlike the Gaussian case, one does not have the chain rule like $Df(F) = f'(F)DF$ for $f \in C_b^1(\mathbb{R})$ and Malliavin differentiable random variable F (see [13, Proposition 2.3.7]), while an approximate chain rule (see (1.13)) is derived in [14] and it requires quite much regularity of the function f. As a consequence, the authors of [14] had to use smooth test functions when they applied the Stein's estimation: roughly speaking, for nice centred Rademacher functional F in the setting (S), for $h \in C_b^2(\mathbb{R})$, $Z \sim \mathscr{N}(0,1)$, one has (see [14, Theorem 3.1])

$$\begin{aligned} &\Big|\mathbb{E}[h(F) - h(Z)]\Big| \\ &\leq \min\left(4\|h\|_\infty, \|h''\|_\infty\right) \cdot \mathbb{E}\Big[\big|1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\big|\Big] + \frac{20}{3}\|h''\|_\infty \mathbb{E}\Big[\big\langle |DL^{-1}F|, |DF|^3\big\rangle_{\mathfrak{H}}\Big], \end{aligned} \tag{1.1}$$

where the precise meaning of the above notation will be explained in the Section 2.

Krokowski, Reichenbachs and Thäle, carefully using a representation of the discrete Malliavin derivative $Df(F)$ and the fundamental theorem of calculus instead of the approximate chain rule (1.13), were able to deduce the Berry-Esséen bound in [8, Theorem 3.1] and its non-symmetric analogue in [9, Proposition 4.1]: roughly speaking, for nice centred Rademacher functional F in the setting (NS),

$$d_K(F,Z) := \sup_{x\in\mathbb{R}} \Big|\mathbb{P}(F \leq x) - \mathbb{P}(Z \leq x)\Big| \tag{1.2}$$

$$\leq \mathbb{E}\Big[\big|1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\big|\Big] + \frac{\sqrt{2\pi}}{8}\mathbb{E}\Big[\big\langle \frac{1}{\sqrt{pq}}|DL^{-1}F|, |DF|^2\big\rangle_{\mathfrak{H}}\Big] \tag{1.3}$$

$$+ \frac{1}{2}\mathbb{E}\Big[\big\langle |F\cdot DL^{-1}F|, \frac{1}{\sqrt{pq}}|DF|^2\big\rangle_{\mathfrak{H}}\Big] + \sup_{x\in\mathbb{R}}\mathbb{E}\Big[\big\langle |DL^{-1}F|, \frac{1}{\sqrt{pq}}(DF)\cdot \mathbf{I}_{(F>x)}\big\rangle_{\mathfrak{H}}\Big]. \tag{1.4}$$

The quantity $d_K(F,Z)$ defined in (1.2) is called the Kolmogorov distance between F and Z.

For the setting (NS), the corresponding analysis including normal approximation and Poisson approximation has been taken up in [5, 7, 9, 18].

In this paper, we give a neat chain rule (see 1.2), from which we are able to derive a bound on the Wasserstein distance

$$d_W(F,Z) := \sup_{\|f'\|_\infty \leq 1} \Big|\mathbb{E}[f(F) - f(Z)]\Big|$$

in both settings (NS) and (S), see 2.1 and related remarks.

Another contribution of this work is the almost sure central limit theorem (ASCLT in the sequel) for Rademacher functionals. We first give the following

Definition 1.1. Given a sequence $(G_n, n \in \mathbb{N})$ of real random variables convergent in law to $Z \sim \mathscr{N}(0,1)$, we say the ASCLT holds for (G_n), if almost surely, for any bounded continuous $f : \mathbb{R} \to \mathbb{R}$, we have

$$\frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} f(G_k) \longrightarrow \mathbb{E}[f(Z)] \ , \tag{1.5}$$

as $n \to +\infty$. In the definition, $\log n$ can be replaced by $\gamma_n := \sum_{k=1}^{n} k^{-1}$, $(\gamma_n - 1 \le \log n \le \gamma_n)$. Note the condition (1.5) is equivalent to that the random probability measure $\gamma_n^{-1} \sum_{k=1}^{n} k^{-1} \cdot \delta_{G_k}$ weakly converges to the standard Gaussian measure almost surely, as $n \to +\infty$.

The following criterion, due to Ibragimov and Lifshits, gives a sufficient condition for the ASCLT.

Ibragimov-Lifshits criterion

$$\sup_{|t| \le r} \sum_{n \ge 2} \frac{\mathbb{E}(|\triangle_n(t)|^2)}{n \gamma_n} < +\infty \ , \quad \text{for every } r > 0, \tag{1.6}$$

where

$$\triangle_n(t) = \frac{1}{\gamma_n} \sum_{k=1}^{n} \frac{1}{k} \Big[e^{itG_k} - e^{-t^2/2} \Big]. \tag{1.7}$$

If $G_k \xrightarrow{\text{law}} Z \sim \mathscr{N}(0,1)$ and (1.6) is satisfied, then the ASCLT holds for (G_k). See [6, Theorem 1.1].

The ASCLT was first stated by Paul Lévy without proof in the 1937 book "*Théorie de l'addition des variables aléatoires* " and rediscovered by Brosamler [2], Schatte [19] independently in 1988. The present form appearing in the above definition was stated by Lacey and Philipp [10] in 1990. And in 1999, Ibragimov and Lifshits [6] gave the above sufficient condition.

Using this criterion, the authors of [1] established the ASCLT for functionals over general Gaussian fields. The Malliavin-Stein approach plays a crucial role in their work. Later, C. Zheng proved the ASCLT on the Poisson space in his Ph.D thesis [20, Chapter 5]. And in this work, we prove the ASCLT in the Rademacher setting, see Section 3.2.

The rest of this paper is organised as follows: Section 2 is devoted to some preliminary knowledge on Rademacher functionals, and we provide a simple but useful approximate chain rule there. In Section 3.1, we establish the Wasserstein distance bound for normal approximation in both setting (S) and (NS); in Section 3.2, the ASCLT for Rademacher chaos is established.

We fix several notation first: $\mathbb{N} = \{1, 2, \cdots\}$, $\mathbf{Y} = (Y_k, k \in \mathbb{N})$ stands for the Rademacher sequence in the setting (S), and it also means the normalised sequence in the setting (NS). Denote by $\mathscr{G}$ the σ-algebra generated by $\mathbf{Y}$, for notational convenience, we write $L^2(\Omega)$ for $L^2(\Omega, \mathscr{G}, \mathbb{P})$ in the sequel. We write $\mathfrak{H} = \ell^2(\mathbb{N})$ for the Hilbert space of square-summable sequences indexed by $\mathbb{N}$. $\mathfrak{H}^{\otimes n}$ means the n^{th} tensor product space and $\mathfrak{H}^{\odot n}$ its symmetric subspace. We denote $\mathfrak{H}_0^{\odot n} = \{ f \in \mathfrak{H}^{\odot n} \ : \ f|_{\triangle_n^c} = 0\}$ with $\triangle_n = \{(i_1, \cdots, i_n) \in \mathbb{N}^n \ : \ i_k \neq i_j$ for different $k, j\}$. Clearly, $\mathfrak{H}_0^{\odot 0} = \mathfrak{H}^0 = \mathbb{R}$. For $u, v, \mathrm{w} \in \mathfrak{H}$, we write $\langle u, v\mathrm{w} \rangle = \sum_{k \in \mathbb{N}} u_k v_k \mathrm{w}_k$.

1.1 Discrete Malliavin calculus

The basic reference for this section is the survey [17] by Privault.

Definition 1.2. The (discrete) n^{th} order multiple stochastic integral $J_n(f)$ of $f \in \mathfrak{H}^{\otimes n}$, $n \geq 1$, is given by

$$J_n(f) = \sum_{(i_1,\cdots,i_n)\in\Delta_n} f(i_1,\cdot\cdot\cdot,i_n) Y_{i_1} Y_{i_2} \cdot\cdot\cdot Y_{i_n} \ . \tag{1.8}$$

We define $J_0(c) = c$ for any $c \in \mathbb{R}$. It is clear that $J_n(f) = J_n(\widetilde{f}\mathbf{I}_{\Delta_n})$, where $\widetilde{f}$ is the standard symmetrisation of f.

For $g \in \mathfrak{H}_0^{\odot n}$, it is easy to check that $\|J_n(g)\|^2_{L^2(\Omega)} = n!\|g\|^2_{\mathfrak{H}^{\otimes n}}$ and $\mathscr{C}_n := \{J_n(g) \, : \, g \in \mathfrak{H}_0^{\odot n}\}$ is isometric to $(\mathfrak{H}_0^{\odot n}, \sqrt{n!}\| \cdot \|_{\mathfrak{H}^{\otimes n}})$. $\mathscr{C}_n$ is called the *Rademacher chaos* of order n, and one can see easily that $\mathscr{C}_n$ is a closed linear subspace of $L^2(\Omega)$ and $\mathscr{C}_n$, $\mathscr{C}_m$ are mutually orthogonal for distinct m, n:

$$\mathbb{E}[J_n(f) \cdot J_m(g)] = n! \cdot \langle f, g\rangle_{\mathfrak{H}^{\otimes n}} \cdot \mathbf{I}_{(m=n)}, \qquad \forall f \in \mathfrak{H}_0^{\odot n} \ , \ g \in \mathfrak{H}_0^{\odot m} \ . \tag{1.9}$$

More notation $\mathscr{C}_0 := \mathbb{R}$. We denote by $\mathcal{S}$ the linear subspace of $L^2(\Omega)$ spanned by multiple integrals and it is a well-known result (*e.g.* see [17, Proposition 6.7]) that $\mathcal{S}$ is dense in $L^2(\Omega)$. In particular, $F \in L^2(\Omega)$ can be expressed as follows:

$$F = \mathbb{E}[F] + \sum_{n\geq 1} J_n(f_n) \ , \quad \text{where } f_n \in \mathfrak{H}_0^{\odot n} \text{ for each } n \in \mathbb{N}. \tag{1.10}$$

We denote by $L^2(\Omega \times \mathbb{N})$ the space of square-integrable random sequences $a = (a_k, k \in \mathbb{N})$, where a_k is a real random variable for each $k \in \mathbb{N}$ and $\|a\|^2_{L^2(\Omega\times\mathbb{N})} := \mathbb{E}[\|a\|^2_{\mathfrak{H}}] = \sum_{k\geq 1} \mathbb{E}[a_k^2] < +\infty$.

Definition 1.3. $\mathbb{D}$ is the set of random variables $F \in L^2(\Omega)$ as in (1.10) satisfying $\sum_{\ell=1}^{\infty} \ell\ell!\|f_\ell\|^2_{\mathfrak{H}^{\otimes\ell}} < +\infty$. For $F \in \mathbb{D}$ as in (1.10), $D_k F = \sum_{n\geq 1} n J_{n-1}(f_n(\cdot, k))$ for each $k \in \mathbb{N}$. $DF = (D_k F, k \in \mathbb{N})$ is called the discrete Malliavin derivative of F.

Remark 1.1. Using (1.10) and (1.9), we can obtain the *Poincaré inequality* for $F \in \mathbb{D}$, $\mathrm{Var}(F) \leq \mathbb{E}[\|DF\|^2_{\mathfrak{H}}]$.

Definition 1.4. We define the *divergence operator* δ as the adjoint operator of D. We say $u \in \mathrm{dom}\delta \subset L^2(\Omega \times \mathbb{N})$ if there exists some constant C such that $|\mathbb{E}[\langle u, DF\rangle_{\mathfrak{H}}]| \leq C\|F\|_{L^2(\Omega)}$ for any $F \in \mathbb{D}$. Then it follows from the Riesz's representation theorem that there exists a unique element in $L^2(\Omega \times \mathbb{N})$, which we denote by $\delta(u)$, such that the duality relation (1.11) holds for any $F \in \mathbb{D}$:

$$\mathbb{E}[\langle u, DF\rangle_{\mathfrak{H}}] = \mathbb{E}[F\delta(u)] \, . \tag{1.11}$$

Definition 1.5. We define the *Ornstein-Uhlenbeck operator* L by $L = -\delta D$. Its domain is given by $\mathrm{dom}L = \{F \in L^2(\Omega)$ admits the chaotic decomposition as in (1.10) such that $\sum_{n=1}^{\infty} n^2 n! \cdot \|f_n\|^2_{\mathfrak{H}^{\otimes n}}$ is finite$\}$. For centred $F \in L^2(\Omega)$ as in (1.10), we define $L^{-1}F = -\sum_{n=1}^{\infty} n^{-1} J_n(f_n)$. It is clear that for such a F, one has $LL^{-1}F = F$. We call L^{-1} the pseudo-inverse of L.

Here is another look at the derivative operator D.

Remark 1.2. We choose $\Omega = \{+1,-1\}^{\mathbb{N}}$ and define $\mathbb{P} = \bigotimes_{k\in\mathbb{N}}(p_k\delta_{+1} + q_k\delta_{-1})$. Then the coordinate projections $\omega = \{\omega_1,\cdots\} \in \Omega \longmapsto \omega_k =: X_k(\omega)$ is an independent sequence of non-symmetric, non-homogeneous Rademacher random variables under $\mathbb{P}$. We can define for $F \in L^2(\Omega)$, $F^{\oplus k} := F(\omega_1,\cdots,\omega_{k-1},1,\omega_{k+1},\cdots)$, that is, by fixing the k^{th} coordinate in the configuration ω to be 1. Similarly, we define $F^{\ominus k} := F(\omega_1,\cdots,\omega_{k-1},-1,\omega_{k+1},\cdots)$. It holds that $D_kF = \sqrt{p_kq_k}(F^{\oplus k} - F^{\ominus k})$, see *e.g.* [17, Proposition 7.3]. The following results are also clear :

- $\left|F^{\oplus k} - F\right| = \mathbf{I}_{(X_k=-1)} \cdot |D_kF|/\sqrt{p_kq_k} \le |D_kF|/\sqrt{p_kq_k}$ and $\left|F^{\ominus k} - F\right| = \mathbf{I}_{(X_k=1)} \cdot |D_kF|/\sqrt{p_kq_k} \le |D_kF|/\sqrt{p_kq_k}$.
- $F \in \mathbb{D}$ if and only if $\sum_{k\in\mathbb{N}} p_kq_k\mathbb{E}[|F^{\oplus k} - F^{\ominus k}|^2] < +\infty$. In particular, if $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, then $f(F) \in \mathbb{D}$.

The following integration-by-part formula is important for our work.

Lemma 1.1. ([14, Lemma 2.12]) For every centred $F, G \in \mathbb{D}$ and $f \in C^1(\mathbb{R})$ with $\|f'\|_\infty < +\infty$, one has $f(F), L^{-1}F \in \mathbb{D}$ and $\mathbb{E}[Gf(F)] = \mathbb{E}[\langle -DL^{-1}G, Df(F)\rangle_{\mathfrak{H}}]$.

In particular, for $f(x) = x$, $\mathbb{E}[F^2] = \mathbb{E}[\langle -DL^{-1}F, DF\rangle_{\mathfrak{H}}]$. The random variable $\langle -DL^{-1}F, DF\rangle_{\mathfrak{H}}$ is crucial in the Malliavin-Stein approach, see *e.g.* [12] and [16].

The term $Df(F)$ is not equal to $f'(F)DF$ in general, unlike the chain rule on Gaussian Wiener space, see *e.g.* [13, Proposition 2.3.7]. The following is our new approximate chain rule.

Lemma 1.2. **(Chain rule)** If $F \in \mathbb{D}$ and $f : \mathbb{R} \longrightarrow \mathbb{R}$ is Lipschitz and of class C^1 such that f' is Lipschitz continuous, then

$$D_kf(F) = f'(F)D_kF + R_k, \tag{1.12}$$

where the remainder term R_k is bounded by $\dfrac{\|f''\|_\infty}{2\sqrt{p_kq_k}} \cdot |D_kF|^2$ in the setting (NS).

Proof. Note first $f(F) \in \mathbb{D}$, since f is Lipschitz. Moreover, since f is of class C^1 with Lipschitz derivative, it follows immediately that $f(y) - f(x) = f'(x)(y-x) + \mathcal{R}(f)$, where the remainder term $\mathcal{R}(f)$ is bounded by $\|f''\|_\infty \cdot |y-x|^2/2$. Therefore, in the setting (NS)

$$\begin{aligned} D_kf(F) &= \sqrt{p_kq_k} \cdot [f(F^{\oplus k}) - f(F^{\ominus k})] \\ &= \sqrt{p_kq_k} \cdot \left\{f(F^{\oplus k}) - f(F) - [f(F^{\ominus k}) - f(F)]\right\} \\ &= \sqrt{p_kq_k} \cdot \left\{ f'(F)(F^{\oplus k} - F) + \mathcal{R}_{1,k} - f'(F)(F^{\ominus k} - F) + \mathcal{R}_{2,k}\right\} \end{aligned}$$

with $|\mathcal{R}_{1,k}| \le \dfrac{\|f''\|_\infty \cdot |D_kF|^2}{2p_kq_k} \cdot \mathbf{I}_{(X_k=-1)}$ and $|\mathcal{R}_{2,k}| \le \dfrac{\|f''\|_\infty \cdot |D_kF|^2}{2p_kq_k} \cdot \mathbf{I}_{(X_k=1)}$.

Whence, $D_kf(F) = \sqrt{p_kq_k} \cdot f'(F)(F^{\oplus k} - F^{\ominus k}) + R_k = f'(F)D_kF + R_k$ and the remainder term $R_k = \sqrt{p_kq_k}(\mathcal{R}_{1,k} + \mathcal{R}_{2,k})$ is bounded by $\dfrac{\|f''\|_\infty \cdot |D_kF|^2}{2\sqrt{p_kq_k}}$. $\square$

It is clear that in the setting (S), the remainder R_k in (1.12) is bounded by $\|f''\|_\infty \cdot |D_kF|^2$.

Remark 1.3. In the setting (S), our approximate chain rule is different from that developed in [14], in which f is assumed to be of class C^3 such that $f(F) \in \mathbb{D}$ and $\|f'''\|_\infty < +\infty$. Moreover, their chain rule is given as follows:

$$D_k f(F) = f'(F) D_k F - \frac{1}{2}\Big[f''(F^{\oplus k}) + f''(F^{\ominus k})\Big] \cdot (D_k F)^2 \cdot Y_k + \widetilde{R}_k \tag{1.13}$$

with $\big|\widetilde{R}_k\big| \leq \frac{10}{3}\|f'''\|_\infty \big|D_k F\big|^3$. Apparently, ours is neater and requires less regularity of f. This is important when we try to get some nice distance bound in Section 3.1. Following [14], the authors of [18] gave an approximate chain rule in the setting (NS): if f is of class C^3 such that $f(F) \in \mathbb{D}$ and $\|f'''\|_\infty < +\infty$, then

$$D_k f(F) = f'(F) D_k F - \frac{|D_k F|^2}{4\sqrt{p_k q_k}}\Big[f''(F^{\oplus k}) + f''(F^{\ominus k})\Big] \cdot (D_k F)^2 \cdot X_k + R_k^F \tag{1.14}$$

with the remainder term R_k^F bounded by $\dfrac{5}{3!}\|f'''\|_\infty \dfrac{|D_k F|^3}{p_k q_k}$. See also Remark 2.1.

1.2 Star-contractions

Fix $m, n \in \mathbb{N}$, and $r = 0, \cdots, n \wedge m$. For $f \in \mathfrak{H}^{\otimes n}$ and $g \in \mathfrak{H}^{\otimes m}$, $f \star_r^r g$ is an element in $\mathfrak{H}^{\otimes n+m-2r}$ defined by

$$f \star_r^r g(i_1, \cdots, i_{n-r}, j_1, \cdots, j_{m-r}) = \sum_{a_1, \cdots, a_r \in \mathbb{N}} f(i_1, \cdots, i_{n-r}, a_1, \cdots, a_r) g(j_1, \cdots, j_{m-r}, a_1, \cdots, a_r) \ .$$

Lemma 1.3. Fix $\ell \in \mathbb{N}$ and $0 \leq r \leq \ell$. If $f, g \in \mathfrak{H}^{\odot \ell}$, then

$$2\big\|f \star_r^r g\big\|^2_{\mathfrak{H}^{\otimes 2\ell-2r}} \leq \big\|f \star_{\ell-r}^{\ell-r} f\big\|^2_{\mathfrak{H}^{\otimes 2r}} + \big\|g \star_{\ell-r}^{\ell-r} g\big\|^2_{\mathfrak{H}^{\otimes 2r}} \,.$$

In particular, $\big\|f \star_r^r g\big\|_{\mathfrak{H}^{\otimes 2\ell-2r}} \leq \big\|f \star_{\ell-r}^{\ell-r} f\big\|_{\mathfrak{H}^{\otimes 2r}} + \big\|g \star_{\ell-r}^{\ell-r} g\big\|_{\mathfrak{H}^{\otimes 2r}}$.

Proof. It follows easily from the definition that

$$\begin{aligned} 2\big\|f \star_r^r g\big\|^2_{\mathfrak{H}^{\otimes 2\ell-2r}} &= 2\langle f \star_{\ell-r}^{\ell-r} f, g \star_{\ell-r}^{\ell-r} g \rangle_{\mathfrak{H}^{\otimes 2r}} \\ &\leq \big\|f \star_{\ell-r}^{\ell-r} f\big\|^2_{\mathfrak{H}^{\otimes 2r}} + \big\|g \star_{\ell-r}^{\ell-r} g\big\|^2_{\mathfrak{H}^{\otimes 2r}}. \end{aligned}$$

□

1.3 Stein's method of normal approximation

A basic reference for Stein's method is the monograph [4]. Let us start with a fundamental fact that a real integrable random variable Z is a standard Gaussian random variable if and only if $\mathbb{E}[f'(Z)] = \mathbb{E}[Z \cdot f(Z)]$ for every bounded differentiable function $f : \mathbb{R} \longrightarrow \mathbb{R}$.

Now suppose that $Z \sim \mathcal{N}(0,1)$, for $h : \mathbb{R} \longrightarrow \mathbb{R}$ measurable such that $\mathbb{E}\big|h(Z)\big| < +\infty$, the differential equation $f'(x) - xf(x) = h(x) - \mathbb{E}[h(Z)]$ with unknown f is called the *Stein's equation* associated with h. We call f its solution, if f is absolutely continuous and one version of f' satisfies the Stein's equation everywhere. More precisely, we take $f'(x) = xf(x) + h(x) - \mathbb{E}[h(Z)]$ for every $x \in \mathbb{R}$.

It is well known (see *e.g.* [4, Chapter 2], [13, Chapter 3]) that given such a function h, there exists a unique solution f_h to the Stein's equation such that $\lim_{|x|\to+\infty} f_h(x)e^{-x^2/2} = 0$. Given a suitable *separating* class $\mathscr{F}$ of nice functions, we define

$$d_{\mathscr{F}}(X, Z) := \sup_{f\in\mathscr{F}} \left|\mathbb{E}[f(X)] - \mathbb{E}[f(Z)]\right| .$$

When $\mathscr{F}$ is set of 1-Lipschitz functions, $d_{\mathscr{F}}$ is the Wasserstein distance; when $\mathscr{F}$ is set of 1-Lipschitz functions that are also uniformly bounded by 1, $d_{\mathscr{F}}$ is called the Fortet-Mourier distance; when $\mathscr{F}$ is the collection of indicator functions $\mathbf{I}_{(-\infty,z]}$, $z \in \mathbb{R}$, $d_{\mathscr{F}}$ corresponds to the Kolmogorov distance appearing in the Berry-Esséen bound. We denote by $d_W, d_{\mathrm{FM}}, d_K$ respectively these distances. It is trivial that $d_{FM} \le d_W$, and it is not difficult to show that $d_K(X, Y) \le \sqrt{2C \cdot d_W(X, Y)}$ if X has density function uniformly bounded by C.

Now we replace the dummy variable x in the Stein's equation by a generic random variable X, then taking expectation on both sides of the equation gives $\mathbb{E}[Xf_h(X) - f_h'(X)] = \mathbb{E}[h(X) - h(Z)]$.

Here we collect several bounds for the Stein's solution f_h:

- For $h : \mathbb{R} \longrightarrow \mathbb{R}$ 1-Lipschitz, f_h is of class C^1 and f_h' is bounded Lipschitz with $\|f_h'\|_\infty \le \sqrt{2/\pi}$, $\|f_h''\|_\infty \le 2$, see *e.g.* [3, Lemma 4.2]. We denote by $\mathscr{F}_W$ the family of differentiable functions ϕ satisfying $\|\phi'\|_\infty \le \sqrt{2/\pi}$, $\|\phi''\|_\infty \le 2$, therefore for any square-integrable random variable F,

$$d_{\mathrm{FM}}(F, Z) \le d_W(F, Z) \le \sup_{\phi\in\mathscr{F}_W} \left|\mathbb{E}[F\phi(F) - \phi'(F)]\right| . \tag{1.15}$$

- If $h = \mathbf{I}_{(-\infty,z]}$ for some $z \in \mathbb{R}$, then $0 < f_h \le \frac{\sqrt{2\pi}}{4}$ and $\|f_h'\|_\infty \le 1$, see [4, Lemma 2.3]. We write $\mathscr{F}_K := \{\phi : \|\phi'\|_\infty \le 1, \|\phi\|_\infty \le \frac{\sqrt{2\pi}}{4}\}$, therefore for any integrable random variable F,

$$d_K(F, Z) \le \sup_{\phi\in\mathscr{F}_K} \left|\mathbb{E}[F\phi(F) - \phi'(F)]\right| . \tag{1.16}$$

  As the density of Z is uniformly bounded by $1/\sqrt{2\pi}$, we have the easy bound $d_K(F, N) \le \sqrt{d_W(F, N)}$.

2 Main results

2.1 Normal approximation in Wasserstein distance

In this subsection, we derive the Wasserstein distance bound for normal approximation of Rademacher functionals. Our new chain rule plays a crucial role.

Theorem 2.1. Given $Z \sim \mathscr{N}(0,1)$ and $F \in \mathbb{D}$ centred, one has in the setting (NS) that

$$d_W(F,Z) \le \sqrt{\frac{2}{\pi}} \cdot \mathbb{E}\Big[\big|1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\big|\Big] + \mathbb{E}\left[\Big\langle \frac{1}{\sqrt{pq}}|DL^{-1}F|, |DF|^2\Big\rangle_{\mathfrak{H}}\right]. \tag{2.1}$$

In particular, if $F \in \mathscr{C}_m$ for some $m \in \mathbb{N}$, then

$$\mathbb{E}\Big[\big|1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\big|\Big] \le \big|1 - \mathbb{E}[F^2]\big| + \frac{1}{m}\sqrt{\mathrm{Var}(\|DF\|^2_{\mathfrak{H}})} \tag{2.2}$$

and

$$\mathbb{E}\left[\Big\langle \frac{1}{\sqrt{pq}}|DL^{-1}F|, |DF|^2\Big\rangle_{\mathfrak{H}}\right] \le \sqrt{\mathbb{E}[F^2]/m} \cdot \sqrt{\sum_{k\in\mathbb{N}} \frac{1}{p_k q_k}\mathbb{E}[|D_kF|^4]}\,. \tag{2.3}$$

Proof. Given $\phi \in \mathscr{F}_W$, it follows from 1.1 and 1.2 that

$$\mathbb{E}[F\phi(F)] = \mathbb{E}[\langle D\phi(F), -DL^{-1}F\rangle_{\mathfrak{H}}] = \mathbb{E}[\phi'(F)\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}} + \langle R, -DL^{-1}F\rangle_{\mathfrak{H}}],$$

where $R = (R_k, k \in \mathbb{N})$ is the remainder satisfying $|R_k| \le |D_kF|^2/\sqrt{p_kq_k}$. Thus,

$$\mathbb{E}[F\phi(F)] - \mathbb{E}[\phi'(F)] = \mathbb{E}\Big[\phi'(F)\Big(\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}} - 1\Big)\Big] + \mathbb{E}[\langle R, -DL^{-1}F\rangle_{\mathfrak{H}}]$$

implying that

$$\Big|\mathbb{E}[F\phi(F) - \phi'(F)]\Big| \le \sqrt{\frac{2}{\pi}} \cdot \mathbb{E}\Big[\big|1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\big|\Big] + \mathbb{E}\left[\Big\langle \frac{1}{\sqrt{pq}}|DL^{-1}F|, |DF|^2\Big\rangle_{\mathfrak{H}}\right].$$

Hence (2.1) follows from (1.15).

If $F = J_m(f)$ with $m \in \mathbb{N}$, $f \in \mathfrak{H}_0^{\odot m}$, then $D_kF = mJ_{m-1}[f(\cdot,k)]$, $D_kL^{-1}F = -J_{m-1}[f(\cdot,k)]$. Recall $\mathbb{E}[F^2] = \mathbb{E}[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}]$, thus (2.2) follows easily from triangle inequality and Cauchy-Schwarz inequality:

$$\begin{aligned}
\mathbb{E}\Big[\big|1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\big|\Big] &\le \big|1 - \mathbb{E}[F^2]\big| + \mathbb{E}\Big[\big|\mathbb{E}[F^2] - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\big|\Big] \\
&\le \big|1 - \mathbb{E}[F^2]\big| + \sqrt{\mathrm{Var}(\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})} \\
&= \big|1 - \mathbb{E}[F^2]\big| + \frac{1}{m}\sqrt{\mathrm{Var}(\|DF\|^2_{\mathfrak{H}})}\,.
\end{aligned}$$

The inequality (2.3) is also an easy consequence of Cauchy-Schwarz inequality:

$$\begin{aligned}
\mathbb{E}\left[\Big\langle \frac{1}{\sqrt{pq}}|DL^{-1}F|, |DF|^2\Big\rangle_{\mathfrak{H}}\right] &= \frac{1}{m}\sum_{k\in\mathbb{N}} \mathbb{E}\left[|D_kF| \cdot \frac{1}{\sqrt{p_kq_k}}|D_kF|^2\right] \\
&\le \frac{1}{m}\sqrt{\sum_{k\in\mathbb{N}} \mathbb{E}[|D_kF|^2]} \cdot \sqrt{\sum_{k\in\mathbb{N}} \frac{1}{p_kq_k}\mathbb{E}[|D_kF|^4]} \\
&= \sqrt{\frac{\mathbb{E}[F^2]}{m}} \cdot \sqrt{\sum_{k\in\mathbb{N}} \frac{1}{p_kq_k}\mathbb{E}[|D_kF|^4]}\,.
\end{aligned}$$

□

Remark 2.1. (1) In the setting (S), the results in 2.1 can be easily deduced by taking $p_k = q_k = 1/2$ for each $k \in \mathbb{N}$. As we have mentioned earlier, our approximate chain rule is neater than (1.13) given in [14, Proposition 2.14], since it requires less regularity (this is the key point for us to get the estimate in Wasserstein distance). Although the authors of [14] were able to derive the Wasserstein distance via some smoothing argument, they imposed some further assumption and their rate of convergence is suboptimal compared to ours.

(2) In the setting (NS), Privault and Torrisi used their approximate chain rule (1.14) and the smoothing argument to obtain the Fortet-Mourier distance, see [18, Section 3.3]. It is suboptimal compared to our estimate in Wasserstein distance, in view of the trivial relation $d_{\mathrm{FM}} \leq d_W$.

(3) Recall that the test function $\phi \in \mathscr{F}_K$ (see (1.16)) may not have Lipschitz derivative, so our approximate chain rule as well as those in [14, 18] does not work to achieve the bound in Kolmogorov distance. Instead of using the chain rule, the authors of [8] carefully used a representation of the discrete Malliavin derivative $D\phi(F)$ and the fundamental theorem of calculus, this turns out to be flexible enough for them to deduce the Berry-Esséen bound in the setting (S). Later they obtained the Berry-Esséen bound in the setting (NS) with applications to random graphs. One can easily see that two terms in (1.3) are almost the same as our bound in Wasserstein distance while there are two extra terms (1.4) in their Kolmogorov distance bound.

Due to a comparison between (1.2) and (2.1), we are able to replace the Kolmogorov distance in many statements in [8, 9] by the Wasserstein distacne (with fewer terms and slightly different multiplicative constants). For example, we will obtain the so-called second-order Poincaré inequality in Wasserstein distance in the following

Remark 2.2. (**Second-order Poincaré inequality**) One can apply the Poincaré's inequality to $\langle -DL^{-1}F, DF\rangle_{\mathfrak{H}}$:

$$\mathrm{Var}(\langle -DL^{-1}F, DF\rangle_{\mathfrak{H}}) \leq \mathbb{E}\Big[\big\|D\langle -DL^{-1}F, DF\rangle_{\mathfrak{H}}\big\|_{\mathfrak{H}}^2\Big], \tag{2.4}$$

provided $\langle -DL^{-1}F, DF\rangle_{\mathfrak{H}} \in \mathbb{D}$.

In [9], Krokowski *et al.* gave the bound on Kolmogorov distance, see (1.2); they also established the so-called second-order Poincaré inequality as follows. For $Z \sim \mathscr{N}(0,1)$ and $F \in \mathbb{D}$ centred with unit variance, and $r, s, t \in (1, \infty)$ such that $r^{-1} + s^{-1} + t^{-1} = 1$, it holds that

$$d_K(F, Z) \leq A_1 + A_2 + \frac{\sqrt{2\pi}}{8} \cdot A_3 + A_4 + A_5 + A_6 + A_7\ , \tag{2.5}$$

where

$$A_1 := \left(\frac{15}{4} \sum_{j,k,\ell=1}^{\infty} \sqrt{\mathbb{E}[(D_jF)^2(D_kF)^2]}\,\sqrt{\mathbb{E}[(D_\ell D_jF)^2(D_\ell D_kF)^2]} \right)^{1/2};$$

$$A_2 := \left(\frac{3}{4} \sum_{j,k,\ell=1}^{\infty} \frac{1}{p_\ell q_\ell}\mathbb{E}\Big[\,(D_\ell D_jF)^2(D_\ell D_kF)^2\,\Big] \right)^{1/2};$$

$$A_3 := \sum_{k=1}^{\infty} \frac{1}{\sqrt{p_kq_k}}\mathbb{E}[|D_kF|^3]\ ; \quad A_4 := \frac{1}{2}\big\|F\big\|_{L^r(\Omega)} \sum_{k=1}^{\infty} \frac{1}{\sqrt{p_kq_k}}\big\|(D_kF)^2\big\|_{L^s(\Omega)}\big\|D_kF\big\|_{L^t(\Omega)};$$

$$A_5 := \left(\sum_{k\in\mathbb{N}} \frac{1}{p_k q_k} \mathbb{E}\Big[(D_k F)^4\Big] \right)^{1/2} \quad ; \quad A_6 = \left(3 \sum_{k,\ell=1}^{\infty} \frac{1}{p_k q_k p_\ell q_\ell} \mathbb{E}\Big[(D_\ell D_k F)^4\Big] \right)^{1/2} ;$$

$$A_7 := \left(6 \sum_{k,\ell=1}^{\infty} \frac{1}{p_k q_k} \sqrt{\mathbb{E}[(D_k F)^4]} \sqrt{\mathbb{E}[(D_\ell D_k F)^4]} \right)^{1/2} .$$

Part of the proof for (2.5) requires fine analysis of the discrete gradients in (2.4). For more details, see [9, Theorem 4.1]. Following exactly the same lines, one can obtain the following second-order Poincaré inequality in Wasserstein distance by simply comparing (1.2) and (2.1), as well as going through the proof of [9, Theorem 4.1]: $d_W(F,Z) \leq \sqrt{2/\pi} \cdot (A_1 + A_2) + A_3$. Note the constants r, s, t are not involved in A_1, A_2, A_3.

In the end of this subsection, we recall from [14] sufficient conditions for CLT (in the setting (S)) inside a fixed Rademacher chaos $\mathscr{C}_\ell$ ($\ell \geq 2$). The analogous result in the setting (NS) was proved in [18, Section 5.3].

Proposition 2.1. ([14, Theorem 4.1]) In the setting (S), fix $\ell \geq 2$. If $F_n := J_\ell(f_n)$ for some $f_n \in \mathfrak{H}_0^{\odot\ell}$, then

$$\text{(2.6)} \qquad \operatorname{Var}(\|DF_n\|^2_{\mathfrak{H}}) \leq C \sum_{m=1}^{\ell-1} \left\| f_n \star_m^m f_n \right\|^2_{\mathfrak{H}^{\otimes 2\ell-2m}} ,$$

and

$$\text{(2.7)} \qquad \sum_{k\in\mathbb{N}} \mathbb{E}[\left|D_k F_n\right|^4] \leq C \sum_{m=1}^{\ell-1} \left\| f_n \star_m^m f_n \right\|^2_{\mathfrak{H}^{\otimes 2\ell-2m}},$$

where the constant C only depends on ℓ.

As a consequence, the following result is straightforward.

Proposition 2.2. ([14, Proposition 4.3]) If $\|f_n\|^2_{\mathfrak{H}^{\otimes\ell}} \cdot \ell! \longrightarrow 1$ and

$$\text{(2.8)} \qquad \left\| f_n \star_m^m f_n \right\|_{\mathfrak{H}^{\otimes(2\ell-2m)}} \longrightarrow 0 \ , \quad \forall m = 1, 2, \cdots, \ell - 1 \ ,$$

as $n \to +\infty$, then F_n converges in law to a standard Gaussian random variable.

2.2 Almost sure central limit theorem for Rademacher chaos

The following lemma is crucial for us to apply the Ibragimov-Lifshits criterion. The Gaussian analogue was proved in [1, Lemma 2.2] and the Poisson case was given in [20, Proposition 5.2.5].

Lemma 2.1. In the setting (NS), if $F \in \mathbb{D}$ is centred such that $\langle -DL^{-1}F, DF \rangle_{\mathfrak{H}} \in L^2(\Omega)$ and

$$\sum_{k\in\mathbb{N}} \frac{1}{p_k q_k} \mathbb{E}[|D_k F|^4] < +\infty \ ,$$

then

$$
\left|\mathbb{E}[e^{itF}] - e^{-t^2/2}\right| \le |t|^2 \cdot \left|1 - \mathbb{E}[F^2]\right| + t^2 \cdot \sqrt{\mathrm{Var}(\langle -DL^{-1}F, DF\rangle_{\mathfrak{H}})} \\
+ |t|^3 \cdot \sum_{k\in\mathbb{N}} \frac{1}{\sqrt{p_k q_k}} \mathbb{E}\Big[|D_k L^{-1} F| \cdot (D_k F)^2\Big] . \tag{2.9}
$$

In particular, if $F = J_\ell(f)$ for some $\ell \in \mathbb{N}$ and $f \in \mathfrak{H}_0^{\odot \ell}$, then

$$
\left|\mathbb{E}[e^{itF}] - e^{-t^2/2}\right| \le |t|^2 \cdot \left|1 - \ell! \|f\|^2_{\mathfrak{H}^{\otimes \ell}}\right| + \frac{|t|^2}{\ell} \sqrt{\mathrm{Var}(\|DF\|^2_{\mathfrak{H}})} \\
+ |t|^3 \cdot \sqrt{\sum_{k\in\mathbb{N}} \frac{1}{p_k q_k} \mathbb{E}[|D_k F|^4]} \sqrt{\mathbb{E}[F^2]/\ell} . \tag{2.10}
$$

Proof. Set $\phi(t) = e^{t^2/2} \cdot \mathbb{E}[e^{itF}]$, $t \in \mathbb{R}$. Then

$$
\left|\mathbb{E}[e^{itF}] - e^{-t^2/2}\right| = |\phi(t) - \phi(0)| \cdot e^{-t^2/2} \le e^{-t^2/2} \cdot |t| \cdot \sup_{|s|\le|t|} \left|\phi'(s)\right| . \tag{2.11}
$$

Clearly,

$$
\phi'(t) = te^{t^2/2}\mathbb{E}[e^{itF}] + ie^{t^2/2}\mathbb{E}[F \cdot e^{itF}] = te^{t^2/2}\mathbb{E}[e^{itF}] + ie^{t^2/2}\mathbb{E}[\langle -DL^{-1}F, De^{itF}\rangle_{\mathfrak{H}}] ,
$$

and it follows from 1.2 that for each $k \in \mathbb{N}$,

$$
D_k e^{itF} = ite^{itF} D_k F + R_k
$$

with $|R_k| \le |t|^2 \cdot |D_k F|^2 / \sqrt{p_k q_k}$. Therefore,

$$
\phi'(t) = te^{t^2/2}\mathbb{E}[e^{itF}] - te^{t^2/2}\mathbb{E}\Big[e^{itF} \cdot \langle -DL^{-1}F, DF\rangle_{\mathfrak{H}}\Big] + ie^{t^2/2}\mathbb{E}[\langle -DL^{-1}F, R\rangle_{\mathfrak{H}}].
$$

Then by triangle inequality, one has

$$
\left|\phi'(t)\right| \le |t| e^{t^2/2} \cdot \mathbb{E}[|1 - \langle -DL^{-1}F, DF\rangle_{\mathfrak{H}}|] + e^{t^2/2}\Big|\mathbb{E}[\langle -DL^{-1}F, R\rangle_{\mathfrak{H}}]\Big|.
$$

It follows from (2.11) that

$$
\left|\mathbb{E}[e^{itF}] - e^{-t^2/2}\right| \le t^2 \mathbb{E}[|1 - \langle -DL^{-1}F, DF\rangle_{\mathfrak{H}}|] + |t|^3 \cdot \sum_{k\in\mathbb{N}} \frac{1}{\sqrt{p_k q_k}} \mathbb{E}\Big[|D_k L^{-1} F| \cdot (D_k F)^2\Big],
$$

then the desired inequality (2.9) follows from the estimate $\mathbb{E}[|1 - \langle -DL^{-1}F, DF\rangle_{\mathfrak{H}}|] \le \left|1 - \mathbb{E}[F^2]\right| + \sqrt{\mathrm{Var}(\langle -DL^{-1}F, DF\rangle_{\mathfrak{H}})}$, see 2.1. The rest is straightforward. □

The following theorem provides sufficient conditions for the ASCLT on a fixed Rademacher chaos in the setting (S). The analogous results in the Gaussian and Poisson settings can be found in [1, 20] respectively.

Theorem 2.2. In the setting (S), fix $\ell \geq 2$ and $F_n = J_\ell(f_n)$ with $f_n \in \mathfrak{H}_0^{\odot \ell}$ for each $n \in \mathbb{N}$. Assume that $\|f_n\|^2_{\mathfrak{H}^{\otimes \ell}} \cdot \ell ! = 1$, and the following two conditions as well as (2.8) are satisfied:

$$\textbf{C-1} \qquad \sum_{n\geq 2} \frac{1}{n\gamma_n^3} \sum_{k,j=1}^n \frac{\left|\langle f_k, f_j\rangle_{\mathfrak{H}^{\otimes \ell}}\right|}{kj} < +\infty \ ;$$

$$\textbf{C-2} \qquad \sum_{n\geq 2} \frac{1}{n\gamma_n^2} \sum_{k=1}^n \frac{1}{k} \cdot \left\| f_k \star_m^m f_k \right\|_{\mathfrak{H}^{\otimes 2\ell - 2m}} < +\infty \ , \quad \forall m = 1, \dots, \ell - 1 \, .$$

Then the ASCLT holds for $(F_n)_{n\in\mathbb{N}}$.

Proof. Note first F_n converges in law to a standard Gaussian random variable, by Proposition 2.2. Observe that

$$\left| \triangle_n(t) \right|^2 = \frac{1}{\gamma_n^2} \sum_{k,j=1}^n \frac{1}{kj}\Big(e^{it(F_k - F_j)} - e^{-t^2} \Big) - \frac{e^{-t^2/2}}{\gamma_n} \sum_{k=1}^n \frac{1}{k} \Big(e^{itF_k} - e^{-t^2/2}\Big) - \frac{e^{-t^2/2}}{\gamma_n} \sum_{j=1}^n \frac{1}{j}\Big(e^{-itF_j} - e^{-t^2/2}\Big). \tag{2.12}$$

Now we fix $r > 0$, $t \in [-r, r]$. For brevity, we omit the subscripts. One can deduce from 2.1 and (2.6), (2.7) that

$$\left| \mathbb{E}[e^{itF_k}] - e^{-t^2/2} \right| \leq C \cdot \sqrt{ \sum_{m=1}^{\ell-1} \left\| f_k \star_m^m f_k \right\|^2 },$$

here and in the following the constant C may vary from line to line but only depend on r, ℓ.

Since $\sqrt{a_1 + \ldots + a_l} \leq \sqrt{a_1} + \ldots + \sqrt{a_l}$ for any $a_1, \ldots, a_l \geq 0$,

$$\left| \mathbb{E}[e^{itF_k}] - e^{-t^2/2} \right| \leq C \cdot \sum_{m=1}^{\ell-1} \left\| f_k \star_m^m f_k \right\|.$$

Similarly, we apply the same argument with $s = \sqrt{2}t$ and $g = (f_k - f_j)/\sqrt{2}$, and we get

$$\left| \mathbb{E}\Big[e^{it(F_k - F_j)} \Big] - e^{-t^2} \right| = \left| \mathbb{E}[e^{is \cdot J_\ell(g)}] - e^{-s^2/2} \right| \leq C \sum_{m=1}^{\ell-1} \left\| g \star_m^m g \right\| + C \cdot \left| \langle f_k, f_j \rangle \right|. \tag{2.13}$$

Clearly, $g \star_m^m g = \frac{1}{2}(f_k \star_m^m f_k + f_j \star_m^m f_j - f_k \star_m^m f_j - f_j \star_m^m f_k)$, then

$$2\left\| g \star_m^m g \right\| \leq \left\| f_k \star_m^m f_k \right\| + \left\| f_j \star_m^m f_j \right\| + 2\left\| f_k \star_m^m f_j \right\| \ .$$

1.3 implies $\left\| f_k \star_m^m f_j \right\| \leq \left\| f_k \star_{\ell-m}^{\ell-m} f_k \right\| + \left\| f_j \star_{\ell-m}^{\ell-m} f_j \right\|$. Therefore,

$$\left| \mathbb{E}\Big[e^{it(F_k - F_j)} \Big] - e^{-t^2} \right| \leq C \left| \langle f_k, f_j \rangle \right| + C \sum_{m=1}^{\ell-1} \left(\left\| f_k \star_m^m f_k \right\| + \left\| f_j \star_m^m f_j \right\| \right). \tag{2.14}$$

Hence,

(2.15)

$$
\begin{aligned}
&\mathbb{E}(|\triangle_n(t)|^2)\\
&\le \frac{C}{\gamma_n^2}\sum_{k,j=1}^{n}\frac{\left|\langle f_k, f_j\rangle\right|}{kj} + \frac{C}{\gamma_n^2}\sum_{k,j=1}^{n}\frac{1}{kj}\sum_{m=1}^{\ell-1}\left(\left\|f_k \star_m^m f_k\right\| + \left\|f_j \star_m^m f_j\right\|\right) + \frac{C}{\gamma_n}\sum_{k=1}^{n}\frac{1}{k}\cdot\sum_{m=1}^{\ell-1}\left\|f_k \star_m^m f_k\right\|\\
&= \frac{C}{\gamma_n^2}\sum_{k,j=1}^{n}\frac{\left|\langle f_k, f_j\rangle\right|}{kj} + \frac{2C}{\gamma_n}\sum_{k=1}^{n}\frac{1}{k}\cdot\sum_{m=1}^{\ell-1}\left\|f_k \star_m^m f_k\right\| .
\end{aligned}
$$

Now we can see that the conditions **C-1, C-2** imply the Ibragimov-Lifshits condition (1.6), so the ASCLT holds for $(F_n, n \in \mathbb{N})$. □

In the setting (S), the normalised partial sum $S_n = (Y_1 + \cdots + Y_n)/\sqrt{n}$ converges in law to a standard Gaussian random variable. Moreover, the ASCLT holds for (S_n), this is a particular case of [10, Theorem 2]. The following result is a slight generalisation of this classic example.

Corollary 2.1. In the setting (NS), let $F_n = J_1(f_n)$ be such that $\|f_n\|_{\mathfrak{H}} = 1$ for all $n \in \mathbb{N}$. Assume that the following conditions hold:

$$
\begin{aligned}
&\text{(i)} \qquad \sum_{n\ge 2}\frac{1}{n\gamma_n^3}\sum_{k,j=1}^{n}\frac{\left|\langle f_k, f_j\rangle_{\mathfrak{H}}\right|}{kj} < +\infty\\
&\text{(ii)} \qquad \sum_{n\ge 2}\frac{1}{n\gamma_n^2}\sum_{k=1}^{n}\frac{1}{k}\cdot\sum_{m=1}^{\infty}\frac{1}{\sqrt{p_m q_m}}|f_k(m)|^3 < +\infty\\
&\text{(iii)} \qquad \sum_{m=1}^{\infty}\frac{1}{\sqrt{p_m q_m}}|f_k(m)|^3 \xrightarrow{k\to+\infty} 0 .
\end{aligned}
$$

Then $F_n \xrightarrow{\text{law}} \mathscr{N}(0,1)$ and the ASCLT holds for $(F_n, n \in \mathbb{N})$.

Proof. Note $-DL^{-1}F_n = f_n$ and $DF = f_n$. Then the quantity $\langle -DL^{-1}F_n, DF_n\rangle_{\mathfrak{H}}$ is deterministic and $|D_m L^{-1}F_n| \cdot (D_m F_n)^2 = |f_n(m)|^3$ for each $m \in \mathbb{N}$. Therefore, it follows from 2.1 that

$$
\left|\mathbb{E}[e^{itF_n}] - e^{-t^2/2}\right| \le |t|^3 \sum_{m=1}^{\infty}\frac{1}{\sqrt{p_m q_m}}|f_n(m)|^3 .
$$

By (iii), the CLT holds for (F_n). Similarly as in the proof of 2.2, we have

$$
\begin{aligned}
\left|\mathbb{E}\left[e^{it(F_k - F_j)}\right] - e^{-t^2}\right| &\le 2t^2\left|\langle f_k, f_j\rangle_{\mathfrak{H}}\right| + 2\sqrt{2}|t|^3\sum_{m=1}^{\infty}\frac{1}{\sqrt{p_m q_m}}|f_k(m) - f_j(m)|^3\\
&\le 2t^2\left|\langle f_k, f_j\rangle_{\mathfrak{H}}\right| + 8\sqrt{2}|t|^3\sum_{m=1}^{\infty}\frac{1}{\sqrt{p_m q_m}}(|f_k(m)|^3 + |f_j(m)|^3) ,
\end{aligned}
$$

where the last inequality follows from the elementary inequality $(a+b)^3 \le 4a^3 + 4b^3$ for any $a, b \ge 0$. The rest of the proof goes along the same lines as in 2.2. □

To conclude this section, we give the following example as an application of 2.2.

Example 2.1. In the setting (S), we consider the symmetric kernels $f_n \in \mathfrak{H}_0^{\odot 2}$ for $n \geq 1$:

$$f_n(i,j) = \begin{cases} \dfrac{1}{2\sqrt{n}} & \text{if } i,j \in \{1,2,\cdots,2n\} \text{ and } |i-j| = n\ ; \\ 0 & \text{otherwise.} \end{cases}$$

Setting $F_n = J_2(f_n)$, we claim that the ASCLT holds for $(F_n, n \geq 1)$.

Proof. It is easy to get $2\|f_n\|^2_{\mathfrak{H}^{\otimes 2}} = 1$ and

$$f_n \star_1^1 f_n(i,j) = \begin{cases} \dfrac{1}{4n} & \text{if } i = j \in \{1,2,\cdots,2n\}; \\ 0 & \text{otherwise.} \end{cases}$$

So $\|f_n \star_1^1 f_n\|_{\mathfrak{H}^{\otimes 2}} = \dfrac{1}{2\sqrt{2n}}$ converges to zero as $n \to +\infty$, thus the CLT follows from Proposition 2.2. If $k < \ell$, then

$$\langle f_k, f_\ell \rangle_{\mathfrak{H}^{\otimes 2}} = \sum_{i,j=1}^{2k} f_k(i,j) f_\ell(i,j) = \sum_{i,j=1}^{2k} f_k(i,j) \mathbf{I}_{(|i-j|=k)} f_\ell(i,j) \mathbf{I}_{(|i-j|=\ell)} = 0\ ,$$

thus

$$\sum_{n\geq 2} \frac{1}{n\gamma_n^3} \sum_{k,\ell=1}^{n} \frac{\left|\langle f_k, f_\ell \rangle_{\mathfrak{H}^{\otimes 2}}\right|}{k\ell} = \sum_{n\geq 2} \frac{1}{n\gamma_n^3} \sum_{k=1}^{n} \frac{\left|\langle f_k, f_k \rangle_{\mathfrak{H}^{\otimes 2}}\right|}{k^2} = \sum_{n\geq 2} \frac{1}{n\gamma_n^3} \sum_{k=1}^{n} \frac{1}{2k^2} < +\infty\ .$$

That is, the condition (**C-1**) in 2.2 is satisfied. It remains to check the condition (**C-2**):

$$\sum_{n\geq 2} \frac{1}{n\gamma_n^2} \sum_{k=1}^{n} \frac{1}{k} \cdot \left\| f_k \star_1^1 f_k \right\|_{\mathfrak{H}^{\otimes 2}} = \sum_{n\geq 2} \frac{1}{n\gamma_n^2} \sum_{k=1}^{n} \frac{1}{k} \cdot \frac{1}{2\sqrt{2k}} < +\infty\ .$$

Hence it follows from 2.2 that the ASCLT holds for $(F_n, n \geq 1)$. □

References

[1] Bercu, B., Nourdin, I. and Taqqu, M.S. Almost sure central limit theorems on the Wiener space. *Stochastic Process. Appl.*, Vol. 120, Issue 9, pages 1607-1628.

[2] Brosamler, G.A. An almost everywhere central limit theorem. *Math. Proc. Cambridge Philos. Soc.* **104**(3)(1988) 561-574.

[3] Chatterjee, S. A new method of normal approximation. *Ann. Probab.* 2008, **36**(4),1584-1610.

[4] Chen, L.H.Y., Goldstein, L. and Shao Q.M. Normal approximation by Stein's method. *Probability and Its Applications*, 2011 Springer-Verlag Berlin Heidelberg.

[5] Eichelsbacher, P. and Thäle, Ch. New Berry-Esseen bounds for non-linear functionals of Poisson random measures. *Electron. J. Probab.* **19**(2014), no. 102, 1-25.

[6] Ibragimov, I.A. and Lifshits, M.A. On almost sure limit theorems. *Theory PROBAB. APPL.* Vol44, No. 2, (1999) translated by M.V. Khatuntseva from Russian journal *Teor. Veroyatnost. i Primenen*. **44**(2), 328-350(1999).

[7] Krokowski, K. Poisson approximation of Rademacher functionals by the Chen-Stein method and Malliavin Calculus. *Commun. Stoch. Anal.* 11 (2017), no. 2, 195-222

[8] Krokowski, K., Reichenbachs, A. and Thäle, Ch. Berry-Esseen bounds and multivariate limit theorems for functionals of Rademacher sequences. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2016) **no. 2**, 763-803.

[9] Krokowski, K., Reichenbachs, A. and Thäle, Ch. Discrete Malliavin-Stein method: Berry-Esseen bounds for random graphs, point processes and percolation. to appear in *Ann. Probab.*

[10] Lacey, M.T. and Phillip, W. A note on the almost sure central limit theorem. *Statist. Probab. Lett.* **9**(3), 201-205 (1990).

[11] Nourdin, I. https://sites.google.com/site/malliavinstein/home (a website maintained by Ivan Nourdin)

[12] Nourdin, I. and Peccati, G. Stein's method on Wiener chaos, *Probab. Theory Related Fields*, 2009, Vol. 145, Issue 1, p. 75-118.

[13] Nourdin, I. and Peccati, G. Normal approximations with Malliavin calculus: from Stein's method to universality, *Cambridge tracts in Mathematics*, Vol. 192, 2012, Cambridge University Press.

[14] Nourdin, I., Peccati, G. and Reinert, G. Stein's method and stochastic analysis of Rademacher functionals, *Electron. J. Probab.*, Vol. 15 (2010), Paper no. 55, pages 1703-1742.

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

[16] Peccati, G., Solé, J.L., Taqqu, M.S. and Utzet, F. Stein's method and normal approximation of Poisson functionals, *Ann. Probab.*, 2010, Vol. 38, No. 2, 443-478.

[17] Privault, N. Stochastic analysis of Bernoulli processes, *Probab. Surv.* Vol. 5 (2008) 435-483

[18] Privault, N. and Torrisi, G.L. The Stein and Chen-Stein methods for functionals of non-symmetric Bernoulli processes, *ALEA Lat. Am. J. Probab. Math. Stat.*, **12**(1), 309-356 (2015)

[19] Schatte, P. On Strong versions of the central limit theorem. *Math. Nachr.* **137**(1988) 249-256.

[20] Zheng, C. *Multi-dimensional "Malliavin-Stein" method on the Poisson space and its applications to limit theorems*, Ph.D dissertation, Université Pierre et Marie Curie, Paris VI, 2011.

This page is left blank

Paper 2: Convergence of random oscillatory integrals in the presence of long-range dependence and application to homogenization

A. Lechiheb, I. Nourdin, G. Zheng and E. Haouala

To appear in *Probab. Math. Stat.*

Abstract

This paper deals with the asymptotic behavior of random oscillatory integrals in the presence of long-range dependence. As a byproduct, we solve the corrector problem in random homogenization of one-dimensional elliptic equations with highly oscillatory random coefficients displaying long-range dependence, by proving convergence to stochastic integrals with respect to Hermite processes.

1 Main results of the paper

1.1 Convergence of random oscillatory integrals

In the present paper one of our goals is to study, once properly normalized, the distributional convergence of some random oscillatory integrals of the form

$$\int_0^1 \Phi[g(x/\varepsilon)]h(x)\,dx \ , \tag{1.1}$$

where

- $h \in C([0,1])$ is deterministic,
- $\{g(x)\}_{x\in\mathbb{R}_+}$ is a certain centred stationary Gaussian process exhibiting long-range correlation,
- $\Phi \in L^2(\mathbb{R},\nu)$ has Hermite rank $m \geq 1$ (with ν the standard Gaussian measure).

As we will see later, the main motivation of this study comes from the random corrector problem studied in [4].

Let us first introduce the Gaussian process $\{g(x)\}_{x\in\mathbb{R}_+}$ we will deal with throughout all this paper. It is constructed as follows:

1. Let $m \in \mathbb{N}^*$ be fixed, let $H_0 \in (1-\frac{1}{2m},1)$, and set $H = 1 + m(H_0-1) \in (1/2,1)$;

2010 *Mathematics Subject Classification*. Primary: 60F05, 80M40 ; Secondary: 60H05,60H20,60G10,60G18.
Key words and phrases. Elliptic equation; Hermite process; oscillatory integral; corrector; homogenization.

2. Fix a slowly varying function $L : (0, +\infty) \to (0, +\infty)$ at $+\infty$, that is, consider a measurable and locally bounded function L such that $L(\lambda x)/L(x) \to 1$ as $x \to +\infty$, for every $\lambda > 0$. Assume furthermore that L is bounded away from 0 and $+\infty$ on every compact subset of $(0, +\infty)$. (See [3] for more details on slowly varying functions.)
3. Let $e : \mathbb{R} \to \mathbb{R}$ be a square-integrable function such that
   (3a) $\int_{\mathbb{R}} e(u)^2\, du = 1$,
   (3b) $|e(u)| \le C u^{H_0 - \frac{3}{2}} L(u)$ for almost all $u > 0$, for some absolute constant C,
   (3c) $e(u) \sim C_0 u^{H_0 - \frac{3}{2}} L(u)$, where $C_0 = \left(\int_0^\infty (u + u^2)^{H_0 - \frac{3}{2}}\, du \right)^{-1/2}$,
   (3d) their exists $0 < \gamma < \min\{H_0 - (1 - \frac{1}{2m}), 1 - H_0\}$ such that
   $$\int_{-\infty}^{0} |e(u)e(xy + u)|\, du = o(x^{2H_0 - 2} L(x)^2) y^{2H_0 - 2 - 2\gamma}$$
   as $x \to \infty$, uniformly in $y \in (0, t]$ for each given $t > 0$.
4. Finally, let W be a two-sided Brownian motion.

Bearing all these ingredients in mind, we can now set, for $x \in \mathbb{R}_+$,

$$g(x) := \int_{-\infty}^{\infty} e(x - \xi) dW_\xi \ . \tag{1.2}$$

Remark 1.1. (i) Assumptions 3a and 4 ensure that $\{g(x)\}_{x \in \mathbb{R}_+}$ is a normalised centred Gaussian process.

(ii) Assumption 3b controls $|e(u)|$ for small u, while Assumption 3d ensures that the "forward" contribution of $e(u)$ is ultimately negligible due to the following computation:

$$\begin{aligned}
\mathbb{E}[g(s)g(s + x)] &= \int_{-\infty}^{\infty} e(s - \xi)e(s + x - \xi)\, d\xi = \int_{-\infty}^{\infty} e(u)e(u + x)\, du \\
&= \int_{-\infty}^{0} e(u)e(u + x)\, du + \int_{0}^{\infty} e(u)e(u + x)\, du \\
&= o(x^{2H_0 - 2} L(x)^2) + x \int_{0}^{\infty} e(xu)e(xu + x)\, du \ .
\end{aligned}$$

(iii) Assumption 3c ensures that the process $\{g(x)\}_{x \in \mathbb{R}_+}$ exhibits the following asymptotic behaviour:

$$R_g(x) := \mathbb{E}[g(s)g(s + x)] \sim x^{2H_0 - 2} L(x)^2 \quad \text{as } x \to +\infty, \tag{1.3}$$

see [12, Equation (2.3)].

In section 3.1, we will show that the random integral given by (1.1) exhibits the following asymptotic behavior as $\varepsilon \to 0$.

Theorem 1.1. Let g be the centred stationary Gaussian process defined by (1.2), and assume that $\Phi \in L^2(\mathbb{R}, \nu)$ has Hermite rank $m \geq 1$. Then, for any $h \in C([0,1])$, the following convergence in law takes place

$$M_h^\varepsilon := \frac{1}{\varepsilon d(1/\varepsilon)} \int_0^1 \Phi[g(x/\varepsilon)]h(x)\, dx \overset{\varepsilon\downarrow 0}{\longrightarrow} M_h^0 := \frac{V_m}{m!} \int_0^1 h(x)\, dZ(x)\,, \tag{1.4}$$

where Z is the mth-Hermite process defined by (2.4) and $d(\cdot)$ is defined by

$$d(x) = \sqrt{\frac{m!}{H(2H-1)}} x^H L(x)^m. \tag{1.5}$$

As we already anticipated, the fine analysis of the asymptotic behavior of (1.4) is motivated by the random corrector problem studied in [4], that we describe now.

1.2 A motivating example

Theorem 1.1 appears to be especially useful and relevant in the study of the following homogenization problem. Consider the following *one-dimensional* elliptic equation displaying random coefficients:

$$\begin{cases} -\dfrac{d}{dx}\left(a(x/\varepsilon, \omega)\dfrac{d}{dx}u^\varepsilon(x,\omega)\right) = f(x)\ , \quad x \in (0,1)\ , \quad \varepsilon > 0 \\ u^\varepsilon(0,\omega) = 0\ , \quad u^\varepsilon(1,\omega) = b \in \mathbb{R}. \end{cases} \tag{1.6}$$

In (1.6), the random potential $\{a(x)\}_{x\in\mathbb{R}_+}$ is assumed to be a uniformly bounded, positive[1] stationary stochastic process, whereas the data f is continuous. This model has received a lot of interests in the literature (see for instance [5, page 13-14]).

Taking strong advantage of the fact that the ambient dimension is one, it is immediate to check that the solution to (1.6) is given explicitly by

$$u^\varepsilon(x,\omega) = c^\varepsilon(\omega) \int_0^x \frac{1}{a(y/\varepsilon,\omega)}\, dy - \int_0^x \frac{F(y)}{a(y/\varepsilon,\omega)}\, dy, \tag{1.7}$$

where $F(x) := \int_0^x f(y)\, dy$ is the antiderivative of f vanishing at zero, and where

$$c^\varepsilon(\omega) := \left(b + \int_0^1 \frac{F(y)}{a(y/\varepsilon,\omega)}\, dy \right)\left(\int_0^1 \frac{1}{a(y/\varepsilon,\omega)}\, dy \right)^{-1}.$$

Under suitable ergodic and stationary assumptions on a, the ergodic theorem applied to (1.7) implies that u^ε converges pointwise to $\bar{u}$ as $\varepsilon \to 0$, where

$$\bar{u}(x) = \frac{c^* x}{a^*} - \int_0^x \frac{F(y)}{a^*}\, dy,$$

[1] That is, there exists $r \in (0,1)$ such that $r \leq a(x) \leq r^{-1}$ for every $(x,\omega) \in \mathbb{R}_+ \times \Omega$.

with $c^* := ba^* + \int_0^1 F(y)\,dy$ and

$$a^* := \frac{1}{\mathbb{E}[1/a(0)]}.$$

The above parameter a^* is usually refered to as the *effective diffusion coefficient* in the literature, see *e.g.* [10]. It is also immediately checked that $\bar{u}$ is the unique solution to the following *deterministic* equation:

$$\begin{cases} -\dfrac{d}{dx}\left(a^* \dfrac{d}{dx}\bar{u}(x)\right) = f(x)\ , \quad x \in (0,1) \\ \bar{u}(0) = 0\ , \quad \bar{u}(1) = b. \end{cases} \tag{1.8}$$

Interested readers can refer to [2] for a recent review on models involving more general elliptic equations.

In this work, we address the random corrector problem for (1.6) in presence of *long-range* media, that is, we analyze the behaviour of the random fluctuations between u^ε and $\bar{u}$ when the random potential a is obtained by means of a long-range process (see below for the details). Taking advantage of the explicit expressions for both (1.6) and (1.8), it is easy but crucial to observe that the random corrector $u^\varepsilon(x) - \bar{u}(x)$ can be fully expressed by means of random oscillatory integrals of the form

$$\int_0^1 \left[\frac{1}{a(y/\varepsilon)} - \frac{1}{a^*} \right] h(y)\,dy \tag{1.9}$$

for some function h. Thus, the random corrector problem for (1.6) reduces in a careful analysis of the asymptotic behaviour of random quantities of the form (1.9) as $\varepsilon \to 0$. To this aim, we need to give a precise description about the form of the process a.

Let ν denote the standard Gaussian measure on $\mathbb{R}$. Every $\Phi \in L^2(\mathbb{R},\nu)$ admits the following series expansion

$$\Phi = \sum_{q=0}^{\infty} \frac{V_q}{q!} H_q, \quad \text{with } V_q := \int_{\mathbb{R}} \Phi(x) H_q(x)\nu(dx), \tag{1.10}$$

and where $H_q(x) = (-1)^q \exp(x^2/2)\frac{d^q}{dx^q}\exp(-x^2/2)$ denotes the qth Hermite polynomial. Recall that the integer $m_\Phi := \inf\{q \geq 0 :\ V_q \neq 0\}$ is called the *Hermite rank* of Φ (with the convention $\inf \emptyset = +\infty$). For any integer $m \geq 1$, we define $\mathscr{G}_m$ to the collection of all square-integrable functions (with respect to the standard Gaussian measure on $\mathbb{R}$) that have Hermite rank m.

Using Theorem 1.1 as main ingredient, we will prove the following result about the asymptotic behaviour of the random corrector associated with (1.6).

Theorem 1.2. Fix an integer $m \geq 1$ as well as two real numbers $H_0 \in (1 - \frac{1}{2m}, 1)$ and $b \in \mathbb{R}$, and let $\{a(x)\}_{x\in\mathbb{R}_+}$ be a uniformly bounded, positive and stationary stochastic process. Assume in addition that $q = \{q(x)\}_{x\in\mathbb{R}_+}$ given by

$$q(x) = \frac{1}{a(x)} - \frac{1}{a^*}, \quad \text{where } a^* := 1/\mathbb{E}[1/a(0)], \tag{1.11}$$

has the form

$$q(x) = \Phi(g(x)), \tag{1.12}$$

where $\Phi \in L^2(\mathbb{R}, \nu)$ belongs to $\mathscr{G}_m$ and $\{g(x)\}_{x\in\mathbb{R}_+}$ is the Gaussian process given by (1.2). Finally, let $f : [0,1] \to \mathbb{R}$ be continuous, and let us consider the solutions u^ε and $\bar{u}$ of (1.6) and (1.8) respectively. Then, for each $\varepsilon > 0$, the random corrector $u^\varepsilon - \bar{u}$ is a continuous process on $[0,1]$. Moreover, we have the following convergence in law on $C([0,1])$ endowed with the supremum norm as $\varepsilon \to 0$:

$$\left\{\frac{u^\varepsilon(x) - \bar{u}(x)}{\varepsilon d(1/\varepsilon)}\right\}_{x\in[0,1]} \Longrightarrow \left\{\frac{V_m}{m!}\int_{\mathbb{R}} F(x,y)\, dZ(y)\right\}_{x\in[0,1]},$$

where d is given by (1.5),

$$F(x) = \int_0^x f(y)dy, \quad c^* = a^* b + \int_0^1 F(y)\, dy,$$
$$F(x,y) = [c^* - F(y)]\mathbf{I}_{[0,x]}(y) + x\Big(F(y) - \int_0^1 F(z)dz - a^* b\Big)\mathbf{I}_{[0,1]}(y),$$

and Z is the Hermite process of order m and self-similar index

$$H := 1 + m(H_0 - 1) \in (1/2, 1).$$

(The definition of Z is given in Theorem 2.1.)

Note that it is not difficult to construct a process a satisfying all the assumptions of Theorem 1.2. Indeed, bearing in mind the notation of Theorem 1.2, we can write

$$a(x) = \left(q(x) + \frac{1}{a^*}\right)^{-1} = \left(\Phi(g(x)) + \frac{1}{a^*}\right)^{-1}. \tag{1.13}$$

Firstly, we note that since g given by (1.2) is stationary, clearly the same holds for a, whatever the expression of Φ. Secondly, given any fixed $a^* > 0$, we can construct a bounded measurable function $\Phi \in \mathscr{G}_2$ with $\|\Phi\|_\infty \le 1/(2a^*)$:

> let h_1, h_2 be two bounded measurable functions, then it is clear that they belong to $L^2(\mathbb{R}, \nu)$ and they admit the series expansion
>
> $$h_1 - \int_{\mathbb{R}} h_1\, d\nu = \sum_{k=1}^\infty a_k H_k \qquad \text{and} \qquad h_2 - \int_{\mathbb{R}} h_2\, d\nu = \sum_{k=1}^\infty b_k H_k\,,$$
>
> where the coefficients a_k, b_k are defined in the obvious manner. Therefore, the function
>
> $$\Psi := b_1\left(h_1 - \int_{\mathbb{R}} h_1\, d\nu\right) - a_1\left(h_2 - \int_{\mathbb{R}} h_2\, d\nu\right)$$
>
> is bounded and belongs to $\mathscr{G}_2$.

Then we pick $\Phi = \dfrac{\Psi}{2a^*\|\Psi\|_\infty} \in \mathscr{G}_2$. Therefore $a(x)$ defined by (1.13) satisfies

$$0 < \frac{2a^*}{3} \leq a(x) \leq 2a^* \ . \tag{1.14}$$

Inductively, one can construct a bounded measurable Φ with Hermite rank $m \geq 3$ (by starting with two bounded functions in $\mathscr{G}_{m-1}$) such that the process $\{a(x), x \in \mathbb{R}\}$ given in (1.13) satisfies (1.14).

Yet another possibility of constructing such a process $\{a(x), x \in \mathbb{R}\}$ is stated (more explicitly) as follows: fix $0 < t_1 < \ldots < t_m$, and consider the unique $(m+1)$-uple $(b_0, \ldots, b_m)$ satisfying

$$\begin{cases} \sum_{l=0}^m b_l\, e^{-kt_l} = 0 \quad \text{for all } k \in \{0, \ldots, m-1\}, \\ \sum_{l=0}^m b_l\, e^{-mt_l} = 1 \ . \end{cases} \tag{1.15}$$

(Existence and uniqueness of a solution to (1.15) is a consequence of a Vandermonde determinant.) Now, consider any measurable function ψ satisfying

$$0 \leq \psi \leq \frac{1}{2a^* \sum_{l=0}^m |b_l|}. \tag{1.16}$$

Since ψ belongs obviously to $L^2(\mathbb{R}, \nu)$, it may be expanded in Hermite polynomials as $\psi = \sum_{k=0}^\infty a_k H_k$. We assume moreover that $a_m \neq 0$. (Existence of ψ satisfying both (1.16) and $a_m \neq 0$ is clear by a contradiction argument.) Now, let

$$\Phi = \sum_{l=0}^m b_l P_{t_l}\psi,$$

where $P_t\psi(x) = \int_{\mathbb{R}} \psi(e^{-t}x + \sqrt{1-e^{-2t}}y)\nu(dy)$ is the classical Ornstein-Uhlenbeck semigroup. Due to (1.15), it is readily checked that the expansion of Φ is

$$\Phi = a_m H_m + \sum_{k=m+1}^\infty \left\{ \sum_{l=0}^m b_l e^{-kt_l} \right\} a_k H_k,$$

so that $\Phi \in \mathscr{G}_m$. Moreover,

$$\|\Phi\|_\infty \leq \sum_{l=0}^m |b_l| \|P_{t_l}\psi\|_\infty \leq \|\psi\|_\infty \sum_{l=0}^m |b_l| \leq \frac{1}{2a^*}$$

and a given by (1.13) is positive and bounded. So, existence of a process a satisfying all the assumptions of Theorem 1.2 is shown.

Theorem 1.2 should be seen as an extension and unified approach of the main results of [4], and it contains these latter as particular cases. More precisely, the case where the Hermite rank of Φ is $m = 1$ corresponds to [1, Theorem 2.5] and involves the fractional Brownian motion in the limit, whereas the case where the Hermite rank of Φ is $m = 2$ corresponds to [4, Theorem 2.2] and involves the Rosenblatt process in the limit. Also, in their last section (entitled *Conclusions* *and further discussion*), the authors of [4] pointed out that "it is natural to ask what would happen if the Hermite rank of Φ was greater than 2". Our Theorem 1.2 answers this question, by showing (as was guessed by the authors of [4]) that, in the case $m \geq 3$, the limit takes the form of an integral with respect to the Hermite process of order m. Finally, we would like to emphasize that our Theorem 1.2, even in the cases $m = 1$ and $m = 2$, is a strict extension of the results of [4], as we allow the possibility to deal with a slowly varying function L. That being said, our proof of Theorem 1.2 is exclusively based on the ideas and results contained in the seminal paper [12] and follows the strategy developed in [4]. In higher dimension, it is usually very hard to study the corrector theory due to the lack of explicit form of the solution. In a recent work [8, 9], the authors considered the discretised version of the corrector problem in higher dimension and they were able to study the scaling limit to some Gaussian fields. For more details, we refer the interested readers to these two papers and the references therein.

1.3 Plan of the paper

The rest of the paper is organized as follows. In Section 2, we give some preliminary results, divided into several subsections. Section 3 contains the proof of Theorems 1.2 and 1.1.

2 Preliminary results

Throughout all this section, we let all the notation and assumptions of Sections 1.1 and 1.2 prevail.

2.1 Asymptotic behaviour of the covariance function of q

For $x \in \mathbb{R}$, set $R_q(x) = \mathbb{E}[q(0)q(x)]$. Also, recall that m is the Hermite rank of Φ. Then, proceeding in similar lines as that of [4, Lemma 2.1], one can show that

$$\left|R_q(x)\right| = (o(1) + V_m^2/m!)L(|x|)^{2m}|x|^{-2(1-H)}, \tag{2.1}$$

as $|x| \to +\infty$. Here $o(1)$ means that the term converges to zero when $x \to \infty$.

The asymptotic relation (2.1) implies the existence of some absolute constant C satisfying

$$\left|R_q(x)\right| \leq C\, L(|x|)^{2m}|x|^{-2(1-H)} \tag{2.2}$$

for any $x \neq 0$.

2.2 Taqqu's theorem and convergence to Hermite process Z

Recall $d(x)$ from (1.5). Its main property is that the variance of $\dfrac{1}{d(x)}\displaystyle\int_0^x H_m(g(y))\,dy$ is asymptotically equal to 1 as $x \to +\infty$.

The following result, due to Taqqu in 1979, is the key ingredient in our proofs.

Theorem 2.1. ([12, Lemma 5.3]) Assume $\Phi \in \mathscr{G}_m$ and let g be given by (1.2). Then, as $T \to +\infty$, the process

$$Y_T(x) = \frac{1}{d(T)} \int_0^{Tx} \Phi[g(y)]\, dy, \quad x \in \mathbb{R}_+, \tag{2.3}$$

converges to $\frac{V_m}{m!} Z(x)$ in the sense of finite-dimensional distributions, where the mth-order Hermite process Z with self-similar index $H = m(H_0 - 1) + 1$ is defined by:

$$\begin{aligned} Z(x) &\\ = K(m, H_0) &\left\{ \int_{-\infty}^{\infty} dB_{\xi_1} \int_{-\infty}^{\xi_1} dB_{\xi_2} \ldots \int_{-\infty}^{\xi_{m-1}} dB_{\xi_m} \int_0^x \prod_{i=1}^m (s - \xi_i)^{H_0 - \frac{3}{2}} \mathbf{1}_{(\xi_i < s)}\, ds \right\}, \end{aligned} \tag{2.4}$$

where

$$K(m, H_0) := \sqrt{\frac{m! H(2H-1)}{\left(\int_0^\infty (u + u^2)^{H_0 - \frac{3}{2}}\, du\right)^m}}$$

is the normalising constant such that $\mathbb{E}[Z(1)^2] = 1$. (See [12, Equation (1.6)])

Note that $Z(x)$ lives in the Wiener chaos of order m, which is non-Gaussian unless $m = 1$ or $x = 0$.

2.3 Wiener integral with respect to Z

Let Z be given as above and let $\mathcal{E}$ be the set of elementary (deterministic) functions, that is, the set of functions h of the form

$$h(x) = \sum_{k=1}^{\ell} a_k \mathbf{I}_{(t_k, t_{k+1}]}(x)$$

with $\ell \in \mathbb{N}^*$, $a_k \in \mathbb{R}$, $t_k < t_{k+1}$. For such h, we define the Wiener integral with respect to Z in the usual way, as a linear functional over $\mathcal{E}$:

$$\int_{\mathbb{R}} h(x)\, dZ(x) = \sum_{k=1}^{\ell} a_k \Big[Z(t_{k+1}) - Z(t_k) \Big] .$$

One can verify easily that this definition is independent of choices of representation for elementary functions. Now we introduce the space of (deterministic) integrands for this Wiener integral:

$$\Lambda^H = \left\{ f : \mathbb{R} \longrightarrow \mathbb{R} \,\Big|\, \int_{\mathbb{R}} \int_{\mathbb{R}} f(u) f(v) |u - v|^{2H-2}\, du\, dv < +\infty \right\}, \tag{2.5}$$

equipped with the norm

$$\|f\|_{\Lambda^H}^2 = H(2H-1) \int_{\mathbb{R}} \int_{\mathbb{R}} f(u) f(v) |u - v|^{2H-2}\, du\, dv\,. \tag{2.6}$$

When $h \in \mathcal{E}$, it is straightforward to check the following isometry property:

$$\mathbb{E}\left[\left(\int_{\mathbb{R}} h(x)dZ(x)\right)^2\right] = \|h\|^2_{\Lambda^H}.$$

As a consequence, one can define the Wiener integral $\int_{\mathbb{R}} f(x)dZ(x)$ for any $f \in \Lambda^H$, by a usual approximation procedure.

It is by now well known (thanks to [11]) that $(\Lambda^H, \|\cdot\|_{\Lambda^H})$ is a Hilbert space that contains distributions in the sense of Schwartz. To overcome this problem, we shall restrict ourselves to the proper subspace

$$|\Lambda^H| = \left\{ f : \mathbb{R} \to \mathbb{R} \middle| \quad \int_{\mathbb{R}} \int_{\mathbb{R}} |f(u)f(v)||u-v|^{2H-2}\, du\, dv < +\infty \right\}$$

equipped with the norm

$$\|f\|^2_{|\Lambda^H|} = H(2H-1) \int_{\mathbb{R}} \int_{\mathbb{R}} |f(u)f(v)||u-v|^{2H-2}\, du\, dv\,.$$

We then have (see [11, Proposition 4.2])

$$L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \subset L^{1/H}(\mathbb{R}) \subset |\Lambda^H| \subset \Lambda^H. \tag{2.7}$$

Moreover, $(|\Lambda^H|, \|\cdot\|_{|\Lambda^H|})$ is a Banach space, in which the set $\mathcal{E}$ is dense. So for $h \in |\Lambda^H|$, we can define

$$\int_{\mathbb{R}} h(x)\, dZ(x) = \lim_{n\to+\infty} \int_{\mathbb{R}} h_n(x)\, dZ(x)\,, \tag{2.8}$$

where (h_n) is any sequence of $\mathcal{E}$ converging to h in $(|\Lambda^H|, \|\cdot\|_{|\Lambda^H|})$; the convergence in (2.8) takes place in $L^2(\Omega)$.

For a detailed account of this integration theory, one can refer to [7, 11].

2.4 Some facts about slowly varying functions

Let $L : (0, +\infty) \to (0, +\infty)$ be a slowly varying function at $+\infty$ and $\alpha > 0$. It is well known (see [3, Proposition 1.3.6(v)]) that

$$x^{\alpha} L(x) \to +\infty \quad \text{and} \quad x^{-\alpha} L(x) \to 0\ ,$$

as $x \to +\infty$. In particular, one can deduce that

$$\lim_{\varepsilon\downarrow 0} \varepsilon^{1-H} L(1/\varepsilon)^m = 0\ . \tag{2.9}$$

The following result is known as Potter's Theorem (see [3, Theorem 1.5.6(ii)]).

Theorem 2.2. Let $L : (0, +\infty) \to (0, +\infty)$ be a slowly varying function at $+\infty$ such that it is bounded away from 0 and $+\infty$ on every compact subset of $(0, +\infty)$. Then for any $\delta > 0$, there exists some constant $C = C(\delta)$ such that

$$\frac{L(y)}{L(x)} \leq C \max\left\{ (x/y)^{\delta}\,,\ (y/x)^{\delta} \right\}$$

for any $x, y \in (0, +\infty)$.

3 Proofs of main results

3.1 Proof of Theorem 1.1

First recall that a typical function h in $\mathcal{E}$ has the form

$$h(x) = \sum_{\ell=1}^{n} a_\ell \mathbf{I}_{(t_\ell, t_{\ell+1}]}(x), \quad t_\ell < t_{\ell+1}, \quad a_\ell \in \mathbb{R}, \quad \ell = 1, ..., n\,.$$

For such a simple function h, we deduce from Taqqu's Theorem 2.1 that

$$\begin{aligned} M_h^\varepsilon &= \frac{1}{\varepsilon d(1/\varepsilon)} \int_{\mathbb{R}} q(x/\varepsilon) \sum_{\ell=1}^{n} a_\ell \mathbf{I}_{(t_\ell, t_{\ell+1}]}(x)\, dx \\ &= \sum_{\ell=1}^{n} a_\ell \frac{1}{d(1/\varepsilon)} \left(\int_0^{t_{\ell+1}/\varepsilon} \Phi(g(x))\, dx - \int_0^{t_\ell/\varepsilon} \Phi(g(x))\, dx \right) \\ &\xrightarrow{\varepsilon\to 0} \frac{V_m}{m!} \sum_{\ell=1}^{n} a_\ell [Z(t_{\ell+1}) - Z(t_\ell)] = \frac{V_m}{m!} \int_{\mathbb{R}} h(x)\, dZ(x)\,. \end{aligned}$$

This proves (1.4) for simple functions $h \in \mathcal{E}$.

Let us now consider $h \in C([0,1])$. It is easy to see that there exists a sequence $(h_n) \subset \mathcal{E}$ such that

$$\lim_{n\to+\infty} \left\| h_n - h \right\|_\infty = 0\ .$$

Let us fix a number $\zeta \in (0,1)$ and show the convergence in $L^2(\Omega)$ of $M_{h_n}^\varepsilon$, uniformly in $\varepsilon \in (0,\zeta)$. First, one can write

$$\begin{aligned} &\sup_{\varepsilon\in(0,\zeta)} \mathbb{E}\Big[|M_{h_n}^\varepsilon - M_h^\varepsilon|^2 \Big] \\ &= \sup_{\varepsilon\in(0,\zeta)} \frac{1}{\varepsilon^2 d(1/\varepsilon)^2} \mathbb{E}\left[\left| \int_0^1 q(x/\varepsilon)[h_n(x) - h(x)]\, dx \right|^2 \right] \\ &\le \left\| h_n - h \right\|_\infty^2 \times \sup_{\varepsilon\in(0,\zeta)} \frac{1}{\varepsilon^2 d(1/\varepsilon)^2} \int_{\mathbb{R}^2\setminus D} \left| R_q\left(\frac{y-x}{\varepsilon}\right) \right| dx\, dy\ , \end{aligned}$$

where $D = \{(x,y) \in [0,1]^2 : x = y\}$ is a negligible subset of $\mathbb{R}^2$. By (2.2),

$$\left| R_q(\frac{y-x}{\varepsilon}) \right| \le \mathrm{Cst}\, L\left(\left| \frac{y-x}{\varepsilon} \right| \right)^{2m} \left| \frac{y-x}{\varepsilon} \right|^{-2(1-H)}\ , \quad \forall (x,y) \in \mathbb{R}^2 \setminus D\,.$$

Secondly, with $\beta > 0$ small enough such that $2m\beta + 2(1-H) \in (0,1)$, we have

$$\begin{aligned} &\sup_{\varepsilon\in(0,\zeta)} \frac{1}{\mathfrak{X}(\varepsilon)^2} \int_{[0,1]^2\setminus D} \left| R_q\left(\frac{y-x}{\varepsilon}\right) \right| dx\, dy \\ &\le \mathrm{Cst} \sup_{\varepsilon\in(0,\zeta)} \int_{[0,1]^2\setminus D} \left\{ \frac{L(|(x-y)/\varepsilon|)}{L(1/\varepsilon)} \right\}^{2m} |x-y|^{-2(1-H)}\, dx\, dy \end{aligned}$$

$$(3.1) \qquad \leq \mathrm{Cst} \int_{[0,1]^2 \setminus D} |x-y|^{-2m\beta-2(1-H)}\,dx\,dy$$
$$< +\infty\ ,$$

where (3.1) follows from Theorem 2.2. It is now clear that, indeed,

$$\lim_{n\to+\infty} \sup_{\varepsilon\in(0,\zeta)} \mathbb{E}[|M^{\varepsilon}_{h_n} - M^{\varepsilon}_{h}|^2] = 0\,. \tag{3.2}$$

To conclude, let $d(\cdot,\cdot)$ denote any distance metrizing the convergence in distribution between real-valued random variables (for instance, the Fortet-Mourier distance). For $h \in C([0,1])$ and $(h_n) \subset \mathcal{E}$ converging to h, one can write, for any $\varepsilon > 0$ and $n \in \mathbb{N}$:

$$d(M^{\varepsilon}_h, M^0_h) \leq d(M^{\varepsilon}_h, M^{\varepsilon}_{h_n}) + d(M^{\varepsilon}_{h_n}, M^0_{h_n}) + d(M^0_{h_n}, M^0_h).$$

Fix $\eta > 0$. By (3.2), one can choose n big enough so that, for any $\varepsilon \in (0,\zeta)$, both $d(M^{\varepsilon}_h, M^{\varepsilon}_{h_n})$ and $d(M^0_{h_n}, M^0_h)$ are less than $\eta/3$. It remains to choose $\varepsilon > 0$ small enough so that $d(M^{\varepsilon}_{h_n}, M^0_{h_n})$ is less that $\eta/3$ (by (1.4) for the simple function $h_n \in \mathcal{E}$), to conclude that (1.4) holds true for any continuous function h.

Remark 3.1. Clearly, the above result still holds true for any function h that is continuous except at finitely many points. Note also that the function $\Phi \in \mathcal{G}_m$ is not necessarily bounded in Theorem 1.1.

3.2 Proof of Theorem 1.2

The proof is divided into five steps. We write $\mathfrak{X}(\varepsilon) = \varepsilon d(1/\varepsilon) = \sqrt{\frac{m!}{H(2H-1)}}\varepsilon^{1-H}L(1/\varepsilon)^m$.

(a) Preparation. Following [4], more precisely identities (5.1) and (5.19) therein, we first rewrite the rescaled corrector as follows:

$$\frac{u^{\varepsilon}(x) - \bar{u}(x)}{\mathfrak{X}(\varepsilon)} = \mathcal{U}^{\varepsilon}(x) + \underline{\frac{1}{\mathfrak{X}(\varepsilon)} r^{\varepsilon}(x) + \frac{1}{\mathfrak{X}(\varepsilon)}\rho^{\varepsilon}\frac{x}{a^*}}\ , \tag{3.3}$$

where the underlined term will be denoted by $\mathcal{R}^{\varepsilon}(x)$, and

$$\mathcal{U}^{\varepsilon}(x) = \frac{1}{\mathfrak{X}(\varepsilon)} \int_{\mathbb{R}} F(x,y) q(y/\varepsilon)\,dy\,,$$
$$r^{\varepsilon}(x) = (c^{\varepsilon} - c^*) \int_0^x q(y/\varepsilon)\,dy\,,$$

and

$$\rho^{\varepsilon} := \frac{a^*}{\int_0^1 a(y/\varepsilon)^{-1}\,dy}\Big[\Big(a^*b + \int_0^1 F(y)dy\Big)\Big(\int_0^1 q(y/\varepsilon)\,dy\Big)^2 - \int_0^1 F(y)q(y/\varepsilon)\,dy \times \int_0^1 q(y/\varepsilon)\,dy\Big].$$

Now, let us first show the weak convergence of $\mathcal{U}^\varepsilon$ to $\mathcal{U}$ in $C([0,1])$ and then prove that $\mathcal{R}^\varepsilon$ is a remainder. In order to prove the first claim, we start by establishing the f.d.d. convergence and then we prove the tightness.

(b) Convergence of finite dimensional distributions of $\mathcal{U}^\varepsilon$. For $x_1,\dots,x_n \in \mathbb{R}$ and $\lambda_1,\dots,\lambda_n \in \mathbb{R}$ $(n \geq 1)$, we have

$$\sum_{k=1}^n \lambda_k \mathcal{U}^\varepsilon(x_k) = \frac{1}{\mathfrak{X}(\varepsilon)} \int_{\mathbb{R}} \sum_{k=1}^n \lambda_k F(x_k, y) q(y/\varepsilon)\, dy.$$

Note that the function $\sum_{k=1}^n \lambda_k F(x_k,\cdot)$ have at most finitely many discontinuities. Thus, Theorem 1.1 and Remark 3.1 imply that $\sum_{k=1}^n \lambda_k \mathcal{U}^\varepsilon(x_k)$ converges in distribution to $\sum_{k=1}^n \lambda_k \mathcal{U}(x_k)$, yielding the desired convergence of finite dimensional distributions.

(c) Tightness of $\mathcal{U}^\varepsilon$. We check Kolmogorov's criterion ([6, Corollary 16.9]). First observe that $\mathcal{U}^\varepsilon(0) = 0$. Now, fix $0 \leq u < v \leq 1$, and set $F_1(y) = c^* - F(y)$ and $F_2(y) = F(y) - \int_0^1 F(t)\,dt - a^* b$, so that $F(x,y) = F_1(y)\mathbf{I}_{[0,x]}(y) + xF_2(y)\mathbf{I}_{[0,1]}(y)$. Then

$$\begin{aligned}
&\mathbb{E}(|\mathcal{U}^\varepsilon(u) - \mathcal{U}^\varepsilon(v)|^2) \\
&= \mathbb{E}\Big[\frac{1}{\mathfrak{X}(\varepsilon)^2} \Big| \int_0^1 \mathbf{I}_{(u,v]}(y) q(y/\varepsilon) F_1(y)\, dy + (v-u) \int_0^1 q(y/\varepsilon) F_2(y)\, dy \Big|^2 \Big] \\
&\leq \frac{2}{\mathfrak{X}(\varepsilon)^2} \mathbb{E}\Big[\Big| \int_0^1 \mathbf{I}_{(u,v]}(y) q(y/\varepsilon) F_1(y)\, dy \Big|^2 + \Big| (v-u) \int_0^1 q(y/\varepsilon) F_2(y)\, dy \Big|^2 \Big] \\
&\leq \frac{2}{\mathfrak{X}(\varepsilon)^2} \int_u^v \int_u^v F_1(x) F_1(y) R_q\Big(\frac{y-x}{\varepsilon}\Big)\, dx\, dy \qquad (3.4)\\
&\qquad + \frac{2(v-u)^2}{\mathfrak{X}(\varepsilon)^2} \int_0^1 \int_0^1 F_2(x) F_2(y) R_q\Big(\frac{y-x}{\varepsilon}\Big)\, dx\, dy\,. \qquad (3.5)
\end{aligned}$$

Note that F_2 is bounded on $[0,1]$. Therefore, as far as (3.5) is concerned, one can write, using Potter's Theorem as in the proof of Theorem 1.1,

$$\sup_{\varepsilon \in (0,\zeta)} \Big| \frac{(v-u)^2}{\mathfrak{X}(\varepsilon)^2} \int_0^1 \int_0^1 F_2(x) F_2(y) R_q\Big(\frac{y-x}{\varepsilon}\Big)\, dx\, dy \Big| \leq \mathrm{Cst}(v-u)^2\,.$$

Now, let us consider the term in (3.4). Similarly,

$$\begin{aligned}
&\sup_{\varepsilon \in (0,\zeta)} \frac{1}{\mathfrak{X}(\varepsilon)^2} \Big| \int_u^v \int_u^v F_1(x) F_1(y) R_q\Big(\frac{y-x}{\varepsilon}\Big)\, dx\, dy \Big| \\
&\leq \mathrm{Cst} \sup_{\varepsilon \in (0,\zeta)} \frac{1}{\mathfrak{X}(\varepsilon)^2} \int_u^v \int_u^v \Big| R_q\Big(\frac{y-x}{\varepsilon}\Big) \Big|\, dx\, dy \quad \text{(since } F_1 \text{ is bounded)} \\
&\leq \mathrm{Cst} \sup_{\varepsilon \in (0,\zeta)} \frac{1}{L(1/\varepsilon)^{2m}} \int_u^v \int_u^v L(|y-x|/\varepsilon)^{2m} \frac{dx\, dy}{|y-x|^{2(1-H)}} \\
&\leq \mathrm{Cst} \int_u^v \int_u^v |y-x|^{-2(1-H)-2m\beta}\, dy\, dx \quad \text{(similarly as in (3.1))}
\end{aligned}$$

$$= \mathrm{Cst}(v-u)^{2-2m\beta-2(1-H)} . \tag{3.6}$$

Since $2-2m(1-H_0)-2m\beta > 1$, this proves the tightness of $(\mathcal{U}^\varepsilon)_\varepsilon$ by means of the usual Kolmogorov's criterion.

(d) Control on the remainder term $\mathcal{R}^\varepsilon$ in (3.3). We shall prove that the process $\mathcal{R}^\varepsilon$ converges in probability to zero in $C([0,1])$. First we claim that if $G \in C([0,1])$, then there exists some constant $C = C(G)$ such that

$$\sup_{x\in[0,1]} \mathbb{E}\left[\left(\int_0^x q(y/\varepsilon)G(y)\,dy\right)^2\right] \le C\,\mathfrak{X}(\varepsilon)^2 . \tag{3.7}$$

Indeed, the same argument we used for bounding (3.5) works here as well:

$$\begin{aligned}
&\sup_{x\in[0,1]} \mathbb{E}\left[\left(\int_0^x q(y/\varepsilon)G(y)\,dy\right)^2\right] \\
&\le \|G\|_\infty^2 \int_{[0,1]^2} \left|R_q(|y-z|/\varepsilon)\right| dy\,dz \\
&\le \|G\|_\infty^2 \mathfrak{X}(\varepsilon)^2 \left(\sup_{\varepsilon\in(0,\zeta)} \frac{1}{\mathfrak{X}(\varepsilon)^2} \int_{[0,1]^2} \left|R_q(|y-z|/\varepsilon)\right| dy\,dz\right) \\
&\le \mathrm{Cst}\mathfrak{X}(\varepsilon)^2 ,
\end{aligned}$$

where the last inequality follows from (3.1).

Now, let us consider $\mathcal{R}^\varepsilon$:

(i) Due to the explicit expression of ρ^ε, it follows from (3.7), the fact that a is bounded from below and Cauchy-Schwarz inequalities that

$$\begin{aligned}
&\mathbb{E}[|\rho^\varepsilon|] \\
&\le \mathrm{Cst}\left\{\left\|\int_0^1 q(y/\varepsilon)\,dy\right\|_{L^2(\Omega)}^2 + \left\|\int_0^1 F(y)q(y/\varepsilon)\,dy\right\|_{L^2(\Omega)} \left\|\int_0^1 q(y/\varepsilon)\,dy\right\|_{L^2(\Omega)}\right\} \\
&\le \mathrm{Cst}\,\mathfrak{X}(\varepsilon)^2 .
\end{aligned}$$

(ii) Observe that

$$\begin{aligned}
c^\varepsilon - c^* &= a^* \int_0^1 \left(F(y) - \int_0^1 F(t)\,dt - ba^*\right) q(y/\varepsilon)\,dy + \rho^\varepsilon \\
&=: \int_0^1 \widehat{F}(y) q(y/\varepsilon)\,dy + \rho^\varepsilon.
\end{aligned}$$

Then

$$\begin{aligned}
\sup_{x\in[0,1]} \mathbb{E}[|r^\varepsilon(x)|] &= \sup_{x\in[0,1]} \mathbb{E}\left[\left|(c^\varepsilon - c^*)\int_0^x q(y/\varepsilon)\,dy\right|\right] \\
\le \sup_{x\in[0,1]} \mathbb{E}\left[\left|\int_0^1 \widehat{F}(y)q(y/\varepsilon)\,dy \int_0^x q(y/\varepsilon)\,dy\right|\right] &+ \mathrm{Cst}\,\mathbb{E}[|\rho^\varepsilon|] \le \mathrm{Cst}\,\mathfrak{X}(\varepsilon)^2 .
\end{aligned}$$

Therefore, as $\varepsilon \to 0$ we have

$$\sup_{x\in[0,1]} \mathbb{E}\Big[\big|\mathcal{R}^\varepsilon(x)\big|\Big] \le \mathrm{Cst}\mathfrak{X}(\varepsilon) \longrightarrow 0 \ . \quad \text{(by (2.9))}$$

In particular, $\{\mathcal{R}^\varepsilon(x), x \in [0,1]\}$ converges to zero in the sense of finite-dimensional distributions. Now, let us check the tightness of $(\mathcal{R}^\varepsilon)_\varepsilon$. Note that $\mathcal{R}^\varepsilon(0) = 0$ and that, for $0 \le u < v \le 1$,

$$\begin{aligned}
&\big\|\mathcal{R}^\varepsilon(u) - \mathcal{R}^\varepsilon(v)\big\|^2_{L^2(\Omega)} \\
\le\ & \frac{2}{\mathfrak{X}(\varepsilon)^2}\left\{ \big\|r^\varepsilon(u) - r^\varepsilon(v)\big\|^2_{L^2(\Omega)} + \frac{2(u-v)^2}{|a^*|^2}\mathbb{E}[|\rho^\varepsilon|^2] \right\} \\
\le\ & \frac{2}{\mathfrak{X}(\varepsilon)^2}\big\|r^\varepsilon(u) - r^\varepsilon(v)\big\|^2_{L^2(\Omega)} \\
&+ \mathrm{Cst}\frac{(u-v)^2}{\mathfrak{X}(\varepsilon)^2}\mathbb{E}[|\rho^\varepsilon|] \quad \text{(since } \rho^\varepsilon \text{ is uniformly bounded)} \\
\le\ & \frac{2}{\mathfrak{X}(\varepsilon)^2}\big\|r^\varepsilon(u) - r^\varepsilon(v)\big\|^2_{L^2(\Omega)} + \mathrm{Cst}(u-v)^2 \qquad \text{(by point (i) above)} \\
\le\ & \mathrm{Cst}\frac{1}{\mathfrak{X}(\varepsilon)^2}\int_{[u,v]^2} \big|R((y-z)/\varepsilon)\big|\,dy\,dz \\
&+ \mathrm{Cst}(u-v)^2 \quad \text{(since } c^\varepsilon - c^* \text{ is uniformly bounded)} \\
\le\ & \mathrm{Cst}(v-u)^{2-2(1-H)-2m\beta} + \mathrm{Cst}(v-u)^2\,,
\end{aligned}$$

where the last inequality follows from the same arguments as in (3.6). Therefore, $\mathcal{R}^\varepsilon$ converges in distribution to 0, as $\varepsilon \downarrow 0$, so it converges in probability to 0.

(e) Conclusion. Combining the results of (**a**) to (**d**), the proof of Theorem 1.2 is concluded by evoking Slutsky lemma.

References

[1] G. Bal, J. Garnier, S. Motsch and V. Perrier (2008): *Random integrals and correctors in homogenization*, Asymptot. Anal. 59 1-26.

[2] G. Bal and Y. Gu (2015): *Limiting models for equations with large random potential: A review*. Communications in Mathematical Sciences **13**, no. 3, pp. 729-748. Special Issue in Honor of George Papanicolaou's 70th Birthday, edited by L. Borcea, J.-P. Fouque, S. Jin, L. Ryzhik, and J. Xin.

[3] N. H. Bingham, C.M. Goldie and J.L. Teugels (1987): *Regular Variation*. Encyclopedia for Mathematics and its Applications **27**, edited by G.C. Rota.

[4] Y. Gu and G. Bal (2012): *Random homogenization and convergence to integrals with respect to the Rosenblatt process*. J. Diff. Equa. **253**, no. 4, pp. 1069-1087.

[5] V.V. Jikov, S.M. Kozlov and O.A. Oleinik (1994): *Homogenization of Differential Operators and Integral Functionals*, Springer, New York, 1994.

[6] O. KALLENBERG (2002): *Foundations of Modern Probability*, second edition, Probability and Its Applications, Springer, 2002.

[7] M. MAEJIMA and C.A. TUDOR (2007): *Wiener integrals with respect to the Hermite process and a non-central limit theorem*. Stoch. Anal. Appl. **25**, pp. 1043-1056.

[8] J.-C. MOURRAT and J. NOLEN (2015): *Scaling limit of the corrector in stochastic homogenization*, to appear in Ann. Appl. Probab.

[9] J.-C. MOURRAT and F. OTTO (2016): *Correlation structure of the corrector in stochastic homogenization*, Ann. Probab. 2016, Vol. 44, No. 5, 3207-3233.

[10] G.C. PAPANICOLAOU and S.R.S. VARADHAN (1981): *Boundary value problems with rapidly oscillating random coefficients*. Proceedings of Conference on Random Fields, (Esztergom, Hungary 1979) Seria Colloquia Mathematica Societatis Janos Bolyai, 27, pp. 835-873, North Holland.

[11] V. PIPIRAS and M.S. TAQQU (2000): *Integration questions related to fractional Brownian motion*. Probab. Theory Relat. Fields **118**, pp. 251-291.

[12] M.S. TAQQU (1979): *Convergence of integrated processes of arbitray Hermite rank*. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **50**, no. 1, pp 53-83.

This page is left blank

Paper 3: Exchangeable pairs on Wiener chaos

Ivan Nourdin and Guangqu Zheng

Submitted

Abstract

In [14], Nourdin and Peccati combined the Malliavin calculus and Stein's method of normal approximation to associate a rate of convergence to the celebrated fourth moment theorem [19] of Nualart and Peccati. Their analysis, known as the Malliavin-Stein method nowadays, has found many applications towards stochastic geometry, statistical physics and zeros of random polynomials, to name a few. In this article, we further explore the relation between these two fields of mathematics. In particular, we construct exchangeable pairs of Brownian motions and we discover a natural link between Malliavin operators and these exchangeable pairs. By combining our findings with E. Meckes' infinitesimal version of exchangeable pairs, we can give another proof of the quantitative fourth moment theorem. Finally, we extend our result to the multidimensional case.

Dedicated to the memory of Charles Stein, in remembrance of his beautiful mind and of his inspiring, creative, very original and deep mathematical ideas, which will, for sure, survive him for a long time.

1 Introduction

At the beginning of the 1970s, Charles Stein, one of the most famous statisticians of the time, introduced in [24] a new revolutionary method for establishing probabilistic approximations (now known as *Stein's method*), which is based on the breakthrough application of characterizing differential operators. The impact of Stein's method and its ramifications during the last 40 years is immense (see for instance the monograph [3]), and touches fields as diverse as combinatorics, statistics, concentration and functional inequalities, as well as mathematical physics and random matrix theory.

Introduced by Paul Malliavin [7], *Malliavin calculus* can be roughly described as an infinite-dimensional differential calculus whose operators act on sets of random objects associated with Gaussian or more general noises. In 2009, Nourdin and Peccati [14] combined the Malliavin calculus and Stein's method for the first time, thus virtually creating a new domain of research, which is now commonly known as the *Malliavin-Stein method*. The success of their method relies crucially on the existence of integration-by-parts formulae on both sides: on one side, the Stein's lemma is built on the Gaussian integration-by-parts formula and it is one of the cornerstones of the Stein's method; on the other side, the integration-by-parts formula on Gaussian space is one of the main tools in Malliavin calculus. Interested readers can refer to the constantly updated website [13] and the monograph [15] for a detailed overview of this active field of research.

Key words and phrases. Stein's method; Exchangeable pairs; Brownian motion; Malliavin calculus.

A prominent example of applying Malliavin-Stein method is the obtention (see also (1.1) below) of a Berry-Esseen's type rate of convergence associated to the celebrated fourth moment theorem [19] of Nualart and Peccati, according to which a standardized sequence of multiple Wiener-Itô integrals converges in law to a standard Gaussian random variable if and only if its fourth moment converges to 3.

Theorem 1.1. (i) (Nualart, Peccati [19]) Let (F_n) be a sequence of multiple Wiener-Itô integrals of order p, for some fixed $p \geq 1$. Assume that $E[F_n^2] \to \sigma^2 > 0$ as $n \to \infty$. Then, as $n \to \infty$, we have the following equivalence:

$$F_n \overset{\text{law}}{\to} N(0, \sigma^2) \quad \Longleftrightarrow \quad E[F_n^4] \to 3\sigma^4.$$

(ii) (Nourdin, Peccati [14, 15]) Let F be any multiple Wiener-Itô integral of order $p \geq 1$, such that $E[F^2] = \sigma^2 > 0$. Then, with $N \sim N(0, \sigma^2)$ and d_{TV} standing for the total variation distance,

$$d_{TV}(F, N) \leq \frac{2}{\sigma^2} \sqrt{\frac{p-1}{3p}} \sqrt{E[F^4] - 3\sigma^4}.$$

Of course, (ii) was obtained several years after (i), and (ii) implies '$\Leftarrow$' in (i). Nualart and Peccati's fourth moment theorem has been the starting point of a number of applications and generalizations by dozens of authors. These collective efforts have allowed one to break several long-standing deadlocks in several domains, ranging from stochastic geometry (see *e.g.* [6, 21, 23]) to statistical physics (see *e.g.* [8, 9, 10]), and zeros of random polynomials (see *e.g.* [1, 2, 4]), to name a few. At the time of writing, more than two hundred papers have been written, which use in one way or the other the Malliavin-Stein method (see again the webpage [13]).

Malliavin-Stein method has become a popular tool, especially within the Malliavin calculus community. Nevertheless, and despite its success, it is less used by researchers who are not specialists of the Malliavin calculus. A possible explanation is that it requires a certain investment before one is in a position to be able to use it, and doing this investment may refrain people who are not originally trained in the Gaussian analysis. This paper takes its root from this observation.

During our attempt to make the proof of Theorem 1.1(ii) more accessible to readers having no background on Malliavin calculus , we discover the following interesting fact for exchangeable pairs of multiple Wiener-Itô integrals. When $p \geq 1$ is an integer and f belongs to $L^2([0,1]^p)$, we write $I_p^B(f)$ to indicate the multiple Wiener-Itô integral of f with respect to Brownian motion B, see Section 2 for the precise meaning.

Proposition 1.1. Let $(B, B^t)_{t\geq 0}$ be a family of exchangeable pairs of Brownian motions (that is, B is a Brownian motion on $[0,1]$ and, for each t, one has $(B, B^t) \overset{\text{law}}{=} (B^t, B)$). Assume moreover that

(a) for any integer $p \geq 1$ and any $f \in L^2([0,1]^p)$,

$$\lim_{t \downarrow 0} \frac{1}{t} E\Big[I_p^{B^t}(f) - I_p^B(f) \Big| \sigma\{B\} \Big] = -p\, I_p^B(f) \quad \text{in } L^2(\Omega).$$

Then, for any integer $p \geq 1$ and any $f \in L^2([0,1]^p)$,

(b) $\lim_{t\downarrow 0} \frac{1}{t} E\Big[(I_p^{B^t}(f) - I_p^B(f))^2 | \sigma\{B\}\Big] = 2p^2 \int_0^1 I_{p-1}^B(f(x,\cdot))^2 dx \quad \text{in } L^2(\Omega);$

(c) $\lim_{t\downarrow 0} \frac{1}{t} E\Big[(I_p^{B^t}(f) - I_p^B(f))^4\Big] = 0.$

Why is this proposition interesting? Because, as it turns out, it combines perfectly well with the following result, which represents the main ingredient from Stein's method we will rely on and which corresponds to a slight modification of a theorem originally due to Elizabeth Meckes (see [11, Theorem 2.1]).

Theorem 1.2 (Meckes [11])**.** Let F and a family of random variables $(F_t)_{t\geq 0}$ be defined on a common probability space $(\Omega, \mathcal{F}, P)$ such that $F_t \overset{law}{=} F$ for every $t \geq 0$. Assume that $F \in L^3(\Omega, \mathcal{G}, P)$ for some σ-algebra $\mathcal{G} \subset \mathcal{F}$ and that in $L^1(\Omega)$,

(a) $\lim_{t\downarrow 0} \frac{1}{t} E[F_t - F | \mathcal{G}] = -\lambda F$ for some $\lambda > 0$,

(b) $\lim_{t\downarrow 0} \frac{1}{t} E[(F_t - F)^2 | \mathcal{G}] = (2\lambda + S)\mathrm{Var}(F)$ for some random variable S,

(c) $\lim_{t\downarrow 0} \frac{1}{t} (F_t - F)^3 = 0.$

Then, with $N \sim N(0, \mathrm{Var}(F))$,

$$d_{TV}(F, N) \leq \frac{E|S|}{\lambda}.$$

To see how to combine Proposition 1.1 with Theorem 1.2 (see also point(ii) in Remark 5.1), consider indeed a multiple Wiener-Itô integral of the form $F = I_p^B(f)$, with $\sigma^2 = E[F^2] > 0$. Assume moreover that we have at our disposal a family $\{(B, B^t)\}_{t\geq 0}$ of exchangeable pairs of Brownian motions, satisfying the assumption (a) in Proposition 1.1. Then, putting Proposition 1.1 and Theorem 1.2 together immediately yields that

$$d_{TV}(F, N) \leq \frac{2}{\sigma^2} E\left[\left| p \int_0^1 I_{p-1}^B(f(x,\cdot))^2 dx - \sigma^2 \right|\right]. \tag{1.1}$$

Finally, to obtain the inequality stated Theorem 1.1(ii) from (1.1), it remains to 'play' cleverly with the (elementary) product formula (2.2), see Proposition 7.1 for the details.

To conclude our elementary proof of Theorem 1.1(ii), we are thus left to construct the family $\{(B, B^t)\}_{t>0}$. Actually, we will offer two constructions with different motivations: the first one is inspired by Mehler's formula from Gaussian analysis, whereas the second one is more in the spirit of the so-called *Gibbs sampling procedure* within Stein's method (see *e.g.* [5, A.2]).

For the first construction, we consider two independent Brownian motions on $[0,1]$ defined on the same probability space $(\Omega, \mathcal{F}, P)$, namely B and $\widehat{B}$. We interpolate between them by considering, for any $t \geq 0$,

$$B^t = e^{-t}B + \sqrt{1 - e^{-2t}}\widehat{B}.$$

It is then easy and straightforward to check that, for any $t \geq 0$, this new Brownian motion B^t, together with B, forms an exchangeable pair (see Lemma 3.1). Moreover, we will compute below (see (3.1)) that $E\big[I_p^{B^t}(f)\big|\sigma\{B\}\big] = e^{-pt}\, I_p^B(f)$ for any $p \geq 1$ and any $f \in L^2([0,1]^p)$, from which (a) in Proposition 1.1 immediately follows.

For the second construction, we consider two independent Gaussian white noise W and W' on $[0,1]$ with Lebesgue intensity measure. For each $n \in \mathbb{N}$, we introduce a uniform partition $\{\Delta_1, \ldots, \Delta_n\}$ and a uniformly distributed index $I_n \sim \mathscr{U}_{\{1,\ldots,n\}}$, independent of W and W'. For every Borel set $A \subset [0,1]$, we define $W^n(A) = W'(A \cap \Delta_{I_n}) + W(A \setminus \Delta_{I_n})$. This will give us a new Gaussian white noise W^n, which will form an exchangeable pair with W. This construction is a particular Gibbs sampling procedure. The analogue of (a) in Proposition 1.1 is satisfied, namely, if $f \in L^2([0,1]^p)$, $F = I_p^W(f)$ is the pth multiple integral with respect to W and $F^{(n)} = I_p^{W^n}(f)$, we have

$$nE\big[F^{(n)} - F\big|\sigma\{W\}\big] \to -pF \quad \text{in } L^2(\Omega) \text{ as } n \to \infty.$$

To apply Theorem 1.2 in this setting, we only need to replace $\frac{1}{t}$ by n and replace F_t by $F^{(n)}$. To get the exchangeable pairs (B, B^n) of Brownian motions in this setting, it suffices to consider $B(t) = W([0,t])$ and $B^n(t) = W^n([0,t])$, $t \in [0,1]$. See Section 4 for more precise statements.

Finally, we discuss the extension of our exchangeable pair approach on Wiener chaos to the multidimensional case. Here again, it works perfectly well, and it allows us to recover the (known) rate of convergence associated with the remarkable Peccati-Tudor theorem [20]. This latter represents a multidimentional counterpart of the fourth moment theorem Theorem 1.1(i), exhibiting conditions involving only the second and fourth moments that ensure a central limit theorem for random *vectors* with chaotic components.

Theorem 1.3 (Peccati, Tudor [20])**.** Fix $d \geq 2$ and $p_1, \ldots, p_d \geq 1$. For each $k \in \{1, \ldots, d\}$, let $(F_n^k)_{n\geq 1}$ be a sequence of multiple Wiener-Itô integrals of order p_k. Assume that $E[F_n^k F_n^l] \to \sigma_{kl}$ as $n \to \infty$ for each pair $(k,l) \in \{1, \ldots, d\}^2$, with $\Sigma = (\sigma_{kl})_{1\leq k,l\leq d}$ non-negative definite. Then, as $n \to \infty$,

$$F_n = (F_n^1, \ldots, F_n^d) \xrightarrow{\text{law}} N \sim N(0,\Sigma) \quad \Longleftrightarrow \quad E[(F_n^k)^4] \to 3\sigma_{kk}^2 \text{ for all } k \in \{1, \ldots, d\}. \tag{1.2}$$

In [17], it is shown that the right-hand side of (1.2) is also equivalent to

$$E[\|F_n\|^4] \to E[\|N\|^4] \quad \text{as } n \to \infty, \tag{1.3}$$

where $\|\cdot\|$ stands for the usual Euclidean ℓ^2-norm of $\mathbb{R}^d$. Combining the main findings of [16] and [17] yields the following quantitative version associated to Theorem 1.3, which we are able to recover by means of our elementary exchangeable approach.

Theorem 1.4 (Nourdin, Peccati, Réveillac, Rosiński [16, 17])**.** Let $F = (F^1, \ldots, F^d)$ be a vector composed of multiple Wiener-Itô integrals F^k, $1 \leq k \leq d$. Assume that the covariance matrix Σ of F is invertible. Then, with $N \sim N(0,\Sigma)$,

$$d_W(F,N) \leq \|\Sigma\|_{op}^{\frac{1}{2}} \|\Sigma^{-1}\|_{op} \sqrt{E[\|F\|^4] - E[\|N\|^4]}, \tag{1.4}$$

where d_W denotes the Wasserstein distance and $\|\cdot\|_{op}$ the operator norm of a matrix.

The currently available proof of (1.4) relies on two main ingredients: (i) simple manipulations involving the product formula (2.2) and implying that

$$\sum_{i,j=1}^{d} \mathrm{Var}\Big(p_j \int_0^1 I_{p_i-1}(f_i(x,\cdot))I_{p_j-1}(f_j(x,\cdot))dx\Big) \le E[\|F\|^4] - E[\|N\|^4],$$

(see [17, Theorem 4.3] for the details) and (ii) the following inequality shown in [16, Corollary 3.6] by means of the Malliavin operators D, δ and L:

$$d_W(F,N) \le \|\Sigma\|_{op}^{\frac{1}{2}}\|\Sigma^{-1}\|_{op} \sqrt{\sum_{i,j=1}^{d} \mathrm{Var}\Big(p_j \int_0^1 I_{p_i-1}(f_i(x,\cdot))I_{p_j-1}(f_j(x,\cdot))dx\Big)}. \tag{1.5}$$

Here, in the spirit of what we have done in dimension one, we also apply our elementary exchangeable pairs approach to prove (1.5), with slightly different constants.

The rest of the paper is organized as follows. Section 2 contains preliminary knowledge on multiple Wiener-Itô integrals. In Section 3 (resp. 4), we present our first (resp. second) construction of exchangeable pairs of Brownian motions, and we give the main associated properties. Section 5 is devoted to the proof of Proposition 1.1, whereas in Section 6 we offer a simple proof of Meckes' Theorem 1.2. Our new, elementary proof of Theorem 1.1(ii) is given in Section 7. In Section 8, we further investigate the connections between our exchangeable pairs and the Malliavin operators. Finally, we discuss the extension of our approach to the multidimensional case in Section 9.

Acknowledgement. We would like to warmly thank Christian Döbler and Giovanni Peccati, for very stimulating discussions on exchangeable pairs since the early stage of this work.

2 Multiple Wiener-Itô integrals: definition and elementary properties

In this subsection, we recall the definition of multiple Wiener-Itô integrals, and then we give a few *soft* properties that will be needed for our new proof of Theorem 1.1(ii). We refer to the classical monograph [18] for the details and missing proofs.

Let $f : [0,1]^p \to \mathbb{R}$ be a square-integrable function, with $p \ge 1$ a given integer. The pth multiple Wiener-Itô integral of f with respect to the Brownian motion $B = (B(x))_{x\in[0,1]}$ is *formally* written as

$$\int_{[0,1]^p} f(x_1,\ldots,x_p)dB(x_1)\ldots dB(x_p). \tag{2.1}$$

To give a precise meaning to (2.1), Itô's crucial idea from the fifties was to first define (2.1) for elementary functions that vanish on diagonals, and then to approximate any f in $L^2([0,1]^p)$ by such elementary functions.

Consider the diagonal set of $[0,1]^p$, that is, $D = \{(t_1,\dots,t_p) \in [0,1]^p : \exists i \neq j,\ t_i = t_j\}$. Let $\mathcal{E}_p$ be the vector space formed by the set of elementary functions on $[0,1]^p$ that vanish over D, that is, the set of those functions f of the form

$$f(x_1,\dots,x_p) = \sum_{i_1,\dots,i_p=1}^{k} \beta_{i_1\dots i_p} \mathbf{I}_{[\tau_{i_1-1},\tau_{i_1})\times\dots\times[\tau_{i_p-1},\tau_{i_p})}(x_1,\dots,x_p),$$

where $k \geq 1$ and $0 = \tau_0 < \tau_1 < \dots < \tau_k$, and the coefficients $\beta_{i_1\dots i_p}$ are zero if any two of the indices $i_1,\dots,i_p$ are equal. For $f \in \mathcal{E}_p$, we define (without ambiguity with respect to the choice of the representation of f)

$$I_p^B(f) = \sum_{i_1,\dots,i_p=1}^{k} \beta_{i_1\dots i_p}(B(\tau_{i_1}) - B(\tau_{i_1-1}))\dots(B(\tau_{i_p}) - B(\tau_{i_p-1})).$$

We also define the symmetrization $\widetilde{f}$ of f by

$$\widetilde{f}(x_1,\dots,x_p) = \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p} f(x_{\sigma(1)},\dots,x_{\sigma(p)}),$$

where $\mathfrak{S}_p$ stands for the set of all permutations of $\{1,\dots,p\}$. The following elementary properties are immediate and easy to prove.

1. If $f \in \mathcal{E}_p$, then $I_p^B(f) = I_p^B(\widetilde{f})$.
2. If $f \in \mathcal{E}_p$ and $g \in \mathcal{E}_q$, then $E[I_p^B(f)] = 0$ and $E[I_p^B(f)I_q^B(g)] = \begin{cases} 0 & \text{if } p \neq q \\ p!\langle \widetilde{f}, \widetilde{g}\rangle_{L^2([0,1]^p)} & \text{if } p = q \end{cases}$.
3. The space $\mathcal{E}_p$ is dense in $L^2([0,1]^p)$. In other words, to each $f \in L^2([0,1]^p)$ one can associate a sequence $(f_n)_{n\geq1} \subset \mathcal{E}_p$ such that $\|f - f_n\|_{L^2([0,1]^p)} \to 0$ as $n \to \infty$.
4. Since $E[(I_p^B(f_n) - I_p^B(f_m))^2] = p!\|\widetilde{f_n} - \widetilde{f_m}\|^2_{L^2([0,1]^p)} \leq p!\|f_n - f_m\|^2_{L^2([0,1]^p)} \to 0$ as $n, m \to \infty$ for f and $(f_n)_{n\geq1}$ as in the previous point 3, we deduce that the sequence $(I_p(f_n))_{n\geq1}$ is Cauchy in $L^2(\Omega)$ and, as such, it admits a limit denoted by $I_p^B(f)$. It is easy to check that $I_p^B(f)$ only depends on f, not on the particular choice of the approximating sequence $(f_n)_{n\geq1}$, and that points 1 to 3 continue to hold for general $f \in L^2([0,1]^p)$ and $g \in L^2([0,1]^q)$.

We will also crucially rely on the following *product formula*, whose proof is elementary and can be made by induction. See, *e.g.*, [18, Proposition 1.1.3].

5. For any $p, q \geq 1$, and if $f \in L^2([0,1]^p)$ and $g \in L^2([0,1]^q)$ are symmetric, then

$$I_p^B(f)I_q^B(g) = \sum_{r=0}^{p\wedge q} r!\binom{p}{r}\binom{q}{r} I^B_{p+q-2r}(f\otimes_r g), \tag{2.2}$$

where $f\otimes_r g$ stands for the rth-contraction of f and g, defined as an element of $L^2([0,1]^{p+q-2r})$ by

$$\begin{aligned}&(f\otimes_r g)(x_1,\dots,x_{p+q-2r})\\ &= \int_{[0,1]^r} f(x_1,\dots,x_{p-r},u_1,\dots,u_r)g(x_{p-r+1},\dots,x_{p+q-2r},u_1,\dots,u_r)du_1\dots du_r.\end{aligned}$$

Product formula (2.2) has a nice consequence, the inequality (2.3) below. It is a very particular case of a more general phenomenon satisfied by multiple Wiener-Itô integrals, the *hypercontractivity* property.

6. For any $p \geq 1$, there exists a constant $c_{4,p} > 0$ such that, for any (symmetric) $f \in L^2([0,1]^p)$,

$$E[I_p^B(f)^4] \leq c_{4,p}\, E[I_p^B(f)^2]^2 . \tag{2.3}$$

Indeed, thanks to (2.2) one can write $I_p^B(f)^2 = \sum_{r=0}^{p} r! \binom{p}{r}^2 I_{2p-2r}^B(f \otimes_r f)$ so that

$$E[I_p^B(f)^4] = \sum_{r=0}^{p} r!^2 \binom{p}{r}^4 (2p-2r)! \|f \widetilde{\otimes}_r f\|^2_{L^2([0,1]^{2p-2r})}.$$

The conclusion (2.3) follows by observing that

$$p!^2 \|f \widetilde{\otimes}_r f\|^2_{L^2([0,1]^{2p-2r})} \leq p!^2 \|f \otimes_r f\|^2_{L^2([0,1]^{2p-2r})} \leq p!^2 \|f\|^4_{L^2([0,1]^p)} = E[I_p^B(f)^2]^2.$$

Furthermore, for each $n \geq 2$, using (2.2) and induction, one can show that, with $c_{2^n,p}$ a constant depending only on p but not on f,

$$E[I_p^B(f)^{2^n}] \leq c_{2^n,p}\, E[I_p^B(f)^2]^{2^{n-1}} .$$

So for any $r > 2$, there exists an absolute constant $c_{r,p}$ depending only on p, r (but not on f) such that

$$E[|I_p^B(f)|^r] \leq c_{r,p}\, E[I_p^B(f)^2]^{r/2} . \tag{2.4}$$

3 Exchangeable pair of Brownian motions: a first construction

As anticipated in the introduction, for this construction we consider two independent Brownian motions on $[0,1]$ defined on the same probability space $(\Omega, \mathcal{F}, P)$, namely B and $\widehat{B}$, and we interpolate between them by considering, for any $t \geq 0$, $B^t = e^{-t}B + \sqrt{1-e^{-2t}}\widehat{B}$.

Lemma 3.1. For each $t \geq 0$, the pair (B, B^t) is exchangeable, that is, $(B, B^t) \overset{\text{law}}{=} (B^t, B)$. In particular, B^t is a Brownian motion.

Proof. Clearly, the bi-dimensional process (B, B^t) is Gaussian and centered. Moreover, for any $x, y \in [0,1]$,

$$\begin{aligned} E[B^t(x)B^t(y)] &= e^{-2t}E[B(x)B(y)] + (1-e^{-2t})E[\widehat{B}(x)\widehat{B}(y)] = E[B(x)B(y)] \\ E[B(x)B^t(y)] &= e^{-t}E[B(x)B(y)] = E[B^t(x)B(y)]. \end{aligned}$$

The desired conclusion follows. □

We can now state that, as written in the introduction, our exchangeable pair indeed satisfies the crucial property (a) of Proposition 1.1.

Theorem 3.1. Let $p \geq 1$ be an integer, and consider a kernel $f \in L^2([0,1]^p)$. Set $F = I_p^B(f)$ and $F_t = I_p^{B^t}(f)$, $t \geq 0$. Then,

$$E\big[F_t \big| \sigma\{B\}\big] = e^{-pt} F. \tag{3.1}$$

In particular, convergence (a) in Proposition 1.1 takes place:

$$\lim_{t \downarrow 0} \frac{1}{t} E\Big[I_p^{B^t}(f) - I_p^B(f) \Big| \sigma\{B\}\Big] = -p\, I_p^B(f) \quad \text{in } L^2(\Omega). \tag{3.2}$$

Proof. Consider first the case where $f \in \mathcal{E}_p$, that is, f has the form

$$f(x_1, \dots, x_p) = \sum_{i_1,\dots,i_p=1}^{k} \beta_{i_1\dots i_p} \mathbf{I}_{[\tau_{i_1-1},\tau_{i_1})\times\dots\times[\tau_{i_p-1},\tau_{i_p})}(x_1,\dots,x_p),$$

with $k \geq 1$ and $0 = \tau_0 < \tau_1 < \dots < \tau_k$, and the coefficients $\beta_{i_1\dots i_p}$ are zero if any two of the indices $i_1, \dots, i_p$ are equal. We then have

$$\begin{aligned} F_t &= \sum_{i_1,\dots,i_p=1}^{k} \beta_{i_1\dots i_p} (B^t(\tau_{i_1}) - B^t(\tau_{i_1-1})) \dots (B^t(\tau_{i_p}) - B^t(\tau_{i_p-1})) \\ &= \sum_{i_1,\dots,i_p=1}^{k} \beta_{i_1\dots i_p} [e^{-t}(B(\tau_{i_1}) - B(\tau_{i_1-1})) + \sqrt{1-e^{-2t}}(\widehat{B}(\tau_{i_1}) - \widehat{B}(\tau_{i_1-1}))] \\ &\qquad \times \dots \times [e^{-t}(B(\tau_{i_1}) - B(\tau_{i_1-1})) + \sqrt{1-e^{-2t}}(\widehat{B}(\tau_{i_p}) - \widehat{B}(\tau_{i_p-1}))]. \end{aligned}$$

Expanding and integrating with respect to $\widehat{B}$ yields (3.1) for elementary f. Thanks to point 4 in Section 2, we can extend it to any $f \in L^2([0,1]^p)$. We then deduce that

$$\frac{1}{t} E\big[F_t - F \big| \sigma\{B\}\big] = \frac{e^{-pt} - 1}{t} F,$$

from which (3.2) now follows immediately. □

4 Exchangeable pair of Brownian motions: a second construction

In this section, we present yet another construction of exchangeable pairs via Gaussian white noise. We believe it is of independent interest, as such a construction can be similarly carried out for other additive noises. This part may be skipped in a first reading, as it is not used in other sections. And we assume that the readers are familiar with the multiple Wiener-Itô integrals with respect to the Gaussian white noise, and refer to [18, Page 8-13] for all missing details.

Let W be a Gaussian white noise on $[0,1]$ with Lebesgue intensity measure ν, that is, W is a centred Gaussian process indexed by Borel subsets of $[0,1]$ such that for any Borel sets $A, B \subset [0,1]$, $W(A) \sim N(0, \nu(A))$ and $E[W(A)W(B)] = \nu(A \cap B)$. We denote by $\mathscr{G} := \sigma\{W\}$

the σ-algebra generated by $\{W(A)$: A Borel subset of $[0,1]\}$. Now let W' be an independent copy of W (denote by $\mathscr{G}' = \sigma\{W'\}$ the σ-algebra generated by W') and I_n be a uniform random variable over $\{1,\dots,n\}$ for each $n \in \mathbb{N}$ such that I_n, W, W' are independent. For each fixed $n \in \mathbb{N}$, we consider the partition $[0,1] = \bigcup_{j=1}^n \Delta_j$ with $\Delta_1 = [0, \frac{1}{n}]$, $\Delta_2 = (\frac{1}{n}, \frac{2}{n}], \dots, \Delta_n = (1-\frac{1}{n}, 1]$.

Definition 4.0. *Set* $W^n(A) := W'(A \cap \Delta_{I_n}) + W(A \setminus \Delta_{I_n})$ *for any Borel set* $A \subset [0,1]$.

Remark 4.1. One can first treat W as the superposition of $\{W|_{\Delta_j}, j = 1,\dots,n\}$, where $W|_{\Delta_j}$ denotes the Gaussian white noise on Δ_j. Then according to $I_n = j$, we (only) replace $W|_{\Delta_j}$ by an independent copy $W'|_{\Delta_j}$ so that we get W^n. This is nothing else but a particular Gibbs sampling procedure (see [5, A.2]), hence heuristically speaking, the new process W^n shall form an exchangeable pair with W.

Lemma 4.1. W and W^n form an exchangeable pair with W, that is, $(W, W^n) \overset{law}{=} (W^n, W)$. In particular, W^n is a Gaussian white noise on $[0,1]$ with Lebesgue intensity measure.

Proof. Let us first consider m mutually disjoint Borel sets $A_1,\dots,A_m \subset [0,1]$. Given D_1, D_2 Borel subsets of $\mathbb{R}^m$, we have

$$\begin{aligned}
&P\Big((W(A_1),\dots,W(A_m)) \in D_1\,,\ (W^n(A_1),\dots,W^n(A_m)) \in D_2\Big)\\
&= \sum_{v=1}^n P\Big((W(A_1),\dots,W(A_m)) \in D_1\,,\ (W^n(A_1),\dots,W^n(A_m)) \in D_2\,,\ I_n = v\Big)\\
&= \frac{1}{n}\ \sum_{v=1}^n P\Big(g(X_v, Y_v) \in D_1, g(X'_v, Y_v) \in D_2\Big)\ ,
\end{aligned}$$

where for each $v \in \{1,\dots,n\}$,

- $X_v := (W(A_1 \cap \Delta_v),\dots,W(A_m \cap \Delta_v))$, $X'_v := (W'(A_1 \cap \Delta_v),\dots,W'(A_m \cap \Delta_v))$,
- $Y_v := (W(A_1 \setminus \Delta_v),\dots,W(A_m \setminus \Delta_v))$, and g is a function from $\mathbb{R}^{2m}$ to $\mathbb{R}^m$ given by $(x_1,\dots,x_m,y_1,\dots,y_m) \mapsto g(x_1,\dots,x_m,y_1,\dots,y_m) = (x_1+y_1,\dots,x_m+y_m)$

It is clear that for each $v \in \{1,\dots,n\}$, X_v, X'_v and Y_v are independent, therefore $g(X_v, Y_v)$ and $g(X'_v, Y_v)$ form an exchangeable pair. It follows from the above equalities that

$$\begin{aligned}
&P\Big((W(A_1),\dots,W(A_m)) \in D_1\,,\ (W^n(A_1),\dots,W^n(A_m)) \in D_2\Big)\\
&= \frac{1}{n}\ \sum_{v=1}^n P\Big(g(X'_v, Y_v) \in D_1, g(X_v, Y_v) \in D_2\Big)\\
&= P\Big((W^n(A_1),\dots,W^n(A_m)) \in D_1\,,\ (W(A_1),\dots,W(A_m)) \in D_2\Big)\ .
\end{aligned}$$

This proves the exchangeability of $(W(A_1),\dots,W(A_m))$ and $(W^n(A_1),\dots,W^n(A_m))$.

Now let $B_1,\dots,B_m$ be Borel subsets of $[0,1]$, then one can find mutually disjoint Borel sets $A_1,\dots,A_p$ (for some $p \in \mathbb{N}$) such that each B_j is a union of some of A_i's. Therefore we can find some measurable $\phi : \mathbb{R}^p \to \mathbb{R}^m$ such that $(W(B_1),\dots,W(B_m)) = \phi(W(A_1),\dots,W(A_p))$. Accordingly, $(W^n(B_1),\dots,W^n(B_m)) = \phi(W^n(A_1),\dots,W^n(A_p))$, hence $(W(B_1),\dots,W(B_m))$ and $(W^n(B_1),\dots,W^n(B_m))$ are exchangeable. Now our proof is complete. $\square$

Remark 4.2. For each $t \in [0,1]$, we set $B(t) := W([0,t])$ and $B^n(t) := W^n([0,t])$. Modulo continuous modifications, one can see from Lemma 4.1 that B, B^n are two Brownian motions that form an exchangeable pair. An important difference between this construction and the previous one is that (B, B^t) is bi-dimensional Gaussian process whereas B, B^n are not *jointly* Gaussian.

Before we state the analogous result to Theorem 3.1, we briefly recall the construction of multiple Wiener-Itô integrals in white noise setting.

1. For each $p \in \mathbb{N}$, we denote by $\mathscr{E}_p$ the set of simple functions of the form

$$f(t_1, \dots, t_p) = \sum_{i_1,\dots,i_p=1}^{m} \beta_{i_1\dots i_p} \mathbf{I}_{A_{i_1}\times\dots\times A_{i_p}}(t_1, \dots, t_p) \ , \tag{4.1}$$

   where $m \in \mathbb{N}$, $A_1, \dots, A_m$ are pair-wise disjoint Borel subsets of $[0,1]$, and the coefficients $\beta_{i_1\dots i_p}$ are zero if any two of the indices $i_1, \dots i_p$ are equal. It is known that $\mathscr{E}_p$ is dense in $L^2([0,1]^p)$.

2. For f given as in (4.1), the pth multiple integral with respect to W is defined as

$$I_p^W(f) := \sum_{i_1,\dots,i_p=1}^{m} \beta_{i_1\dots i_p} W(A_{i_1}) \dots W(A_{i_p}) \ ,$$

   and one can extend I_p^W to $L^2([0,1]^p)$ via usual approximation argument. Note $I_p^W(f)$ is nothing else but $I_p^B(f)$ with the Brownian motion B constructed in Remark 4.2.

Theorem 4.1. If $F = I_p^W(f)$ for some symmetric $f \in L^2([0,1]^p)$ and we set $F^{(n)} := I_p^{W^n}(f)$, then in $L^2(\Omega, \mathscr{G}, P)$ and as $n \to +\infty$, $n\, E[F^{(n)} - F \big| \mathscr{G}] \to -pF$.

Proof. First we consider the case where $f \in \mathscr{E}_p$, we assume moreover that $F = \prod_{j=1}^p W(A_j)$ with $A_1, \dots, A_p$ mutually disjoint Borel subsets of $[0,1]$, and accordingly we define $F^{(n)} = \prod_{j=1}^p W^n(A_j)$. Then, (we write $[p] = \{1, \dots, p\}$, $A^v = A \cap \Delta_v$ for any $A \subset [0,1]$ and $v \in \{1, \dots, n\}$)

$$\begin{aligned}
n\, E[F^{(n)} \big| \mathscr{G}] &= n\, E\left\{ \sum_{v=1}^{n} \mathbf{1}_{\{I_n = v\}} \prod_{j=1}^{p} [W'(A_j^v) + W(A_j \setminus \Delta_v)] \ \big| \mathscr{G} \right\} \\
&= \sum_{v=1}^{n} E\left\{ \prod_{j=1}^{p} [W'(A_j^v) + W(A_j \setminus \Delta_v)] \ \big| \mathscr{G} \right\} = \sum_{v=1}^{n} \prod_{j=1}^{p} W(A_j \setminus \Delta_v) \\
&= \sum_{v=1}^{n} \left\{ \left(\prod_{j=1}^{p} W(A_j) \right) - \sum_{k=1}^{p} W(A_k^v) \left(\prod_{j\in[p]\setminus\{k\}} W(A_j) \right) \right. \\
&\qquad \left. + \sum_{\ell=2}^{p} (-1)^\ell \sum_{\substack{k_1,\dots,k_\ell \in [p] \\ \text{all distinct}}} \left(\prod_{j\in[p]\setminus\{k_1,\dots,k_\ell\}} W(A_j) \right) W(A_{k_1}^v) \cdots W(A_{k_\ell}^v) \right\} \\
&= n\,F - p\,F + R_n(F) \ ,
\end{aligned}$$

$$\text{where} \quad R_n(F) = \sum_{\ell=2}^{p} (-1)^\ell \sum_{\substack{k_1,\dots,k_\ell \in [p] \\ \text{all distinct}}} \left(\prod_{j \in [p] \setminus \{k_1,\dots,k_\ell\}} W(A_j) \right) \sum_{v=1}^{n} W(A^v_{k_1}) \cdots W(A^v_{k_\ell}) \,.$$

Then $R_n(F)$ converges in $L^2(\Omega, \mathscr{G}, P)$ to 0, due to the fact that $\sum_{v=1}^n \prod_{i=1}^q W(A^v_{k_i})$ converges in $L^2(\Omega)$ to 0, as $n \to +\infty$, if $q \geq 2$ and all k_i's are distinct numbers. This proves our theorem when $f \in \mathscr{E}_p$.

By the above computation, we can see that if $F = I^W_p(f)$ with f given in (4.1), then

$$R_n(F) = \sum_{i_1,\dots,i_p=1}^{m} \beta_{i_1 i_2 \dots i_p} \sum_{\ell=2}^{p} (-1)^\ell \sum_{\substack{k_1,\dots,k_\ell=1 \\ \text{all distinct}}}^{p} \left(\prod_{j \in [p] \setminus \{k_1,\dots,k_\ell\}} W(A_{i_j}) \right) \sum_{v=1}^{n} W(A^v_{i_{k_1}}) \cdots W(A^v_{i_{k_\ell}}).$$

Therefore, using Wiener-Itô isometry, we can first write $\left\|R_n(F)\right\|^2_{L^2(\Omega)}$ as

$$p! \sum_{i_1,\dots,i_p=1}^{m} (\beta_{i_1 i_2 \dots i_p})^2 \sum_{v=1}^{n} \left\| \sum_{\ell=2}^{p} (-1)^\ell \sum_{\substack{k_1,\dots,k_\ell \in [p] \\ \text{all distinct}}} \left(\prod_{j \in [p] \setminus \{k_1,\dots,k_\ell\}} W(A_{i_j}) \right) W(A^v_{i_{k_1}}) \cdots W(A^v_{i_{k_\ell}}) \right\|^2_{L^2(\Omega)} ,$$

and then using the elementary inequality $(a_1 + \ldots + a_m)^\beta \leq m^{\beta-1} \sum_{i=1}^m |a_i|^\beta$ for $a_i \in \mathbb{R}$, $\beta > 1$, $m \in \mathbb{N}$, we have

$$\begin{aligned}
&\left\| \sum_{\ell=2}^{p} (-1)^\ell \sum_{\substack{k_1,\dots,k_\ell \in [p] \\ \text{all distinct}}} \left(\prod_{j \in [p] \setminus \{k_1,\dots,k_\ell\}} W(A_{i_j}) \right) W(A^v_{i_{k_1}}) \cdots W(A^v_{i_{k_\ell}}) \right\|^2_{L^2(\Omega)} \\
&\leq \Theta_1 \sum_{\ell=2}^{p} \sum_{\substack{k_1,\dots,k_\ell \in [p] \\ \text{all distinct}}} \left\| \left(\prod_{j \in [p] \setminus \{k_1,\dots,k_\ell\}} W(A_{i_j}) \right) W(A^v_{i_{k_1}}) \cdots W(A^v_{i_{k_\ell}}) \right\|^2_{L^2(\Omega)} \\
&= \Theta_1 \sum_{\ell=2}^{p} \sum_{\substack{k_1,\dots,k_\ell \in [p] \\ \text{all distinct}}} \left(\prod_{j \in [p] \setminus \{k_1,\dots,k_\ell\}} \nu(A_{i_j}) \right) \nu(A^v_{i_{k_1}}) \cdots \nu(A^v_{i_{k_\ell}}) \\
&\leq \Theta_2 \sum_{\substack{k_1,k_2 \in [p] \\ k_1 \neq k_2}} \left(\prod_{j \in [p] \setminus \{k_1,k_2\}} \nu(A_{i_j}) \right) \nu(A^v_{i_{k_1}})\, \nu(A^v_{i_{k_2}})
\end{aligned}$$

where Θ_1, Θ_2 (and Θ_3 in the following) are some absolute constants that do not depend on n or F. Note now for $k_1 \neq k_2$, $\sum_{v=1}^n \nu(A^v_{i_{k_1}}) \cdot \nu(A^v_{i_{k_2}}) \leq \nu(A_{i_{k_1}}) \sum_{v=1}^n \nu(A^v_{i_{k_2}}) = \nu(A_{i_{k_1}}) \cdot \nu(A_{i_{k_2}})$, thus,

$$\begin{aligned}
\left\|R_n(F)\right\|^2_{L^2(\Omega)} &\leq p! \sum_{i_1,\dots,i_p=1}^{m} (\beta_{i_1 i_2 \dots i_p})^2 \Theta_2 \sum_{\substack{k_1,k_2 \in [p] \\ k_1 \neq k_2}} \left(\prod_{j \in [p] \setminus \{k_1,k_2\}} \nu(A_{i_j}) \right) \nu(A_{i_{k_1}}) \nu(A_{i_{k_2}}) \\
&\leq p! \sum_{i_1,\dots,i_p=1}^{m} (\beta_{i_1 i_2 \dots i_p})^2 \Theta_3 \prod_{j \in [p]} \nu(A_{i_j}) = \Theta_3 \cdot \|F\|^2_{L^2(\Omega)} \,.
\end{aligned}$$

Since $\{I_p^W(f) : f \in \mathscr{E}_p\}$ is dense in the pth Wiener chaos $\mathscr{H}_p$, $R_n : \mathscr{H}_p \to L^2(\Omega)$ is a bounded linear operator with operator norm $\|R_n\|_{\text{op}} \leq \sqrt{\Theta_3}$ for each $n \in \mathbb{N}$. Note the linearity follows from its definition $R_n(F) := n\,E[F^{(n)} - F \big| \mathscr{G}] + pF$, $F \in \mathscr{H}_p$.

Now we define

$$\mathscr{C}_p := \left\{ F \in \mathscr{H}_p \ : \ R_\infty(F) := \lim_{n\to+\infty} R_n(F) \quad \text{is well defined in } L^2(\Omega) \right\} .$$

It is easy to see that $\mathscr{C}_p$ is a dense linear subspace of $\mathscr{H}_p$ and for each $f \in \mathscr{E}_p$, $I_p^W(f) \in \mathscr{C}_p$ and $R_\infty(I_p^W(f)) = 0$. As

$$\sup_{n\in\mathbb{N}} \|R_n\|_{\text{op}} \leq \sqrt{\Theta_3} < +\infty ,$$

R_∞ has a unique extension to $\mathscr{H}_p$ and by density of $\{I_p^W(f) : f \in \mathscr{E}_p\}$ in $\mathscr{H}_p$, $R_\infty(F) = 0$ for each $F \in \mathscr{H}_p$. In other words, for any $F \in \mathscr{H}_p$, $n\,E[F^{(n)} - F \big| \mathscr{G}]$ converges in $L^2(\Omega)$ to $-pF$, as $n \to +\infty$. □

5 Proof of Proposition 1.1

We now give the proof of Proposition 1.1, which has been stated in the introduction. We restate it for the convenience of the reader.

Proposition 1.1 Let $(B, B^t)_{t\geq 0}$ be a family of exchangeable pairs of Brownian motions (that is, B is a Brownian motion on $[0,1]$ and, for each t, one has $(B, B^t) \overset{\text{law}}{=} (B^t, B)$). Assume moreover that

(a) for any integer $p \geq 1$ and any $f \in L^2([0,1]^p)$,

$$\lim_{t\downarrow 0} \frac{1}{t}\, E\Big[I_p^{B^t}(f) - I_p^B(f)\Big|\sigma\{B\}\Big] = -p\, I_p^B(f) \quad \text{in } L^2(\Omega).$$

Then, for any integer $p \geq 1$ and any $f \in L^2([0,1]^p)$,

(b) $\displaystyle \lim_{t\downarrow 0} \frac{1}{t}\, E\Big[(I_p^{B^t}(f) - I_p^B(f))^2 |\sigma\{B\}\Big] = 2p^2 \int_0^1 I_{p-1}^B(f(x,\cdot))^2 dx \quad \text{in } L^2(\Omega);$

(c) $\displaystyle \lim_{t\downarrow 0} \frac{1}{t}\, E\Big[(I_p^{B^t}(f) - I_p^B(f))^4\Big] = 0.$

Proof. We first concentrate on the proof of (b). Fix $p \geq 1$ and $f \in L^2([0,1]^p)$, and set $F = I_p^B(f)$ and $F_t = I_p^{B^t}(f)$. First, we observe that

$$\frac{1}{t}\, E[(F_t - F)^2 \big| \sigma\{B\}] = \frac{1}{t}\, E[F_t^2 - F^2 \big| \sigma\{B\}] - \frac{2}{t} F\, E[F_t - F \big| \sigma\{B\}].$$

Also, as an immediate consequence of the product formula (2.2) and the definition of $f \otimes_r f$, we have

$$p^2 \int_0^1 I_{p-1}^B(f(x,\cdot))^2 dx = \sum_{r=1}^p r r! \binom{p}{r}^2 I_{2p-2r}^B(f \otimes_r f).$$

Given (a) and the previous two identities, in order to prove (b) we are thus left to check that

$$\lim_{t\downarrow 0}\frac{1}{t}E[F_t^2-F^2\big|\sigma\{B\}]=-2p\,F^2+2\sum_{r=1}^{p}rr!\binom{p}{r}^2 I_{2p-2r}^{B}(f\otimes_r f)\quad\text{in } L^2(\Omega).\tag{5.1}$$

The product formula (2.2) used for multiple integrals with respect to B^t (resp. B) yields

$$F_t^2=\sum_{r=0}^{p}r!\binom{p}{r}^2 I_{2p-2r}^{B^t}(f\otimes_r f)\quad\Big(\text{resp. } F^2=\sum_{r=0}^{p}r!\binom{p}{r}^2 I_{2p-2r}^{B}(f\otimes_r f)\Big).$$

Hence it follows from (a) that

$$\begin{aligned}\frac{1}{t}E[F_t^2-F^2\big|\sigma\{B\}] &= \sum_{r=0}^{p-1}r!\binom{p}{r}^2\frac{1}{t}E[I_{2p-2r}^{B^t}(f\otimes_r f)-I_{2p-2r}^{B}(f\otimes_r f)|\sigma\{B\}]\\ &\longrightarrow \sum_{r=0}^{p-1}r!\binom{p}{r}^2(2r-2p)I_{2p-2r}^{B}(f\otimes_r f)\\ &= -2p(F^2-E[F^2])+2\sum_{r=1}^{p-1}rr!\binom{p}{r}^2 I_{2p-2r}^{B}(f\otimes_r f),\end{aligned}$$

which is exactly (5.1). The proof of (b) is complete.

Let us now turn to the proof of (c). Fix $p\geq 1$ and $f\in L^2([0,1]^p)$, and set $F=I_p^B(f)$ and $F_t=I_p^{B^t}(f)$, $t\geq 0$. We claim that the pair (F,F_t) is exchangeable for each t. Indeed, thanks to point 4 in Section 2, we first observe that it is enough to check this claim when f belongs to $\mathcal{E}_p$, that is, when f has the form

$$f(x_1,\ldots,x_p)=\sum_{i_1,\ldots,i_p=1}^{k}\beta_{i_1\ldots i_p}\mathbf{I}_{[\tau_{i_1-1},\tau_{i_1})\times\ldots\times[\tau_{i_p-1},\tau_{i_p})}(x_1,\ldots,x_p),$$

with $k\geq 1$ and $0=\tau_0<\tau_1<\ldots<\tau_k$, and the coefficients $\beta_{i_1\ldots i_p}$ are zero if any two of the indices $i_1,\ldots,i_p$ are equal. But, for such an f, one has

$$\begin{aligned}F=I_p^B(f)&=\sum_{i_1,\ldots,i_p=1}^{k}\beta_{i_1\ldots i_p}(B(\tau_{i_1})-B(\tau_{i_1-1}))\ldots(B(\tau_{i_p})-B(\tau_{i_p-1}))\\ F_t=I_p^{B^t}(f)&=\sum_{i_1,\ldots,i_p=1}^{k}\beta_{i_1\ldots i_p}(B^t(\tau_{i_1})-B^t(\tau_{i_1-1}))\ldots(B^t(\tau_{i_p})-B^t(\tau_{i_p-1})),\end{aligned}$$

and the exchangeability of (F,F_t) follows immediately from those of (B,B^t). Since the pair (F,F_t) is exchangeable, we can write

$$\begin{aligned}E[(F_t-F)^4]&=E[F_t^4+F^4-4F_t^3F-4F^3F_t+6F_t^2F^2]\\ &=2E[F^4]-8E[F^3F_t]+6E[F^2F_t^2]\quad\text{by exchangeability;}\\ &=4E[F^3(F_t-F)]+6E[F^2(F_t-F)^2]\quad\text{after rearrangement;}\end{aligned}$$

$$= 4E[F^3 E[(F_t - F)|\sigma\{B\}]] + 6E[F^2 E[(F_t - F)^2|\sigma\{B\}]].$$

Dividing by t and taking the limit $t \downarrow 0$ into the previous identity, we deduce, thanks to (a) and (b) as well, that

$$\lim_{t\downarrow 0} \frac{1}{t} E\Big[(F_t - F)^4\Big] = -4pE[F^4] + 12p^2 E\left[F^2 \int_0^1 I_{p-1}^B(f(x,\cdot))^2 dx\right]. \tag{5.2}$$

In particular, it appears that the limit of $\frac{1}{t} E[(F_t - F)^4]$ is always the same, irrespective of the choice of our exchangeable pair of Brownian motions (B, B^t) satisfying (a). To compute it, we can then choose the pair (B, B^t) we want, for instance, the pair constructed in Section 3. This is why, starting from now and for the rest of the proof, (B, B^t) refers to the pair defined in Section 3 (which satisfies (a), that is, (3.2)). What we gain by considering this particular pair is that it satisfies a hypercontractivity-type inequality. More precisely, there exists $c_p > 0$ (only depending on p) such that, for all $t \geq 0$,

$$E[(F_t - F)^4] \leq c_p E[(F_t - F)^2]^2. \tag{5.3}$$

Indeed, going back to the definition of multiple Wiener-Itô integrals as given in Section 2 (first for elementary functions and then by approximation for the general case), we see that $F_t - F$ is a multiple Wiener-Itô integral of order p with respect to the *two-sided* Brownian motion $\overline{B} = (\overline{B}(s))_{s\in[-1,1]}$, defined as

$$\overline{B}(s) = B(s)\mathbf{I}_{[0,1]}(s) + \widehat{B}(-s)\mathbf{I}_{[-1,0]}(s).$$

But product formula (2.2) is also true for a two-sided Brownian motion, so the claim (5.3) follows from (2.3) applied to $\overline{B}$. On the other hand, it follows from (b) that $\frac{1}{t} E[(F_t - F)^2]$ converges to a finite number, as $t \downarrow 0$. Hence, combining this fact with (5.3) yields

$$\frac{1}{t} E\Big[(F_t - F)^4\Big] \leq c_p\, t \left(\frac{1}{t} E\Big[(F_t - F)^2\Big]\right)^2 \to 0\,,$$

as $t \downarrow 0$. □

Remark 5.1. (i) A byproduct of (5.2) in the previous proof is that

$$\frac{1}{3}(E[F^4] - 3\sigma^4) = E\left[F^2\left(p\int_0^1 I_{p-1}^B(f(x,\cdot))^2 dx - \sigma^2\right)\right]. \tag{5.4}$$

Note (5.4) was originally obtained by chain rule, see [15, equation (5.2.9)].

(ii) As a consequence of (c) in Proposition 1.1 , we have $\lim_{t\downarrow 0} \frac{1}{t} E[|I_p^{B^t}(f) - I_p^B(f)|^3] = 0$. Indeed,

$$\frac{1}{t} E\Big[|I_p^{B^t}(f) - I_p^B(f)|^3\Big] \leq \left(\frac{1}{t} E\Big[(I_p^{B^t}(f) - I_p^B(f))^2\Big]\right)^{\frac{1}{2}} \left(\frac{1}{t} E\Big[(I_p^{B^t}(f) - I_p^B(f))^4\Big]\right)^{\frac{1}{2}}$$
$$\to 0\,, \quad \text{as } t \downarrow 0.$$

(iii) For any $r > 2$, in view of (2.4) and (5.3), there exists an absolute constant $c_{r,p}$ depending only on p, r (but not on f) such that

$$E[|I_p^B(f) - I_p^{B^t}(f)|^r] \le c_{r,p}\, E[(I_p^B(f) - I_p^{B^t}(f))^2]^{r/2} .$$

Moreover, if $F \in L^2(\Omega, \sigma\{B\}, P)$ admits a *finite* chaos expansion, say, (for some $p \in \mathbb{N}$) $F = E[F] + \sum_{q=1}^p I_q^B(f_q)$, and we set $F_t = E[F] + \sum_{q=1}^p I_q^{B^t}(f_q)$, then there exists some absolute constant $C_{r,p}$ that only depends on p and r such that

$$E[|F - F_t|^r] \le C_{r,p}\, E[(F - F_t)^2]^{r/2} .$$

6 Proof of E. Meckes' Theorem 1.2

In this section, for sake of completeness and because our version slightly differs from the original one given in [11, Theorem 2.1], we provide a proof of Theorem 1.2, which we restate here for convenience.

Theorem 1.2 Let F and a family of random variables $(F_t)_{t\ge 0}$ be defined on a common probability space $(\Omega, \mathcal{F}, P)$ such that $F_t \overset{law}{=} F$ for every $t \ge 0$. Assume that $F \in L^3(\Omega, \mathcal{G}, P)$ for some σ-algebra $\mathcal{G} \subset \mathcal{F}$ and that in $L^1(\Omega)$,

(a) $\lim_{t\downarrow 0} \frac{1}{t} E[F_t - F|\mathcal{G}] = -\lambda F$ for some $\lambda > 0$,

(b) $\lim_{t\downarrow 0} \frac{1}{t} E[(F_t - F)^2|\mathcal{G}] = (2\lambda + S)\mathrm{Var}(F)$ for some random variable S,

(c) $\lim_{t\downarrow 0} \frac{1}{t} (F_t - F)^3 = 0$.

Then, with $N \sim N(0, \mathrm{Var}(F))$,

$$d_{TV}(F, N) \le \frac{E|S|}{\lambda}.$$

Proof. Without loss of generality, we may and will assume that $\mathrm{Var}(F) = 1$. It is known that

$$d_{TV}(F, N) = \frac{1}{2} \sup E[\varphi(F) - \varphi(N)] , \tag{6.1}$$

where the supremum runs over all smooth functions $\varphi : \mathbb{R} \to \mathbb{R}$ with compact support and such that $\|\varphi\|_\infty \le 1$. For such a φ, recall (see, e.g. [3, Lemma 2.4]) that

$$g(x) = e^{x^2/2} \int_{-\infty}^{x} (\varphi(y) - E[\varphi(N)])e^{-y^2/2}\, dy , \quad x \in \mathbb{R},$$

satisfies

$$g'(x) - xg(x) = \varphi(x) - E[\varphi(N)] \tag{6.2}$$

as well as $\|g\|_\infty \leq \sqrt{2\pi}$, $\|g'\|_\infty \leq 4$ and $\|g''\|_\infty \leq 2\|\varphi'\|_\infty < +\infty$. In what follows, we fix such a pair (φ, g) of functions. Let G be a differentiable function such that $G' = g$, then due to $F_t \overset{\text{law}}{=} F$, it follows from the Taylor formula in mean-value form that

$$0 = E[G(F_t) - G(F)] = E[g(F)(F_t - F)] + \frac{1}{2}E[g'(F)(F_t - F)^2] + E[R] \ ,$$

with remainder R bounded by $\frac{1}{6}\|g''\|_\infty |F_t - F|^3$.

By assumption (c) and as $t \downarrow 0$,

$$\left|\frac{1}{t}E[R]\right| \leq \frac{1}{6}\|g''\|_\infty \frac{1}{t}E[|F_t - F|^3] \to 0.$$

Therefore as $t \downarrow 0$, assumptions (a) and (b) imply that

$$\lambda\, E[g'(F) - Fg(F)] + \frac{1}{2}E[g'(F)S] = 0.$$

Plugging this into Stein's equation (6.2) and then using (6.1), we deduce the desired conclusion, namely,

$$d_{TV}(F, N) \leq \frac{1}{2}\frac{\|g'\|_\infty}{2\lambda}E|S| \leq \frac{E|S|}{\lambda}.$$

□

Remark 6.1. Unlike the original Meckes' theorem, we do not assume the exchangeability condition $(F_t, F) \overset{law}{=} (F, F_t)$ in our Theorem 1.2. Our consideration is motivated by [22].

7 Quantitative fourth moment theorem revisited via exchangeable pairs

We give an elementary proof to the quantitative fourth moment theorem, that is, we explain how to prove the inequality of Theorem 1.1(ii) by means of our exchangeable pairs approach. For sake of convenience, let us restate this inequality: for any multiple Wiener-Itô integral F of order $p \geq 1$ such that $E[F^2] = \sigma^2 > 0$, we have, with $N \sim N(0, \sigma^2)$,

$$d_{TV}(F, N) \leq \frac{2}{\sigma^2}\sqrt{\frac{p-1}{3p}}\,\sqrt{E[F^4] - 3\sigma^4}. \tag{7.1}$$

To prove (7.1), we consider, for instance, the exchangeable pairs of Brownian motions $\{(B, B^t)\}_{t>0}$ constructed in Section 3. We deduce, by combining Proposition 1.1 with Theorem 1.2 and Remark 5.1-(ii), that

$$d_{TV}(F, N) \leq \frac{2}{\sigma^2}E\left[\left|p\int_0^1 I^B_{p-1}(f(x, \cdot))^2 dx - \sigma^2\right|\right]. \tag{7.2}$$

To deduce (7.1) from (7.2), we are thus left to prove the following result.

Proposition 7.1. Let $p \geq 1$ and consider a symmetric function $f \in L^2([0,1]^p)$. Set $F = I_p^B(f)$ and $\sigma^2 = E[F^2]$. Then

$$E\left[\left(p\int_0^1 I_{p-1}^B(f(x,\cdot))^2dx - \sigma^2\right)^2\right] \leq \frac{p-1}{3p}(E[F^4] - 3\sigma^4).$$

Proof. Using the product formula (2.2), we can write

$$F^2 = \sum_{r=0}^{p} r!\binom{p}{r}^2 I_{2p-2r}^B(f\otimes_r f) = \sigma^2 + \sum_{r=0}^{p-1} r!\binom{p}{r}^2 I_{2p-2r}^B(f\otimes_r f),$$

as well as

$$p\int_0^1 I_{p-1}^B(f(x,\cdot))^2dx = p\sum_{r=0}^{p-1} r!\binom{p-1}{r}^2 I_{2p-2r-2}^B\left(\int_0^1 f(x,\cdot)\otimes_r f(x,\cdot)dx\right)$$
$$= p\sum_{r=1}^{p}(r-1)!\binom{p-1}{r-1}^2 I_{2p-2r}^B\,(f\otimes_r f) = \sigma^2 + \sum_{r=1}^{p-1}\frac{r}{p}\,r!\binom{p}{r}^2 I_{2p-2r}^B\,(f\otimes_r f)\,.$$

Hence, by the isometry property (point 2 in Section 2),

$$E\left[\left(p\int_0^1 I_{p-1}^B(f(x,\cdot))^2dx - \sigma^2\right)^2\right] = \sum_{r=1}^{p-1}\frac{r^2}{p^2}\,r!^2\binom{p}{r}^4(2p-2r)!\|f\widetilde{\otimes}_r f\|^2_{L^2([0,1]^{2p-2r})}.$$

On the other hand, one has from (5.4) and the isometry property again that

$$\frac{1}{3}(E[F^4] - 3\sigma^4) = E\left[F^2\left(p\int_0^1 I_{p-1}^B(f(x,\cdot))^2dx - \sigma^2\right)\right]$$
$$= \frac{1}{3}(E[F^4] - 3\sigma^4) = \sum_{r=1}^{p-1}\frac{r}{p}\,r!^2\binom{p}{r}^4(2p-2r)!\|f\widetilde{\otimes}_r f\|^2_{L^2([0,1]^{2p-2r})}.$$

The desired conclusion follows. □

8 Connections with Malliavin operators

Our main goal in this paper is to provide an elementary proof of Theorem 1.1(ii). Nevertheless, in this section we further investigate the connections we have found between our exchangeable pair approach and the operators of Malliavin calculus. This part may be skipped in a first reading, as it is not used in other sections. It is directed to readers who are already familiar with Malliavin calculus. We use classical notation and so do not introduce them in order to save place. We refer to [18] for any missing detail.

In this section, to stay on the safe side we only consider random variables F belonging to

(8.1) $$\mathcal{A} := \bigcup_{p\in\mathbb{N}}\bigoplus_{r\leq p}\mathscr{H}_r\,,$$

where $\mathscr{H}_r$ is the rth chaos associated to the Brownian motion B. In other words, we only consider random variables that are $\sigma\{B\}$-measurable and that admit a *finite* chaotic expansion. Note that $\mathscr{A}$ is an algebra (in view of product formula) that is dense in $L^2(\Omega, \sigma\{B\}, P)$.

As is well-known, any $\sigma\{B\}$-measurable random variable F can be written $F = \psi_F(B)$ for some measurable mapping $\psi_F : \mathbb{R}^{\mathbb{R}_+} \to \mathbb{R}$ determined $P \circ B^{-1}$ almost surely. For such an F, we can then define $F_t = \psi_F(B^t)$, with B^t defined in Section 3. Another equivalent description of F_t is to define it as $F_t = E[F] + \sum_{r=1}^{p} I_r^{B^t}(f_r)$, if the family $(f_r)_{1\leq r\leq p}$ is such that $F = E[F] + \sum_{r=1}^{p} I_r^{B}(f_r)$.

Our main findings are summarized in the statement below.

Proposition 8.1. Consider $F, G \in \mathscr{A}$, and define F_t, G_t for each $t \in \mathbb{R}_+$ as is done above. Then, in $L^2(\Omega)$,

(a) $\displaystyle \lim_{t\downarrow 0} \frac{1}{t} E\Big[F_t - F\Big|\sigma\{B\}\Big] = LF,$

(b) $\displaystyle \lim_{t\downarrow 0} \frac{1}{t} E\Big[(F_t - F)(G_t - G)|\sigma\{B\}\Big] = L(FG) - FLG - GLF = 2\,\langle DF, DG\rangle.$

Proof. The proof of (a) is an immediate consequence of (3.2), the linearity of conditional expectation, and the fact that $LI_r^B(f_r) = -r\,I_r^B(f_r)$ by definition of L. Let us now turn to the proof of (b). Using elementary algebra and then (a), we deduce that, as $t \downarrow 0$ and in $L^2(\Omega)$,

$$\begin{aligned}
&\frac{1}{t} E[(F_t - F)(G_t - G)\big|\sigma\{B\}] \\
&= \frac{1}{t} E[F_tG_t - FG\big|\sigma\{B\}] - \frac{1}{t}F\, E[G_t - G\big|\sigma\{W\}] - \frac{1}{t}G\, E[F_t - F\big|\sigma\{B\}] \\
&\to L(FG) - FLG - GLF\ .
\end{aligned}$$

Using $L = -\delta D$, $D(FG) = FDG + GDF$ (Leibniz rule) and $\delta(FDG) = F\delta(DG) - \langle DF, DG\rangle$ (see [18, Proposition 1.3.3]), it is easy to check that $L(FG) - FLG - GLF = 2\langle DF, DG\rangle$, which concludes the proof of Proposition 8.1. □

Remark 8.1. The expression appearing in the right-hand side of (b) is nothing else but $2\,\Gamma(F, G)$, the (doubled) carré du champ operator.

To conclude this section, we show how our approach allows to recover the diffusion property of the Ornstein-Uhlenbeck operator.

Proposition 8.2. Fix $d \in \mathbb{N}$, let $F = (F_1, \dots, F_d) \in \mathscr{A}^d$ (with $\mathscr{A}$ given in (8.1)), and $\Psi : \mathbb{R}^d \to \mathbb{R}$ be a polynomial function. Then

$$L\Psi(F) = \sum_{j=1}^{d} \partial_j \Psi(F) LF_j + \sum_{i,j=1}^{d} \partial_{ij}\Psi(F)\langle DF_i, DF_j\rangle\ . \tag{8.2}$$

Proof. We first define $F_t = (F_{1,t}, \dots, F_{d,t})$ as explained in the beginning of the present section. Using classical multi-index notations, Taylor formula yields that

$$\Psi(F_t) - \Psi(F) = \sum_{j=1}^{d} \partial_j \Psi(F)(F_{j,t} - F_j) + \frac{1}{2}\sum_{i,j=1}^{d} \partial_{i,j}\Psi(F)(F_{j,t} - F_j)(F_{i,t} - F_i)$$

$$+\sum_{|\beta|=3}\frac{3}{\beta_1!\ldots\beta_d!}(F_t-F)^\beta\int_0^1(1-s)^k(\partial_1^{\beta_1}\ldots\partial_d^{\beta_d}\Psi)(F+s(F_t-F))\,ds\,. \tag{8.3}$$

In view of the previous proposition, the only difficulty in establishing (8.2) is about controlling the last term in (8.3) while passing $t\downarrow 0$. Note first $(\partial_1^{\beta_1}\ldots\partial_d^{\beta_d}\Psi)(F+s(F_t-F))$ is polynomial in F and (F_t-F), so our problem reduces to show

$$\lim_{t\downarrow 0}\frac{1}{t}E[|F^\alpha(F_t-F)^\beta|]=0\,, \tag{8.4}$$

for $\alpha=(\alpha_1,\ldots,\alpha_d),\beta=(\beta_1,\ldots,\beta_d)\in(\mathbb{N}\cup\{0\})^d$ with $|\beta|\geq 3$.

Indeed, (assume $\beta_j>0$ for each j)

$$\begin{aligned}\frac{1}{t}E[|F^\alpha(F_t-F)^\beta|]&\leq\frac{1}{t}E[|F^\alpha|^2]^{1/2}E[|(F_t-F)^\beta|^2]^{1/2}\quad\text{by Cauchy-Schwarz inequality;}\\&\leq E[|F^\alpha|^2]^{1/2}\,\frac{1}{t}\left(\prod_{j=1}^d E\Big[(F_{j,t}-F_j)^{2|\beta|}\Big]^{\frac{\beta_j}{|\beta|}}\right)^{1/2}\quad\text{by Hölder inequality;}\\&\leq C\,E[|F^\alpha|^2]^{1/2}\,t^{\frac{|\beta|}{2}-1}\left(\prod_{j=1}^d\frac{1}{t^{\beta_j}}E\Big[(F_{j,t}-F_j)^2\Big]^{\beta_j}\right)^{1/2}\,,\end{aligned}$$

where the last inequality follows from point-(iii) in Remark 5.1 with $C>0$ independent of t. Since $F^\alpha\in\mathcal{A}$ and $|\beta|\geq 3$, (8.4) follows immediately from the above inequalities. □

9 Peccati-Tudor theorem revisited too

In this section, we combine a multivariate version of Meckes' abstract exchangeable pairs [12] with our results from Section 3 to prove (1.5), thus leading to a fully elementary proof of Theorem 1.4 as well.

First, we recall the following multivariate version of Meckes' theorem (see [12, Theorem 4]). Unlike in the one-dimensional case, it seems inevitable to impose the exchangeability condition in the following proposition, as we read from its proof in [12].

Proposition 9.1. For each $t>0$, let (F,F_t) be an exchangeable pair of centered d-dimensional random vectors defined on a common probability space. Let $\mathcal{G}$ be a σ-algebra that contains $\sigma\{F\}$. Assume that $\Lambda\in\mathbb{R}^{d\times d}$ is an invertible deterministic matrix and Σ is a symmetric, non-negative definite deterministic matrix such that

(a) $\lim_{t\downarrow 0}\frac{1}{t}\,E[F_t-F|\mathcal{G}]=-\Lambda F$ in $L^1(\Omega)$,

(b) $\lim_{t\downarrow 0}\frac{1}{t}\,E[(F_t-F)(F_t-F)^T|\mathcal{G}]=2\Lambda\Sigma+S$ in $L^1(\Omega,\|\cdot\|_{\mathrm{HS}})$ for some matrix $S=S(F)$, and with $\|\cdot\|_{\mathrm{HS}}$ the Hilbert-Schmidt norm

(c) $\lim_{t\downarrow 0}\sum_{i=1}^{d}\frac{1}{t}\,E[|F_{i,t}-F_i|^3]=0$, where $F_{i,t}$ (resp. F_i) stands for the ith coordinate of F_t (resp. F).

Then, with $N\sim N_d(0,\Sigma)$,

(1) for $g\in C^2(\mathbb{R}^d)$,

$$\left|E[g(F)]-E[g(N)]\right|\leq\frac{\|\Lambda^{-1}\|_{\mathrm{op}}\sqrt{d}\,M_2(g)}{4}E\left[\sqrt{\sum_{i,j=1}^{d}S_{ij}^2}\right],$$

where $M_2(g):=\sup_{x\in\mathbb{R}^d}\left\|D^2g(x)\right\|_{\mathrm{op}}$ with $\|\cdot\|_{\mathrm{op}}$ the operator norm.

(2) if, in addition, Σ is positive definite, then

$$d_{\mathrm{W}}(F,N)\leq\frac{\|\Lambda^{-1}\|_{\mathrm{op}}\|\Sigma^{-1/2}\|_{\mathrm{op}}}{\sqrt{2\pi}}\,E\left[\sqrt{\sum_{i,j=1}^{d}S_{ij}^2}\right].$$

Remark 9.1. Constant in (2) is different from Meckes' paper [12] . We took this better constant from Christian Döbler's dissertation [5], see page 114 therein.

By combining the previous proposition with our exchangeable pairs, we get the following result, whose point 2 corresponds to (1.5).

Theorem 9.1. Fix $d\geq 2$ and $1\leq p_1\leq\ldots\leq p_d$. Consider a vector $F:=(I^B_{p_1}(f_1),\ldots,I^B_{p_d}(f_d))$ with $f_i\in L^2([0,1]^{p_i})$ symmetric for each $i\in\{1,\ldots,d\}$. Let $\Sigma=(\sigma_{ij})$ be the covariance matrix of F, and $N\sim N_d(0,\Sigma)$. Then

(1) for $g\in C^2(\mathbb{R}^d)$,

$$\left|E[g(F)]-E[g(N)]\right|\leq\frac{\sqrt{d}\,M_2(g)}{2p_1}\sqrt{\sum_{i,j=1}^{d}\mathrm{Var}\Big(p_ip_j\int_0^1 I_{p_i-1}(f_i(x,\cdot))I_{p_j-1}(f_j(x,\cdot))dx\Big)},$$

where $M_2(g):=\sup_{x\in\mathbb{R}^d}\left\|D^2g(x)\right\|_{\mathrm{op}}$.

(2) if in addition, Σ is positive definite, then

$$d_{\mathrm{W}}(F,N)\leq\frac{2\|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1\sqrt{2\pi}}\sqrt{\sum_{i,j=1}^{d}\mathrm{Var}\Big(p_ip_j\int_0^1 I_{p_i-1}(f_i(x,\cdot))I_{p_j-1}(f_j(x,\cdot))dx\Big)}.$$

Proof. We consider $F_t=(I^{B^t}_{p_1}(f_1),\ldots,I^{B^t}_{p_d}(f_d))$, where B^t is the Brownian motion constructed in Section 3. We deduce from (3.1) that

$$\frac{1}{t}\,E[F_t-F|\sigma\{B\}]=\left(\frac{e^{-p_1t}-1}{t}\,I^{B^t}_{p_1}(f_1),\ldots,\frac{e^{-p_dt}-1}{t}\,I^{B^t}_{p_d}(f_d)\right)$$

implying in turn that, in $L^2(\Omega)$ and as $t \downarrow 0$,

$$\frac{1}{t}\, E[F_t - F|\sigma\{B\}] \to -\Lambda F,$$

with $\Lambda = \mathrm{diag}(p_1, \ldots, p_d)$ (in particular, $\|\Lambda^{-1}\|_{\mathrm{op}} = p_1^{-1}$). That is, assumption (a) in Proposition 9.1 is satisfied (with $\mathcal{G} = \sigma\{B\}$). That assumption (c) in Proposition 9.1 is satisfied as well follows from Proposition 1.1(c). Let us finally check that assumption (b) in Proposition 9.1 takes place too. First, using the product formula (2.2) for multiple integrals with respect to B^t (resp. B) yields

$$\begin{aligned}
F_i F_j &= \sum_{r=0}^{p_i \wedge p_j} r! \binom{p_i}{r}\binom{p_j}{r} I^B_{p_i+p_j-2r}(f_i \otimes_r f_j)\\
F_{i,t} F_{j,t} &= \sum_{r=0}^{p_i \wedge p_j} r! \binom{p_i}{r}\binom{p_j}{r} I^{B^t}_{p_i+p_j-2r}(f_i \otimes_r f_j).
\end{aligned}$$

Hence, using (3.2) for passing to the limit,

$$\begin{aligned}
&\frac{1}{t}\, E\big[(F_{i,t} - F_i)(F_{j,t} - F_j)\big|\sigma\{B\}\big] - \frac{1}{t}\, E\big[F_{i,t}F_{j,t} - F_iF_j\big|\sigma\{B\}\big]\\
&= -\frac{1}{t}\, F_i\, E\big[F_{j,t} - F_j|\sigma\{B\}\big] - \frac{1}{t}\, F_j\, E\big[F_{i,t} - F_i\big|\sigma\{B\}\big]\\
&\to (p_i + p_j)F_iF_j = \sum_{r=0}^{p_i \wedge p_j} r! \binom{p_i}{r}\binom{p_j}{r}(p+q) I^B_{p_i+p_j-2r}(f_i \otimes_r f_j) \quad \text{as } t \downarrow 0.
\end{aligned}$$

Now, note in $L^2(\Omega)$,

$$\begin{aligned}
&\frac{1}{t}\, E\big[F_{i,t}F_{j,t} - F_iF_j\big|\sigma\{B\}\big]\\
&= \sum_{r=0}^{p_i \wedge p_j} r! \binom{p_i}{r}\binom{p_j}{r} \frac{1}{t}\, E\Big[I^{B^t}_{p_i+p_j-2r}(f_i \otimes_r f_j) - I^B_{p_i+p_j-2r}(f_i \otimes_r f_j)\Big|\sigma\{B\}\Big]\\
&\to \sum_{r=0}^{p_i \wedge p_j} r! \binom{p_i}{r}\binom{p_j}{r}(2r - p_i - p_j) I^B_{p_i+p_j-2r}(f_i \otimes_r f_j)\ , \quad \text{as } t \downarrow 0, \text{ by (3.2)}\ .
\end{aligned}$$

Thus, as $t \downarrow 0$,

$$\begin{aligned}
\frac{1}{t}\, E\big[(F_{i,t} - F_i)(F_{j,t} - F_j)\big|\sigma\{B\}\big] &\to 2\sum_{r=1}^{p_i \wedge p_j} r!r \binom{p_i}{r}\binom{p_j}{r} I^B_{p_i+p_j-2r}(f_i \otimes_r f_j)\\
&= 2p_ip_j \int_0^1 I^B_{p_i-1}(f_i(x,\cdot)) I^B_{p_j-1}(f_j(x,\cdot))dx\,,
\end{aligned}$$

where the last equality follows from a straightforward application of the product formula (2.2). As a result, if we set

$$S_{ij} = 2p_ip_j \int_0^1 I_{p_i-1}(f_i(x,\cdot)) I_{p_j-1}(f_j(x,\cdot))dx - 2p_i\sigma_{ij}$$

for each $i, j \in \{1, \ldots, d\}$, then assumption (b) in Proposition 9.1 turns out to be satisfied as well. By the isometry property (point 2 in Section 2), it is straightforward to check that

$$p_j \int_0^1 E\Big[I_{p_i-1}(f_i(x,\cdot))I_{p_j-1}(f_j(x,\cdot))\Big]dx = \sigma_{ij} \ .$$

Therefore,

$$E\left[\sqrt{\sum_{i,j=1}^{d} S_{ij}^2}\right] \leq \sqrt{\sum_{i,j=1}^{d} E[S_{ij}^2]} = 2\sqrt{\sum_{i,j=1}^{d} \mathrm{Var}\Big(p_i p_j \int_0^1 I_{p_i-1}(f_i(x,\cdot))I_{p_j-1}(f_j(x,\cdot))dx\Big)} \ .$$

Hence the desired results in (1) and (2) follow from Proposition 9.1. □

References

[1] J.-M. Azaïs, F. Dalmao and J.R. León. CLT for the zeros of classical random trigonometric polynomials, Ann. Inst. H. Poincaré Probab. Statist. Volume 52, Number 2 (2016), 804-820.

[2] J.-M. Azaïs and J.R. León. CLT for crossing of random trigonometric polynomials, Electron. J. Probab. 18, no. 68 (2013), 1-17.

[3] L.H.Y. Chen, L. Goldstein and Q.M. Shao. Normal approximation by Stein's method. *Probability and Its Applications*, 2011 Springer-Verlag Berlin Heidelberg.

[4] F. Dalmao. Asymptotic variance and CLT for the number of zeros of Kostlan Shub Smale random polynomials, Comptes Rendus Mathematique, Volume 353, Issue 12 (2015), Pages 1141-1145

[5] C. Döbler. New developments in Stein's method with applications. Ph.D dissertation, Ruhr-Universität Bochum (2012). Available online at https://tinyurl.com/y834hxx7.

[6] D. Hug, G. Last, and M. Schulte. Second-order properties and central limit theorems for geometric functionals of Boolean models, Ann. Appl. Probab. Volume 26, Number 1 (2016), 73-135.

[7] P. Malliavin. Stochastic calculus of variations and hypoelliptic operators. In: Proc. Inter. Symp. on Stoch. Diff. Equations, Kyoto 1976, Wiley 1978, 195-263.

[8] D. Marinucci and G. Peccati. Ergodicity and Gaussianity for Spherical Random Fields, J. Math. Phys. 51, (2010)

[9] D. Marinucci, G. Peccati, M. Rossi, I. Wigman. Non-Universality of Nodal Length Distribution for Arithmetic Random Waves, Geom. Funct. Anal. (2016).

[10] D. Marinucci, M. Rossi. Stein-Malliavin approximations for nonlinear functionals of random eigenfunctions on $\mathbf{S}^d$, J. Funct. Anal. 268 (8) (2015), 2379-2420.

[11] E. Meckes. An Infinitesimal Version of Stein's Method of Exchangeable Pairs. Ph.D dissertation, Stanford University (2006)

[12] E. Meckes. On Stein's method for multivariate normal approximation. IMS collections, High dimensional Probability V: The Luminy Volume, Vol. 5 (2009) 153-178.

[13] I. Nourdin. Malliavin-Stein approach: a webpage maintained by Ivan Nourdin. `http://tinyurl.com/kvpdgcy`

[14] I. Nourdin and G. Peccati. Stein's method on Wiener chaos, Probab. Theory Relat. Fields (2009), Vol. 145, Issue 1, p. 75-118.

[15] I. Nourdin and G. Peccati. Normal approximations with Malliavin calculus: from Stein's method to universality, *Cambridge tracts in Mathematics*, Vol. 192, 2012, Cambridge University Press.

[16] I. Nourdin, G. Peccati and A. Réveillac. Multivariate normal approximation using Stein's method and Malliavin calculus. Ann. Instit. H. Poincaré - Probab. Statist. 46, no. 1 (2010), 45-58.

[17] I. Nourdin and J. Rosiński. Asymptotic independence of multiple Wiener-Itô integrals and the resulting limit laws. Ann. Probab. 42, no. 2 (2014), 497-526

[18] D. Nualart. The Malliavin Calculus and Related Topics, second edition. *Probability and Its Applications*, Springer-Verlag Berlin Heidelberg (2006)

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

[20] G. Peccati and C.A. Tudor. Gaussian limits for vector-valued multiple stochastic integrals, Séminaire de Probabilités XXXVIII, 2005

[21] M. Reitzner and M. Schulte. Central Limit Theorems for U-Statistics of Poisson Point Processes, Ann. Probab.41, no. 6 (2013), 3879-3909

[22] A. Röllin. A note on the exchangeability condition in Stein's method. Statist. Probab. Lett. 78 (2008), 1800-1806.

[23] M. Schulte. A Central Limit Theorem for the Poisson-Voronoi Approximation, Adv. Appl. Math. 49, no. 3-5 (2012), 285-306

[24] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, *Berkeley Symp. on Math. Statist. and Prob. Proc. Sixth Berkeley Symp. on Math. Statist. and Prob., Vol. 2* (Univ. of Calif. Press, 1972), 583-602.

This page is left blank

Paper 4: Fourth moment theorems on the Poisson space in any dimension

Christian Döbler, Anna Vidotto and Guangqu Zheng

to appear in *Eletron. J. Probab.*

Abstract

We extend to any dimension the quantitative fourth moment theorem on the Poisson setting, recently proved by C. Döbler and G. Peccati (2017). In particular, by adapting the exchangeable pairs couplings construction introduced by I. Nourdin and G. Zheng (2017) to the Poisson framework, we prove our results under the weakest possible assumption of finite fourth moments. This yields a Peccati-Tudor type theorem, as well as an optimal improvement in the univariate case.

Finally, a transfer principle "from-Poisson-to-Gaussian" is derived, which is closely related to the universality phenomenon for homogeneous multilinear sums.

1 Introduction and main results

1.1 Outline

In the recent paper [13], the authors succeeded in proving exact quantitative *fourth moment theorems* for multiple Wiener-Itô integrals on the Poisson space. Briefly, their method consisted in extending the spectral framework initiated by the remarkable paper [22], and further refined by [1], from the situation of a diffusive Markov generator to the non-diffusive Ornstein-Uhlenbeck generator on the Poisson space. The principal aim of the present article is to extend the results from [13] to the multivariate case of vectors of multiple integrals. In view of the result of Peccati and Tudor [35] on vectors of multiple integrals on a Gaussian space, we are in particular interested in discussing the relationship between coordinatewise convergence and joint convergence to normality. Indeed, one of our main achievements is a complete quantitative version of a Peccati-Tudor type theorem on the Poisson space (see Theorem 1.7 and Corollary 1.8).

Furthermore, still keeping the spectral point of view as in [13], by replacing the rather intrinsic techniques used there with an adaption of a recent construction of *exchangeable pairs couplings* from [32], we can even remove certain technical conditions which seem inevitable in order to justify the computations in [13]. In this way, we are able to prove our results under the *weakest*

Key words and phrases. Stein's method; Exchangeable pairs; Brownian motion; Malliavin calculus.
AMS 2000 Classification: 60F05; 60H07; 60H05

possible assumption of finite fourth moments. In the univariate case, our strategy provides an optimal improvement of the Wasserstein bound given in Theorem 1.3 of [13] and, a fortiori, of the associated qualitative fourth moment theorem on the Poisson space (see Corollary 1.4 in [13]).

1.2 Motivation and related works

The so-called *fourth moment theorem* by Nualart and Peccati [33] states that a normalised sequence of multiple Wiener-Itô integrals of fixed order on a Gaussian space converges in distribution to a standard normal random variable N, if and only if the corresponding sequence of fourth moments converges to 3, i.e. to the fourth moment of N. For future reference, we give a precise statement of this result:

THEOREM **1.1** ([33]). *Let $F_n = I_q^W(f_n)$ be a sequence of multiple Wiener-Itô integral of order $q \geq 2$, associated with a Brownian motion $(W_t, t \in \mathbb{R}_+)$ such that $f_n \in L^2(\mathbb{R}_+^q)$ is symmetric for each $n \in \mathbb{N}$, and $\lim_{n\to+\infty} \mathbb{E}[F_n^2] = \sigma^2 > 0$. Then, the following statements are equivalent:*

(1) $\mathbb{E}[F_n^4] \to 3\sigma^4$, *as $n \to +\infty$.*

(2) *F_n converges in law to a Gaussian distribution $\mathcal{N}(0, \sigma^2)$, as $n \to +\infty$.*

(3) *For each $r \in \{1, 2, \dots, q-1\}$, $\|f_n \otimes_r f_n\|_{L^2(\mathbb{R}_+^{2q-2r})} \to 0$, as $n \to +\infty$.*

The contraction $f_n \otimes_r f_n$ is defined as in Section 2. See [27] *for any unexplained notions and notation of Gaussian analysis.*

Note that such a result significantly simplifies the method of moments for sequences of random variables inside a fixed Wiener chaos. In the years after the appearance of [33], this result has been extended and refined in many respects. While [35] provided a significant multivariate extension (see Theorem 1.10), the paper [26] combined Stein's method of normal approximation and Malliavin calculus in order to yield quantitative bounds for the normal approximation of general smooth functionals on the Wiener space. We refer to the monograph [27] for a comprehensive treatment of the so-called *Malliavin-Stein approach* on the Wiener space and of results obtained in this way. One remarkable result quoted from [27] is that, if F is a normalised multiple Wiener-Itô integral of order $q \geq 1$ on a Gaussian space, then one has the bound

$$d_{\mathrm{TV}}(F, N) \leq 2\sqrt{\frac{q-1}{3q}(\mathbb{E}[F^4] - 3)}\,, \tag{1.1}$$

where d_{TV} denotes the total variation distance between the laws of two real random variables. The techniques developed in [26] have also been adapted to non-Gaussian spaces which admit a Malliavin calculus structure: for instance, the papers [16, 34, 36, 41] deal with the Poisson space case, whereas [17, 18, 30, 44] develop the corresponding techniques for sequences of independent Rademacher random variables. The question about general fourth moment theorems on these spaces, however, has remained open in general, until the two recent articles [13] and [11].

1.3 General framework

Let us fix a measurable space $(\mathcal{Z}, \mathscr{Z})$, endowed with a σ-finite measure μ. We let

$$\mathscr{Z}_\mu := \{B \in \mathscr{Z} \,:\, \mu(B) < \infty\}$$

and define

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

to be a *Poisson random measure* on $(\mathcal{Z}, \mathscr{Z})$ with *control* μ, defined on a suitable probability space $(\Omega, \mathcal{F}, \mathbb{P})$. By definition, the distribution of η is completely determined by the following two properties:

(i) for each finite sequence $B_1, \dots, B_m \in \mathscr{Z}$ of pairwise disjoint sets, the random variables $\eta(B_1), \dots, \eta(B_m)$ are independent;

(ii) for every $B \in \mathscr{Z}$, the random variable $\eta(B)$ has the Poisson distribution with mean $\mu(B)$.

Here, we have extended the family of Poisson distributions to the parameter region $[0, +\infty]$ in the usual way. For $B \in \mathscr{Z}_\mu$, we also define $\widehat{\eta}(B) := \eta(B) - \mu(B)$ and denote by

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

the *compensated Poisson random measure* associated with η. Before stating our main results, we need to define some objects from stochastic analysis on the Poisson space. For a detailed discussion see, among others, [19] and [21].

For $q \in \mathbb{N}_0 := \{0, 1, 2, \dots\}$ and $f \in L^2(\mu^q)$, we denote by $I_q^\eta(f)$ the q-th order *multiple Wiener-Itô integral* of f with respect to $\widehat{\eta}$. Let L be the generator of the *Ornstein-Uhlenbeck semigroup* with respect to η, then it is well known that the spectrum of $-L$ is given by the set of nonnegative integers $\mathbb{N}_0$ and that, for $q \in \mathbb{N}_0$, F is an eigenfunction of $-L$ with eigenvalue q, if and only if $F = I_q^\eta(f)$ for some $f \in L^2(\mu^q)$. The corresponding eigenspace C_q will be called the *q-th Poisson Wiener chaos* associated with η. In particular, $C_0 = \mathbb{R}$.

1.4 Main results in the one-dimensional case

Recall that the *Wasserstein distance* between (the distributions of) two real random variables X and Y in $L^1(\mathbb{P})$ is defined by

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

where Lip(1) denotes the class of all 1-Lipschitz functions on $\mathbb{R}$.

In the univariate case, our main result reads as follows.

Theorem **1.2** (Fourth moment bound on the Poisson space)**.** *Fix an integer $q \geq 1$ and let $F \in C_q$ be such that $\sigma^2 := \mathbb{E}[F^2] > 0$ and $\mathbb{E}[F^4] < \infty$. Then, with N denoting a standard normal random variable, we have the bounds:*

$$d_{\mathcal{W}}(F, \sigma N) \leq \left(\frac{2q-1}{\sigma q\sqrt{2\pi}} + \frac{2}{3\sigma}\sqrt{\frac{4q-3}{q}} \right) \sqrt{\mathbb{E}[F^4] - 3\sigma^4} \tag{1.2}$$

$$\leq \left(\frac{1}{\sigma}\sqrt{\frac{2}{\pi}} + \frac{4}{3\sigma} \right) \sqrt{\mathbb{E}[F^4] - 3\sigma^4}\,. \tag{1.3}$$

Theorem 1.2 immediately implies the following qualitative statement, which is analogous to the Nualart-Peccati theorem [33] on a Gaussian space.

Corollary **1.3** (Fourth moment theorem on the Poisson space)**.** *For each $n \in \mathbb{N}$, let $q_n \in \mathbb{N}$ and $F_n \in C_{q_n}$ satisfy*

$$\lim_{n\to\infty} \mathbb{E}[F_n^2] = 1 \quad \textit{and} \quad \lim_{n\to\infty} \mathbb{E}[F_n^4] = 3\,.$$

Then, the sequence $(F_n)_{n\in\mathbb{N}}$ converges in distribution to a standard normal random variable N.

Remark 1.4. (a) Theorem 1.2 and Corollary 1.3 are genuine improvements of Theorem 1.3 and Corollary 1.4 from [13], respectively, since they do not require any additional regularity from the involved multiple integrals like e.g. **Assumption A** in [13]. The main reason for the appearance of such a condition in [13] was that certain intrinsic tools used there, notably the *Mecke formula* and a pathwise representation of the Ornstein-Uhlenbeck generator L, require $L^1(\mathbb{P}\otimes\mu)$-integrability conditions. It is the reconciliation of such conditions with the L^2 nature of the objects under consideration which necessitated these assumptions. As will become clear from our proofs in Section 4, such conditions can be completely avoided using an adaptation of exchangeable pairs couplings introduced in [32].

(b) In view of the well-known relation $d_{\mathcal{K}}(F, N) \leq \sqrt{d_{\mathcal{W}}(F, N)}$ between the Kolmogorov distance and the Wasserstein distance, one can obtain the fourth moment bound in the Kolmogorov distance with order 1/4 from Theorem 1.2, under the weakest possible assumption of finite fourth moment. However, the techniques applied in the present paper do not seem capable of proving a bound of order 1/2 in the Kolmogorov distance $d_{\mathcal{K}}(F, N)$.

(c) Indeed, it is an open question whether there is a general bound in the Kolmogorov distance via Stein's method of exchangeable pairs leading to the same accuracy as the one in the Wasserstein distance. It is worth noting that the authors of [13] were able to obtain, under a certain *local version* of **Assumption A** therein, the fourth moment bound in the Kolmogorov distance:

$$d_{\mathcal{K}}(F, N) \leq C\,\sqrt{\mathbb{E}[F^4] - 3\sigma^4}\,,$$

where $F \in C_q$ with $q \in \mathbb{N}$ and C is a numerical constant. See [13] for more details.

In the particular case where

$$\eta \text{ is a Poisson random measure on } \mathbb{R}_+ \text{ with Lebesgue intensity,} \tag{\#}$$

we observe the following *transfer principle* that is of independent interest.

Proposition **1.5.** *Assume* (#) *and* $(W_t, t \in \mathbb{R}_+)$ *is a standard Brownian motion. Given* $p \in \mathbb{N}$, $f_n \in L^2(\mathbb{R}_+^p)$ *symmetric for each* $n \in \mathbb{N}$ *such that*

$$\lim_{n\to+\infty} p!\, \|f_n\|^2_{L^2(\mathbb{R}_+^p)} = 1 \ ,$$

then the following implications holds ($N \sim \mathcal{N}(0,1)$)

$$\lim_{n\to+\infty} \mathbb{E}[I_p^\eta(f_n)^4] = 3 \Longrightarrow \lim_{n\to+\infty} \mathbb{E}[I_p^W(f_n)^4] = 3 \Longrightarrow \lim_{n\to+\infty} d_{\mathrm{TV}}(I_p^W(f_n), N) = 0 \ .$$

Remark 1.6. This transfer principle "from-Poisson-to-Gaussian" is closely related to the universality of Gaussian Wiener chaos and Poisson Wiener chaos, see Section 1.6. It is also worth pointing out that the transfer principle "from-Gaussian-to-Poisson" does *not* hold true, due to a counterexample given in [4], See Proposition 5.4 therein.

1.5 Main results in the multivariate case

In this subsection, let us fix integers $d \geq 2$ and $1 \leq q_1 \leq q_2 \leq \ldots \leq q_d$ and consider a random vector

$$F := (F_1, \ldots, F_d)^T \ ,$$

where $F_j \in C_{q_j}$, $1 \leq j \leq d$. We will further assume that $F_j \in L^4(\mathbb{P})$ for each $j \in \{1, \ldots, d\}$. Furthermore, we denote by $\Sigma := (\Sigma_{i,j})_{i,j=1,\ldots,d}$ the covariance matrix of F, i.e. $\Sigma_{i,j} = \mathbb{E}[F_i F_j]$ for $1 \leq i, j \leq d$. Note that $\Sigma_{i,j} = 0$ whenever $q_i \neq q_j$ due to the orthogonality properties of multiple integrals (see Section 2.1), and hence Σ is always a block diagonal matrix. Denote by $N = (N_1, \ldots, N_d)^T$ a centred Gaussian random vector with the same covariance matrix Σ.

In order to formulate our bounds, we need to fix some further notation: for a vector $x = (x_1, \ldots, x_d)^T \in \mathbb{R}^d$, we denote by $\|x\|_2$ its *Euclidean norm* and for a matrix $A \in \mathbb{R}^{d\times d}$, we denote by $\|A\|_{\mathrm{op}}$ the *operator norm* induced by the Euclidean norm, i.e.,

$$\|A\|_{\mathrm{op}} := \sup\{\|Ax\|_2 \ : \ \|x\|_2 = 1\} \, .$$

More generally, for a k-multilinear form $\psi : (\mathbb{R}^d)^k \to \mathbb{R}$, $k \in \mathbb{N}$, we define the *operator norm*

$$\|\psi\|_{\mathrm{op}} := \sup\left\{|\psi(u_1, \ldots, u_k)| \ : \ u_j \in \mathbb{R}^d,\ \|u_j\|_2 = 1,\ j = 1, \ldots, k\right\} .$$

Recall that for a function $h : \mathbb{R}^d \to \mathbb{R}$, its minimum Lipschitz constant $M_1(h)$ is given by

$$M_1(h) := \sup_{x\neq y} \frac{|h(x) - h(y)|}{\|x - y\|_2} \in [0, \infty] \, .$$

If h is differentiable, then $M_1(h) = \sup_{x\in\mathbb{R}^d} \|Dh(x)\|_{\mathrm{op}}$. More generally, for $k \geq 1$ and a $(k-1)$-times differentiable function $h : \mathbb{R}^d \to \mathbb{R}$, we set

$$M_k(h) := \sup_{x\neq y} \frac{\|D^{k-1}h(x) - D^{k-1}h(y)\|_{\mathrm{op}}}{\|x - y\|_2} \ ,$$

viewing the $(k-1)$-th derivative $D^{k-1}h$ of h at any point x as a $(k-1)$-multilinear form. Then, if h is k-times differentiable, we have

$$M_k(h) = \sup_{x\in\mathbb{R}^d} \|D^k h(x)\|_{\mathrm{op}} .$$

Recall that, for two matrices $A, B \in \mathbb{R}^{d\times d}$, their *Hilbert-Schmidt inner product* is defined by

$$\langle A, B\rangle_{\mathrm{H.S.}} := \mathrm{Tr}(AB^T) = \mathrm{Tr}(BA^T) = \mathrm{Tr}(B^T A) = \sum_{i,j=1}^{d} A_{i,j}B_{i,j} .$$

Thus, $\langle\cdot,\cdot\rangle_{\mathrm{H.S.}}$ is just the standard inner product on $\mathbb{R}^{d\times d} \cong \mathbb{R}^{d^2}$. The corresponding *Hilbert-Schmidt norm* will be denoted by $\|\cdot\|_{\mathrm{H.S.}}$. With this notion at hand, following [7] and [24], for $k = 2$ we finally define

$$\widetilde{M}_2(h) := \sup_{x\in\mathbb{R}^d} \|\mathrm{Hess}\, h(x)\|_{\mathrm{H.S.}} ,$$

where $\mathrm{Hess}\, h$ is the *Hessian matrix* corresponding to h. Note that for a symmetric matrix $A \in \mathbb{R}^{d\times d}$ with eigenvalues $\lambda_1(A) \le \ldots \le \lambda_d(A)$, one has

$$\|A\|^2_{\mathrm{H.S.}} = \sum_{j=1}^{d} \lambda_j(A)^2 \le d \, \max\{|\lambda_1(A)|^2, \ldots, |\lambda_d(A)|^2\} = d\|A\|^2_{\mathrm{op}} .$$

From this, it follows immediately that $\widetilde{M}_2(h) \le \sqrt{d}\, M_2(h)$.

The next statement is our main result in the multivariate setting.

Theorem **1.7.** *Under the above assumptions and notation, we have the following bounds:*

(i) *For every $g \in C^3(\mathbb{R}^d)$ such that $g(F), g(N) \in L^1(\mathbb{P})$, we have*

$$\begin{aligned}\left|\mathbb{E}[g(F)] - \mathbb{E}[g(N)]\right| \le\ & B_3(g) \sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2} \\ & + A_2(g)\left(\sum_{i=1}^{d-1} \mathbb{E}[F_i^4]^{1/4}\right) \sum_{j=2}^{d} \left(\mathbb{E}[F_j^4] - 3\mathbb{E}[F_j^2]^2\right)^{1/4} ,\end{aligned} \tag{1.4}$$

with $B_3(g) = A_2(g) + \dfrac{2q_d\sqrt{d\mathrm{Tr}(\Sigma)}}{9q_1} M_3(g)$ *and* $A_2(g) = \dfrac{(2q_d-1)\sqrt{2d}}{4q_1} M_2(g)$.

(ii) *If in addition Σ is positive definite, then for every $g \in C^2(\mathbb{R}^d)$ such that $g(F), g(N) \in L^1(\mathbb{P})$, we have*

$$\begin{aligned}\left|\mathbb{E}[g(F)] - \mathbb{E}[g(N)]\right| \le\ & B_2(g) \sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2} \\ & + A_1(g)\left(\sum_{i=1}^{d-1} \mathbb{E}[F_i^4]^{1/4}\right) \sum_{j=2}^{d} \left(\mathbb{E}[F_j^4] - 3\mathbb{E}[F_j^2]^2\right)^{1/4} ,\end{aligned} \tag{1.5}$$

with

$$B_2(g) = A_1(g) + \frac{q_d \sqrt{2\pi}\|\Sigma^{-1/2}\|_{\mathrm{op}} \sqrt{\mathrm{Tr}(\Sigma)}}{6q_1} M_2(g)$$

and

$$A_1(g) = \frac{(2q_d - 1)\|\Sigma^{-1/2}\|_{\mathrm{op}}}{q_1 \sqrt{\pi}} M_1(g)\,.$$

The qualitative statement in the multivariate situation reads as follows.

COROLLARY **1.8.** *Fix $d \in \mathbb{N}$ and $q_1, \dots, q_d \in \mathbb{N}$ and suppose that, for each $n \in \mathbb{N}$, $F^{(n)} := (F_1^{(n)}, \dots, F_d^{(n)})^T$ is a random vector such that each $F_k^{(n)}$ belongs to the q_k-th Poisson Wiener chaos. Moreover, assume that $C = C(i,j)_{1 \le i,j \le d}$ is a fixed nonnegative definite matrix and that $N = (N_1, \dots, N_d)^T$ is a centred Gaussian vector with covariance matrix C. Assume that the following two conditions hold true:*

(i) *The covariance matrix of $F^{(n)}$ converges to C as $n \to \infty$.*

(ii) *For each $1 \le k \le d$ it holds that $\lim_{n\to\infty} \mathbb{E}[(F_k^{(n)})^4] = 3C(k,k)^2$.*

Then, as $n \to \infty$, the random vector $F^{(n)}$ converges in distribution to N.

REMARK 1.9. (a) Comparing the bounds in Theorem 1.7 with the one provided in Theorem 1.2, one observes that in the multivariate case the order of dependence on the fourth cumulants of the respective coordinates is $1/4$ instead of $1/2$. This phenomenon, which technically results from an application of the Cauchy-Schwarz inequality in order to disentangle certain joint moments of the coordinate variables, is nothing peculiar of the Poisson framework but also arises in the Gaussian situation [25] and in the recent multivariate de Jong type CLT for vectors of degenerate non-symmetric U-statistics [12]. Moreover, this phenomenon only arises in the case when there are components belonging to different chaoses (see Remark 4.3).

(b) We stress that it is remarkable that, as in the Gaussian case [35], the bounds and conditions in Theorem 1.7 and Corollary 1.8 can be expressed just in terms of the individual fourth cumulants of the components of the random vectors. Indeed, both in the general situation of diffusive Markov generators (see [5, Theorem 1.2]) and for the multivariate CLT for vectors of degenerate non-symmetric U-statistics (see [12, Theorem 1.7]), one additionally needs to assume the convergence of mixed fourth moments of those entries, which are of the same chaos and Hoeffding order, respectively.

(c) Corollary 1.8 is a full Poisson space analogue of the Peccati-Tudor theorem [35] for vectors of multiple integrals on a Gaussian space, which boils down the question about joint convergence of the whole vector to conditions guaranteeing coordinatewise convergence (via Corollary 1.3).

For the convenience of later reference, we state below the theorem of Peccati-Tudor on a Gaussian space.

Theorem **1.10** ([35]). *Let $(W_t, t \in \mathbb{R}_+)$ be a real standard Brownian motion, and we fix integers $d \geq 2$ and $1 \leq q_1 \leq \ldots \leq q_d$. Let $C = C(i,j)_{1\leq i,j\leq d}$ be a $d \times d$ symmetric nonnegative definite matrix and for any $n \geq 1$, $i \in \{1, \ldots, d\}$, let $f_{n,i} \in L^2(\mathbb{R}_+^{q_i}, dx)$ be symmetric. Assume that the d-dimensional random vectors*

$$F^{(n)} = (F_1^{(n)}, \ldots, F_d^{(n)})^T := (I_{q_1}^W(f_{n,1}), \ldots, I_{q_d}^W(f_{n,d}))^T$$

satisfy

$$\lim_{n\to+\infty} \mathbb{E}[F_i^{(n)} F_j^{(n)}] = C(i,j)\,, \quad i,j \in \{1, \ldots, d\}\,.$$

Then, as $n \to +\infty$, the following assertions are **equivalent***:*

(1) *The vector $F^{(n)}$ converges in distribution to a d-dimensional Gaussian vector $\mathcal{N}(0, C)$;*

(2) *for every $i \in \{1, \ldots, d\}$, $F_i^{(n)}$ converges in distribution to a real Gaussian random variable $\mathcal{N}(0, C(i,i))$;*

(3) *for every $i \in \{1, \ldots, d\}$, $\mathbb{E}[(F_i^{(n)})^4] \to 3C(i,i)^2$;*

(4) *for every $i \in \{1, \ldots, d\}$ and each $1 \leq r \leq q_i - 1$, $\left\| f_{n,i} \otimes_r f_{n,i} \right\|_{L^2(\mathbb{R}_+^{2q_i-2r})} \to 0$.*

For a proof, one can refer to [27].

1.6 Universality of Homogeneous sums

The transfer principle in Proposition 1.5 is closely related to the *universality phenomenon* for multilinear homogeneous sums in independent Poisson random variables. We refer to the papers [29], [37], [28] and [2] for the universality results on multilinear homogeneous sums. Before we can state the result, we need to introduce some notation.

Notation. Suppose that $d \geq 2$, $N \in \mathbb{N}$, and that $f \in \ell^2(\mathbb{N}^d)$ is a function, which is symmetric in its arguments and vanishes on diagonals, *i.e.* for any $i_1, \ldots, i_d \in \mathbb{N}$, $f(i_1, \ldots, i_d) = f(i_{\sigma(1)}, \ldots, i_{\sigma(d)})$ for any $\sigma \in \mathbb{S}_d$ and $f(i_1, \ldots, i_d) = 0$, whenever $i_p = i_q$ for some $p \neq q$. For a sequence $\mathbf{X} = (X_i, i \in \mathbb{N})$ of real random variables, we define the multilinear homogeneous sum of order d, based on the kernel f and on the first N elements of $\mathbf{X}$ by

$$Q_d(f, N, \mathbf{X}) := \sum_{1\leq i_1,\ldots,i_d\leq N} f(i_1, \ldots, i_d) X_{i_1} \cdots X_{i_d}\,. \tag{1.6}$$

Now let us consider an independent sequence $\mathbf{P} = (P_i, i \in \mathbb{N})$ of normalised Poisson random variables, which can be realised via our Poisson random measure η on $\mathbb{R}_+$. More precisely, let $(t_i, i \in \mathbb{N})$ be a strictly increasing sequence of positive numbers. Set

$$P_i := \frac{\widehat{\eta}([t_i, t_{i+1}))}{\sqrt{t_{i+1} - t_i}} = I_1^\eta\left(\frac{1}{\sqrt{t_{i+1} - t_i}} \mathbf{I}_{[t_i, t_{i+1})} \right),$$

$i \in \mathbb{N}$. We are now in the position to state the universality result.

Theorem **1.11.** *Let the above notation prevail. Fix integers $d \geq 2$ and $q_d \geq \ldots \geq q_1 \geq 2$. For each $j \in \{1, \ldots, d\}$, let $(N_{n,j}, n \geq 1)$ be a sequence of natural numbers diverging to infinity, and let $f_{n,j} : \{1, \ldots, N_{n,j}\}^{q_j} \to \mathbb{R}$ be symmetric and vanishing on diagonals such that*

$$\lim_{n\to+\infty} \mathbf{I}_{(q_k=q_l)} q_k! \sum_{i_1,\ldots,i_{q_k} \leq N_{n,k}} f_{n,k}(i_1, \ldots, i_{q_k}) f_{n,l}(i_1, \ldots, i_{q_k}) = \Sigma(k,l)\,,$$

where $\Sigma = \Sigma(i,j)_{1\leq i,j\leq d}$ is a symmetric nonnegative definite d by d matrix. Then the following condition (A_0) implies two **equivalent** *statements (A_1), (A_2) :*

(A_0) *For each $j \in \{1, \ldots, d\}$, one has $\lim_{n\to+\infty} \mathbb{E}[Q_{q_j}(f_{n,j}, N_{n,j}, \mathbf{P})^4] = 3\Sigma(j,j)^2$.*

(A_1) *Let $\mathbf{G}$ be a sequence of i.i.d. standard Gaussian random variables, then, as $n \to +\infty$, $(Q_{q_1}(f_{n,1}, N_{n,1}, \mathbf{G}), \ldots, Q_{q_d}(f_{n,d}, N_{n,d}, \mathbf{G}))^T$ converges in distribution to $\mathcal{N}(0,\Sigma)$.*

(A_2) *For every sequence $\mathbf{X} = (X_i, i \in \mathbb{N})$ of independent centred random variables with unit variance and $\sup_{i\in\mathbb{N}} \mathbb{E}[|X_i|^3] < +\infty$, the sequence of d-dimensional random vectors $(Q_{q_1}(f_{n,1}, N_{n,1}, \mathbf{X}), \ldots, Q_{q_d}(f_{n,d}, N_{n,d}, \mathbf{X}))^T$ converges in distribution to $\mathcal{N}(0,\Sigma)$, as $n \to +\infty$.*

If, in addition, $\inf\{t_{i+1} - t_i : i \in \mathbb{N}\} > 0$, then (A_0) and (A_1)-(A_2) are equivalent to the following assertion

(A_3) *$(Q_{q_1}(f_{n,1}, N_{n,1}, \mathbf{P}), \ldots, Q_{q_d}(f_{n,d}, N_{n,d}, \mathbf{P}))^T$ converges to $\mathcal{N}(0,\Sigma)$ in distribution, as $n \to +\infty$.*

Remark 1.12. The authors of [37] established a fourth moment theorem for sequences of homogeneous sums in independent Poisson random variables whose variance is bounded away from zero, namely, $\inf\{t_{i+1} - t_i : i \in \mathbb{N}\} > 0$ in our language. In particular, in order to get the implication "$(A_0) \Rightarrow (A_1)$", they relied heavily on the assumption that $\inf\{t_{i+1} - t_i : i \in \mathbb{N}\} > 0$, which is inevitable due to their use of product formula. As a consequence, our Theorem 1.11 is an improvement of the results in [37].

Plan of the paper. In Section 2, we review some necessary definitions and facts about multiple integrals and Malliavin operators on the Poisson space. Section 3 is devoted to the essential construction of a suitable family of exchangeable pairs for the concrete purpose of establishing fourth moment bounds on the Poisson space. In order to make use of it, we also state two new abstract plug-in results for such families of exchangeable pairs. In Section 4 we give the proofs of our main results, whereas Section 5 presents the proofs of Proposition 1.5, Theorem 1.11 as well as certain technical auxiliary results.

2 Some stochastic analysis on the Poisson space

2.1 Basic operators and notation

For a positive integer p, we denote by $L^2(\mu^p)$ the Hilbert space of all square-integrable and real-valued functions on $\mathcal{Z}^p$, and we denote by $L^2_s(\mu^p)$ the subspace of $L^2(\mu^p)$ whose elements are μ^p-a.e. symmetric. Moreover, we indicate by $\|\cdot\|_2$ and $\langle\cdot,\cdot\rangle_2$ respectively the usual norm and scalar product on $L^2(\mu^p)$ for any value of p. We also set $L^2(\mu^0) := \mathbb{R}$. For $f \in L^2(\mu^p)$, we define $I_p^\eta(f)$ to be the *multiple Wiener-Itô integral* of f with respect to the compensated Poisson random measure $\widehat{\eta}$. If $p = 0$, then, by convention, $I_0^\eta(c) := c$ for each $c \in \mathbb{R}$.

The multiple Wiener-Itô integrals satisfy the following properties:

1) For $p \in \mathbb{N}$ and $f \in L^2(\mu^p)$, $I_p^\eta(f) = I_p^\eta(\widetilde{f})$, where $\widetilde{f}$ denotes the *symmetrization* of $f \in L^2(\mu^p)$, i.e.
$$\widetilde{f}(z_1,\dots,z_p) = \frac{1}{p!}\sum_{\pi\in\mathbb{S}_p} f(z_{\pi(1)},\dots,z_{\pi(p)}) \ ,$$
where $\mathbb{S}_p$ is the symmetric group acting on $\{1,\dots,p\}$. Note $\widetilde{c} = c$ for any $c \in \mathbb{R}$.

2) For $p, q \in \mathbb{N}_0$ and $f \in L^2(\mu^p)$, $g \in L^2(\mu^q)$, one has $I_p^\eta(f)$, $I_q^\eta(g) \in L^2(\mathbb{P})$ and $\mathbb{E}[I_p^\eta(f)I_q^\eta(g)] = \delta_{p,q}\, p!\,\langle\widetilde{f},\widetilde{g}\rangle_2$, where $\delta_{p,q}$ denotes *Kronecker's delta symbol*.

See Section 3 of [19] for the proofs of the above well known results.

For $p \in \mathbb{N}_0$, the Hilbert space $C_p := \{I_p^\eta(f),\ f \in L^2(\mu^p)\}$, is called the *p-th Poisson Wiener chaos* associated with η. The well-known *Wiener-Itô chaotic decomposition* states that every $F \in L^2(\mathbb{P}) := L^2(\Omega, \sigma\{\eta\}, \mathbb{P})$ admits a unique representation

$$F = \mathbb{E}[F] + \sum_{p=1}^{\infty} I_p^\eta(f_p) \quad \text{in } L^2(\mathbb{P}), \text{ where } f_p \in L_s^2(\mu^p),\ p \geq 1. \tag{2.1}$$

Let $F \in L^2(\mathbb{P})$ and $p \in \mathbb{N}_0$, then we define by $J_p(F)$ the orthogonal projection of F on C_p. Note that, if F has the chaotic decomposition as in (2.1), then $J_p(F) = I_p^\eta(f_p)$ for all $p \geq 1$ and $J_0(F) = \mathbb{E}[F]$.

For $F \in L^2(\mathbb{P})$ with the chaotic decomposition as in (2.1), we define
$$P_tF = \mathbb{E}[F] + \sum_{p\geq 1} e^{-pt}\, I_p^\eta(f_p)\ .$$
This gives us the *Ornstein-Uhlenbeck semigroup* $(P_t, t \in \mathbb{R}_+)$. The domain $\operatorname{dom} L$ of the *Ornstein-Uhlenbeck generator* L is the set of those $F \in L^2(\mathbb{P})$ with the chaotic decomposition (2.1) verifing $\sum_{p=1}^\infty p^2\, p!\|f_p\|_2^2 < +\infty$, and for $F \in \operatorname{dom} L$, one has

$$LF = -\sum_{p=1}^{\infty} pI_p^\eta(f_p)\,. \tag{2.2}$$

We conclude from (2.2) that LF is always centred, $\mathbb{N}_0$ is the spectrum of $-L$ and $F \in \operatorname{dom} L$ is an eigenfunction of $-L$ with corresponding eigenvalue p if and only if $F = I_p^\eta(f_p)$ for some $f_p \in L_s^2(\mu^p)$, *i.e.* $C_p = \operatorname{Ker}(L + pI)$.

Moreover, it is easy to see that L is *symmetric* in the sense that $\mathbb{E}[GLF] = \mathbb{E}[FLG]$ for all $F, G \in \operatorname{dom} L$. Finally, for $F, G \in \operatorname{dom} L$ with $FG \in \operatorname{dom} L$, we define the *carré du champ* operator Γ associated with L by

$$\Gamma(F,G) := \frac{1}{2}(L(FG) - FLG - GLF)\,, \tag{2.3}$$

and it is easy to verify that $\mathbb{E}[\Gamma(F,G)] = -\mathbb{E}[FLG] = -\mathbb{E}[GLF]$. It follows from Lemma 2.1 below that $\Gamma(F,G)$ is always well-defined whenever $F,G \in L^4(\mathbb{P})$ and both have a finite chaotic decomposition.

In the book [3], the authors develop Dirichlet form method for the Poisson point process, and starting from the Dirichlet form associated with the Ornstein-Uhlenbeck structure, they obtain an expression of carré du champs operator that is close to the one derived in [13]. As readers will see, we only need the spectral decomposition rather than the intrinsic tools in [3, 13]. This highlights the elementary feature of our method.

For $p,q \in \mathbb{N}$, $0 \le r \le p \wedge q$ and $f \in L^2_s(\mu^p)$ and $g \in L^2_s(\mu^q)$, we define the r-th *contraction* $f \otimes_r g : \mathcal{Z}^{p+q-2r} \to \mathbb{R}$ by

$$f \otimes_r g(x_1,\dots,x_{p-r},y_1,\dots,y_{q-r}) := \int_{\mathcal{Z}^r} f(x_1,\dots,x_{p-r},z_1,\dots,z_r) \cdot g(y_1,\dots,y_{q-r},z_1,\dots,z_r) d\mu^r(z_1,\dots,z_r)\,.$$

Observe that $f \otimes_r g \in L^2(\mu^{p+q-2r})$ is in general not symmetric and that $f \otimes_0 g = f \otimes g$ is simply the tensor product of f and g.

Lemma **2.1** (Lemma 2.4 of [13])**.** *Let $p,q \in \mathbb{N}$ and $F = I^\eta_p(f)$, $G = I^\eta_q(g)$ be in $L^4(\mathbb{P})$ with f,g symmetric, then FG has a finite chaotic decomposition of the form*

$$FG = \sum_{r=0}^{p+q} J_r(FG) = \sum_{r=0}^{p+q} I^\eta_r(h_r)\,,$$

where $h_r \in L^2_s(\mu^r)$ for each r. In particular, $h_{p+q} = f\widetilde{\otimes}g$.

2.2 Useful estimates via spectral decomposition

To conclude the section, we state several lemmas that are useful for our proofs.

Lemma **2.2.** *Let $F \in L^4(\mathbb{P}) \cap C_p$ and $G \in L^4(\mathbb{P}) \cap C_q$ for $p,q \in \mathbb{N}$. Then,*

$$\operatorname{Var}(\Gamma(F,G)) \le \frac{(p+q-1)^2}{4}\left(\mathbb{E}[F^2G^2] - 2\,\mathbb{E}[FG]^2 - \operatorname{Var}(F)\operatorname{Var}(G)\right)\,, \tag{2.4}$$

and

$$0 \le \frac{3}{p}\mathbb{E}[F^2\Gamma(F,F)] - \mathbb{E}[F^4] \le \frac{4p-3}{2p}\left(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2\right)\,. \tag{2.5}$$

In particular, for $F = G$, we obtain

$$\operatorname{Var}(\Gamma(F,F)) \le \frac{(2p-1)^2}{4}\left(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2\right). \tag{2.6}$$

Note the authors of [13] provided a proof of (2.5) under the **Assumption A** therein, while we only require the assumption of finite fourth moment. Although (2.6) is the content of Lemma 3.1 in [13], we will provide another proof, in which we deduce a nice relation between contractions of kernels and the fourth cumulant. Such a relation is crucial for us to obtain the transfer principle "from-Poisson-to-Gaussian". The proof of Lemma 2.2 as well as that of the next lemma will be presented in Section 5.

LEMMA **2.3.** *Under the same assumptions of Lemma 2.2, we have that*

(1) *If $p < q$, then*

$$\begin{aligned} \mathrm{Cov}(F^2, G^2) &= \mathbb{E}[F^2G^2] - \mathrm{Var}(F)\mathrm{Var}(G) \\ &\leq \sqrt{\mathbb{E}[F^4]}\sqrt{\mathbb{E}[G^4] - 3\mathbb{E}[G^2]^2}\ ; \end{aligned} \tag{2.7}$$

(2) *if $p = q$, then*

$$\mathrm{Cov}(F^2, G^2) - 2\,\mathbb{E}[FG]^2 \leq 2\sqrt{(\mathbb{E}[F^4] - 3\,\mathbb{E}[F^2]^2)(\mathbb{E}[G^4] - 3\mathbb{E}[G^2]^2)}\,. \tag{2.8}$$

This lemma is motivated by Proposition 3.6 in [5].

3 Stein's method of exchangeable pairs

The exchangeable pairs approach within Stein's method was first used in the paper [9] which, however, attributes the method to Charles Stein himself. Later, this technique was presented in a systematic way in Stein's monograph [43]. We recall that a pair (X, X') of random elements on a common probability space is said to be *exchangeable*, if (X, X') has the same distribution as (X', X). In the book [43], it is highlighted that a given real random variable W is close in distribution to a standard normal variable N, whenever one can construct an exchangeable pair (W, W') such that W' is close to W in *some* sense and that the *linear regression property*

$$\mathbb{E}\big[W' - W \,\big|\, W\big] = -\lambda W$$

is satisfied for some small $\lambda > 0$ and $\mathrm{Var}\left(\frac{1}{2\lambda}\mathbb{E}[(W' - W)^2|W]\right)$ is small. For a precise statement, we refer to Theorem 1 in [43, Lecture III].

In recent years, the method of exchangeable pairs has been generalised for other distributions and multi-dimensional settings in many papers like [6–8, 10, 15, 38–40, 42], to name a few.

Moreover, the articles [7, 14, 23, 24] develop versions of the exchangeable pairs method suitable for situations, in which one can construct a continuous family $(W, W_t)_{t>0}$ of exchangeable pairs. By their continuity assumptions, these papers succeed in reducing the order of smoothness of test functions and hence in obtaining bounds in more sophisticated probabilistic distances. For instance, the bounds from [23] are expressed in terms of the total variation distance. It is this framework of exchangeable pairs that is most closely related to the variant of the method developed in the present paper. In contrast to the quoted papers, however, our abstract results on exchangeable pairs do not make such strong continuity assumptions and hence, allow us to deal with the inherent discreteness of the Poisson space, which, in general, does not even allow for convergence in the total variation distance, see Section 3.2 for details.

3.1 Exchangeable pairs constructed via continuous thinning

Recall that, in our general framework, η is a Poisson random measure on some σ-finite measure space $(\mathcal{Z}, \mathscr{Z}, \mu)$. As a consequence, we can assume that η is a *proper* Poisson point process, that is, almost surely

$$\eta = \sum_{n=1}^{\kappa} \delta_{X_n} , \tag{3.1}$$

where X_n, $n \geq 1$ are random variables with values in $\mathcal{Z}$ and κ is a $\mathbb{N}_0 \cup \{+\infty\}$-valued random variable. Indeed, according to Corollary 3.7 in [21], any Poisson random measure η on some σ-finite measure space is equal in distribution to some proper Poisson point process. As in this work, we are only concerned with distributional properties, we will always assume that η is of the form (3.1).

Let $\mathbf{N}_\sigma$ be the collection of σ-finite measures $\nu : \mathscr{Z} \to \mathbb{N}_0 \cup \{+\infty\}$ and $\mathscr{N}_\sigma(\mathcal{Z})$ be the σ-algebra generated by the maps $\nu \in \mathbf{N}_\sigma \longmapsto \nu(B)$, $B \in \mathscr{Z}$. We consider the Poisson point process η as a random element in $(\mathbf{N}_\sigma, \mathscr{N}_\sigma(\mathcal{Z}))$. Moreover, for any $F \in L^0(\Omega, \sigma\{\eta\}, \mathbb{P})$, one can find a ($\mathbb{P}$-a.s. unique) *representative* $\mathfrak{f}$ of F such that $F = \mathfrak{f}(\eta)$, see [19] for more details.

Now let $\mathbb{Q}$ be a standard exponential measure on $\mathbb{R}_+$ with density $\exp(-y)\,dy$, and let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of *i.i.d.* random variables with distribution $\mathbb{Q}$, independent of (κ, X_n). Then the *marked* point process ξ, given by

$$\xi := \sum_{n=1}^{\kappa} \delta_{(X_n,Y_n)} ,$$

is a Poisson point process with control $\mu \otimes \mathbb{Q}$. For each $t \in \mathbb{R}_+$, we define

$$\eta_{e^{-t}}(A) := \xi(A \times [t, +\infty)) ,$$

which is called the e^{-t}-thinning of η: it is obtained by removing the atoms (X_n) in η independently of each other with probability $1 - e^{-t}$. Moreover, $\eta_{e^{-t}}$ and $\eta - \eta_{e^{-t}}$ are two independent Poisson point processes with control measure $e^{-t}\mu$, $(1 - e^{-t})\mu$ respectively. One can refer to Chapter 5 in [21] for more details.

For any fixed $t \geq 0$, let $\eta'_{1-e^{-t}}$ be a Poisson point process on $\mathcal{Z}$ with control $(1 - e^{-t})\mu$ such that it is independent of $(\eta, \eta_{e^{-t}})$. Then the *Mehler formula* gives a useful representation of the Ornstein-Uhlenbeck semigroup (P_t): for $F \in L^2(\Omega, \sigma\{\eta\}, \mathbb{P})$,

$$P_t F = \mathbb{E}\big[\mathfrak{f}(\eta_{e^{-t}} + \eta'_{1-e^{-t}})\big|\sigma\{\eta\}\big] ,$$

where $\mathfrak{f}$ is a *representative* of F, see [19] for more details. We remark that the Mehler formula on the Poisson space has already been effectively used in [20] in order to obtain a pathwise representation for the pseudo-inverse of the Ornstein-Uhlenbeck generator L on the Poisson space, which has led to second-order Poincaré inequalities.

We record an important observation in the following lemma.

Lemma 3.1. *For each $t \in \mathbb{R}_+$, set $\eta^t := \eta_{e^{-t}} + \eta'_{1-e^{-t}}$. Then, (η, η^t) is an exchangeable pair of Poisson point processes.*

Proof. To prove this lemma, it suffices to notice that $\eta = \eta_{e^{-t}} + \eta - \eta_{e^{-t}}$ and that $\eta - \eta_{e^{-t}}, \eta'_{1-e^{-t}}$ have the same law, and are both independent of $\eta_{e^{-t}}$.

Let $\mathfrak{f} : \mathbf{N}_\sigma \to \mathbb{R}$ be $\mathscr{N}_\sigma(\mathcal{Z})$-measurable; then, for any Borel subsets A_1, A_2 of $\mathbb{R}$, one has that

$$\begin{aligned}
&\mathbb{P}(\mathfrak{f}(\eta) \in A_1\,,\ \mathfrak{f}(\eta^t) \in A_2)\\
&= \mathbb{P}(\mathfrak{f}(\eta_{e^{-t}} + \eta - \eta_{e^{-t}}) \in A_1\,,\ \mathfrak{f}(\eta_{e^{-t}} + \eta'_{1-e^{-t}}) \in A_2)\\
&= \mathbb{P}(\mathfrak{f}(\eta_{e^{-t}} + \eta'_{1-e^{-t}}) \in A_1\,,\ \mathfrak{f}(\eta_{e^{-t}} + \eta - \eta_{e^{-t}}) \in A_2) \quad \text{(by conditioning on } \eta_{e^{-t}})\\
&= \mathbb{P}(\mathfrak{f}(\eta^t) \in A_1\,,\ \mathfrak{f}(\eta) \in A_2)\ .
\end{aligned}$$

This implies the exchangeability of (η, η^t). □

The following result is a consequence of Lemma 3.1 and Mehler formula: it is a key ingredient for us to obtain exact fourth moment theorems in any dimension. Indeed, it fits extremely well with the abstract results for exchangeable pairs that are presented in Section 3.2.

PROPOSITION **3.2.** *Let $F = I_q^\eta(f) \in L^4(\mathbb{P})$ for some $f \in L_s^2(\mu^q)$ and define $F_t = I_q^{\eta^t}(f)$. Then, (F, F_t) is an exchangeable pair for each $t \in \mathbb{R}_+$. Moreover,*

(a) $\lim_{t\downarrow 0} \frac{1}{t}\mathbb{E}[F_t - F|\sigma\{\eta\}] = LF = -qF$ *in* $L^4(\mathbb{P})$.

(b) *If* $G = I_p^\eta(g) \in L^4(\mathbb{P})$ *and* $G_t = I_p^{\eta^t}(g)$ *for some* $g \in L_s^2(\mu^p)$*, then we have* $\lim_{t\downarrow 0} \frac{1}{t}\mathbb{E}[(F_t - F)(G_t - G)|\sigma\{\eta\}] = 2$ *with the convergence in* $L^2(\mathbb{P})$.

(c) $\lim_{t\downarrow 0} \frac{1}{t}\mathbb{E}[(F_t - F)^4] = -4q\,\mathbb{E}[F^4] + 12\,\mathbb{E}[F^2\Gamma(F, F)] \geq 0.$

Proof. The exchangeability of F, F_t is an immediate consequence of Lemma 3.1. Relation (a) is a direct consequence of the Mehler formula:

$$\frac{1}{t}\mathbb{E}[F_t - F|\sigma\{\eta\}] = \frac{P_t(F) - F}{t} = \frac{e^{-qt} - 1}{t}F\,,$$

and such a quantity converges almost surely, and in $L^4(\mathbb{P})$ to $LF = -qF$, as $t \downarrow 0$.

By Lemma 2.1, $FG = \sum_{k=0}^{p+q} J_k(FG) = \sum_{k=0}^{p+q} I_k^\eta(h_k)$ for some $h_k \in L_s^2(\mu^k)$, and consequently $F_t G_t = \sum_{k=0}^{p+q} I_k^{\eta^t}(h_k)$, so that

$$\frac{1}{t}\mathbb{E}[F_t G_t - FG|\sigma\{\eta\}] = \frac{1}{t}\sum_{k=0}^{p+q}\mathbb{E}[I_k^{\eta^t}(h_k) - I_k^\eta(h_k)|\sigma\{\eta\}]$$

converges almost surely and in $L^2(\mathbb{P})$ to $\sum_{k=0}^{p+q} -k\,J_k(FG) = L(FG)$, as $t \downarrow 0$. Hence almost surely and in $L^2(\mathbb{P})$, we infer that

$$\begin{aligned}
&\frac{1}{t}\mathbb{E}[(F_t - F)(G_t - G)|\sigma\{\eta\}]\\
&= \frac{1}{t}\mathbb{E}[F_t G_t - FG|\sigma\{\eta\}] - F\frac{\mathbb{E}[G_t - G|\sigma\{\eta\}]}{t} - G\frac{\mathbb{E}[F_t - F|\sigma\{\eta\}]}{t}
\end{aligned}$$

$$\to L(FG) - FLG - GLF = 2\Gamma(F,G),$$

as $t \downarrow 0$. Since the pair (F, F_t) is exchangeable, we can write

$$\begin{aligned}
\mathbb{E}[(F_t - F)^4] &= \mathbb{E}[F_t^4 + F^4 - 4F_t^3 F - 4F^3 F_t + 6F_t^2 F^2] \\
&= 2\mathbb{E}[F^4] - 8\mathbb{E}[F^3 F_t] + 6\mathbb{E}[F^2 F_t^2] \quad \text{(by exchangeability)} \\
&= 4\mathbb{E}[F^3(F_t - F)] + 6\mathbb{E}[F^2(F_t - F)^2] \quad \text{(after rearrangement)} \\
&= 4\mathbb{E}[F^3\mathbb{E}[F_t - F|\sigma\{\eta\}]] + 6\mathbb{E}[F^2\mathbb{E}[(F_t - F)^2|\sigma\{\eta\}]].
\end{aligned}$$

so (c) follows immediately from (a),(b) and the fact that $F \in L^4(\mathbb{P})$. □

3.2 Abstract results for exchangeable pairs

As indicated in the introductory part of this section, the following two Propositions should be seen as complements to [23, Theorem 1.4] and [24, Theorem 4] as well as [7, Theorem 2.4], respectively. The main difference with respect to these results, as mentioned above, is that we do not assume any continuity from the respective families of exchangeable pairs, which precisely means that we allow for non-zero limits in the respective conditions (c) below.

PROPOSITION **3.3.** *Let Y and a family of random variables $(Y_t)_{t\geq 0}$ be defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $Y_t \overset{law}{=} Y$ for every $t \geq 0$. Assume that $Y \in L^4(\Omega, \mathcal{G}, \mathbb{P})$ for some σ-algebra $\mathcal{G} \subset \mathcal{F}$ and that, in $L^1(\mathbb{P})$,*

(a) $\lim_{t\downarrow 0} \frac{1}{t}\mathbb{E}[Y_t - Y|\mathcal{G}] = -\lambda Y$ *for some* $\lambda > 0$,

(b) $\lim_{t\downarrow 0} \frac{1}{t}\mathbb{E}[(Y_t - Y)^2|\mathcal{G}] = (2\lambda + S)\mathrm{Var}(Y)$ *for some random variable* S,

(c) $\lim_{t\downarrow 0} \frac{1}{t}\mathbb{E}[(Y_t - Y)^4] = \varrho(Y)\,\mathrm{Var}(Y)^2$ *for some* $\varrho(Y) \geq 0$.

Then, with $N \sim \mathcal{N}(0, \mathrm{Var}(Y))$, we have

$$d_W(Y,N) \leq \frac{\sqrt{\mathrm{Var}(Y)}}{\lambda\sqrt{2\pi}}\mathbb{E}[|S|] + \frac{\sqrt{(2\lambda + \mathbb{E}[S])\mathrm{Var}(Y)}}{3\lambda}\sqrt{\varrho(Y)}\,.$$

REMARK 3.4. If the quantity $\varrho(Y) = 0$ in (c), then Proposition 3.3 reduces to Theorem 1.3 in [32] and one has

$$d_{\mathrm{TV}}(Y,N) := \sup_{A\subset\mathbb{R}\text{ Borel}} \left|\mathbb{P}(Y \in A) - \mathbb{P}(N \in A)\right| \leq \frac{\mathbb{E}[|S|]}{\lambda}\,.$$

The following result is a multivariate extension of Proposition 3.3. The proofs will be postponed to Section 5.5 and 5.6.

PROPOSITION **3.5.** *For each $t > 0$, let (X, X_t) be an exchangeable pair of centred d-dimensional random vectors defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{G}$ be a σ-algebra that contains $\sigma\{X\}$. Assume that $\Lambda \in \mathbb{R}^{d\times d}$ is an invertible deterministic matrix and Σ is a symmetric, non-negative definite deterministic matrix such that*

(a) $\lim_{t\downarrow 0} \frac{1}{t}\,\mathbb{E}[X_t - X|\mathscr{G}] = -\Lambda X$ *in* $L^1(\mathbb{P})$,

(b) $\lim_{t\downarrow 0} \frac{1}{t}\,\mathbb{E}[(X_t - X)(X_t - X)^T|\mathscr{G}] = 2\Lambda\Sigma + S$ *in* $L^1(\Omega, \|\cdot\|_{\mathrm{H.S.}})$ *for some random matrix* S,

(c) *for each* $i \in \{1, \dots, d\}$, *there exists some real number* $\varrho_i(X)$ *such that*

$$\lim_{t\downarrow 0} \frac{1}{t}\,\mathbb{E}[(X_{i,t} - X_i)^4] = \varrho_i(X)\ ,$$

where $X_{i,t}$ *(resp.* X_i*) stands for the i-th coordinate of* X_t *(resp.* X*).*

Then, with $N \sim \mathcal{N}(0, \Sigma)$*, we have the following bounds:*

(1) *For* $g \in C^3(\mathbb{R}^d)$ *such that* $g(X), g(N) \in L^1(\mathbb{P})$*, one has*

$$\begin{aligned}&\Big|\mathbb{E}[g(X) - g(N)]\Big|\\ &\le \Theta_1(g)\,\mathbb{E}[\|S\|_{\mathrm{H.S.}}] + \Theta_2(g)\ \sqrt{\sum_{i=1}^d 2\Lambda_{i,i}\Sigma_{i,i} + \mathbb{E}[S_{i,i}]}\sqrt{\sum_{i=1}^d \varrho_i(X)}\,,\end{aligned}$$

where the constants $\Theta_1(g)$ *and* $\Theta_2(g)$ *are given by*

$$\Theta_1(g) = \frac{\|\Lambda^{-1}\|_{\mathrm{op}}\, M_2(g)\,\sqrt{d}}{4} \quad \text{and} \quad \Theta_2(g) = \frac{\sqrt{d}M_3(g)\|\Lambda^{-1}\|_{\mathrm{op}}}{18}\ . \tag{3.2}$$

(2) *If, in addition,* Σ *is positive definite, then for* $g \in C^2(\mathbb{R}^d)$ *such that* $g(X), g(N) \in L^1(\mathbb{P})$*, one has*

$$\begin{aligned}&\Big|\mathbb{E}[g(X) - g(N)]\Big|\\ &\le K_1(g)\,\mathbb{E}[\|S\|_{\mathrm{H.S.}}] + K_2(g)\ \sqrt{\sum_{i=1}^d 2\Lambda_{i,i}\Sigma_{i,i} + \mathbb{E}[S_{i,i}]}\sqrt{\sum_{i=1}^d \varrho_i(X)}\ ,\end{aligned}$$

where the constants $K_1(g)$ *and* $K_2(g)$ *are given by*

$$K_1(g) = \frac{M_1(g)\|\Lambda^{-1}\|_{\mathrm{op}}\|\Sigma^{-1/2}\|_{\mathrm{op}}}{\sqrt{2\pi}}\,, \tag{3.3}$$

$$K_2(g) = \frac{\sqrt{2\pi}M_2(g)\|\Lambda^{-1}\|_{\mathrm{op}}\|\Sigma^{-1/2}\|_{\mathrm{op}}}{24}\,. \tag{3.4}$$

4 Proofs of main results

4.1 Proof of Theorem 1.2

Without loss of generality, we assume $F = I_q^\eta(f)$ for some $f \in L_s^2(\mu^q)$, and we define $F_t = I_q^{\eta^t}(f)$ for $t \in \mathbb{R}_+$. Then, by Proposition 3.2, (F, F_t) is an exchangeable pair and the assumptions (a), (b), (c) in Proposition 3.3 are satistified with

- $\lambda = q$ • $S = 2\dfrac{\Gamma(F,F)}{\sigma^2} - 2q$ • $\varrho(F) = \dfrac{-4q\mathbb{E}[F^4] + 12\mathbb{E}[F^2\Gamma(F,F)]}{\sigma^4}$.

More precisely,

(a) $\displaystyle\lim_{t\downarrow 0} \frac{1}{t}\, \mathbb{E}[F_t - F|\sigma\{\eta\}] = -qF,$

(b) $\displaystyle\lim_{t\downarrow 0} \frac{1}{t}\, \mathbb{E}[(F_t - F)^2|\sigma\{\eta\}] = 2\Gamma(F,F)\,,$

(c) $\displaystyle\lim_{t\downarrow 0} \frac{1}{t}\, \mathbb{E}[(F_t - F)^4] = \varrho(F)\sigma^4.$

Therefore, one has (using that $\mathbb{E}[\Gamma(F,F)] = q\,\mathbb{E}[F^2]$)

$$d_{\mathcal{W}}(F, \mathcal{N}(0,\sigma^2)) \le \frac{\sqrt{2/\pi}}{\sigma q}\sqrt{\mathrm{Var}(\Gamma(F,F))} + \frac{2\sqrt{2}}{3\sigma}\sqrt{\frac{3}{q}\mathbb{E}[F^2\Gamma(F,F)] - \mathbb{E}[F^4]}\ .$$

The desired result follows immediately from Lemma 2.2.

4.2 Proof of Theorem 1.7

Assume that

$$F = (F_1, \ldots, F_d)^T = (I_{q_1}^\eta(f_1), \ldots, I_{q_d}^\eta(f_d))^T$$

with $1 \le q_1 \le \ldots \le q_d$ and $f_j \in L_s^2(\mu^{q_j})$ for each j, and for each $t \in \mathbb{R}_+$, define

$$F_t = (F_{1,t}, \ldots, F_{d,t})^T = (I_{q_1}^{\eta^t}(f_1), \ldots, I_{q_d}^{\eta^t}(f_d))^T\ .$$

Then, by Lemma 3.1, (F_t, F) is an exchangeable pair and by Proposition 3.2, we deduce

$$\mathbb{E}\left[\frac{1}{t}(F_{i,t} - F_i)(F_{j,t} - F_j) - 2\Gamma(F_i, F_j)\,\Big|\, \sigma\{\eta\}\right] \to 0\ ,$$

as $t \downarrow 0$, where the convergence takes place in $L^2(\mathbb{P})$. Therefore, as $t \downarrow 0$ and in $L^1(\mathbb{P})$, we have

$$\left\|\frac{1}{t}\mathbb{E}[(F_t - F)(F_t - F)^T|\sigma\{\eta\}] - (2\Gamma(F_i, F_j))_{1\le i,j\le d}\right\|_{\mathrm{H.S.}}^2$$

$$= \sum_{i,j=1}^{d} \left(\mathbb{E}\left[\frac{1}{t}(F_{i,t} - F_i)(F_{j,t} - F_j) - 2\Gamma(F_i, F_j) \,\middle|\, \sigma\{\eta\} \right] \right)^2 \to 0 \ .$$

It is easy to see that for each $j \in \{1, \ldots, d\}$,

$$\lim_{t \downarrow 0} \frac{1}{t} \mathbb{E}[F_{j,t} - F_j | \sigma\{\eta\}] = -q_j F_j \quad \text{in } L^4(\mathbb{P}),$$

from which we deduce that as $t \downarrow 0$ and in $L^2(\mathbb{P})$, we have

$$\left\| \frac{1}{t}\mathbb{E}[F_t - F|\sigma\{\eta\}] - \Lambda F \right\|_2^2 = \sum_{j=1}^{d} \left(\mathbb{E}\left[\frac{F_{j,t} - F_j}{t} + q_j F_j \,\middle|\, \sigma\{\eta\} \right] \right)^2 \to 0 \ ,$$

with $\Lambda = \mathrm{diag}(q_1, \ldots, q_d)$ in such a way that $\|\Lambda^{-1}\|_{\mathrm{op}} = 1/q_1$.

It is also clear that, for each $i \in \{1, \ldots, d\}$,

$$\begin{aligned} \varrho_i(F) := \lim_{t \downarrow 0} \frac{1}{t} \mathbb{E}[(F_{i,t} - F_i)^4] &= -4q_i \mathbb{E}[F_i^4] + 12\mathbb{E}[F_i^2 \Gamma(F_i, F_i)] \\ &\le 2(4q_i - 3)\Big(\mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2\Big) \quad \text{by (2.5).} \end{aligned}$$

Now define $S_{i,j} := 2\Gamma(F_i, F_j) - 2q_i \Sigma_{i,j}$ for $i, j \in \{1, \ldots, d\}$, and observe in particular that $S_{i,j}$ has zero mean. Thus,

$$\begin{aligned} &\sqrt{\sum_{i=1}^{d} 2\Lambda_{i,i}\Sigma_{i,i} + \mathbb{E}[S_{i,i}]} \sqrt{\sum_{i=1}^{d} \varrho_i(F)} \\ &\le \sqrt{\sum_{i=1}^{d} 2q_i \Sigma_{i,i}} \sqrt{\sum_{i=1}^{d} 2(4q_i - 3)\Big(\mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2\Big)} \\ &\le \sqrt{4q_d(4q_d - 3)\mathrm{Tr}(\Sigma)} \sqrt{\sum_{i=1}^{d} \Big(\mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2\Big)} \\ &\le 4q_d \sqrt{\mathrm{Tr}(\Sigma)} \sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2} \ , \end{aligned} \tag{4.1}$$

where the last inequality follows from the elementary fact that $\sqrt{a_1 + \ldots + a_d} \le \sqrt{a_1} + \ldots + \sqrt{a_d}$ for any nonnegative reals $a_1, \ldots, a_d$.

Now we consider $\mathbb{E}[\|S\|_{\mathrm{H.S.}}]$:

$$\begin{aligned} \mathbb{E}[\|S\|_{\mathrm{H.S.}}] = \mathbb{E}\left(\sqrt{\sum_{i,j=1}^{d} S_{i,j}^2} \right) &\le \left(\sum_{i,j=1}^{d} \mathbb{E}[S_{i,j}^2] \right)^{1/2} \\ &= 2\left(\sum_{i,j=1}^{d} \mathrm{Var}\big(\Gamma(F_i, F_j)\big) \right)^{1/2} \ . \end{aligned} \tag{4.2}$$

It follows from (2.4) that

$$\sum_{i,j=1}^{d} \mathrm{Var}\left(\Gamma(F_i, F_j)\right) \le \sum_{i,j=1}^{d} \frac{(q_i+q_j-1)^2}{4}\Big(\mathbb{E}[F_i^2F_j^2] - 2\mathbb{E}[F_iF_j]^2 - \mathrm{Var}(F_i)\,\mathrm{Var}(F_j)\Big)$$

$$\le \frac{(2q_d-1)^2}{4} \sum_{i,j=1}^{d} \Big(\mathbb{E}[F_i^2F_j^2] - 2\mathbb{E}[F_iF_j]^2 - \mathrm{Var}(F_i)\,\mathrm{Var}(F_j)\Big)$$

$$= \frac{(2q_d-1)^2}{4}\mathbb{E}[\|F\|_2^4 - \|N\|_2^4]\ , \tag{4.3}$$

where the last equality is a consequence of the fact that (see *e.g.* (4.2) in [31])

$$\mathbb{E}[\|N\|_2^4] = \sum_{i,j=1}^{d} (\Sigma_{i,i}\Sigma_{j,j} + 2\Sigma_{i,j}^2)\ .$$

Lemma **4.1.** *Let F, N be given as before, then*

$$\mathbb{E}[\|F\|_2^4] - \mathbb{E}[\|N\|_2^4]$$
$$\le 2\left(\sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\right)^2 + 2\left(\sum_{i=1}^{d-1} \sqrt{\mathbb{E}[F_i^4]}\right)\sum_{j=2}^{d} \sqrt{\mathbb{E}[F_j^4] - 3\,\mathbb{E}[F_j^2]^2}\ .$$

In particular, if $q_1 = \ldots = q_d$, one has,

$$\mathbb{E}[\|F\|_2^4] - \mathbb{E}[\|N\|_2^4] \le 2\left(\sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\right)^2\ .$$

Proof. Let us first consider the particular case where $q_1 = \ldots = q_d$. One obtains from Lemma 2.3 that

$$\mathbb{E}[\|F\|_2^4] - \mathbb{E}[\|N\|_2^4] = \sum_{i,j=1}^{d} \Big(\mathbb{E}[F_i^2F_j^2] - 2\mathbb{E}[F_iF_j]^2 - \mathrm{Var}(F_i)\,\mathrm{Var}(F_j)\Big)$$

$$\le 2\sum_{i,j=1}^{d} \sqrt{(\mathbb{E}[F_i^4] - 3\mathbb{E}[F_i^2]^2)(\mathbb{E}[F_j^4] - 3\mathbb{E}[F_j^2]^2)}$$

$$= 2\left(\sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\right)^2\ .$$

In the general case where $q_1 \le \ldots \le q_d$, Lemma 2.3 implies

$$\mathbb{E}[\|F\|_2^4] - \mathbb{E}[\|N\|_2^4]$$
$$= \sum_{i,j=1}^{d} \mathbf{I}_{(q_i=q_j)}\Big(\mathrm{Cov}(F_i^2, F_j^2) - 2\mathbb{E}[F_iF_j]^2\Big) + 2\sum_{1\le i<j\le d} \mathbf{I}_{(q_i<q_j)}\mathrm{Cov}(F_i^2, F_j^2) \tag{4.4}$$

$$\leq 2\left(\sum_{i=1}^{d}\sqrt{\mathbb{E}[F_i^4]-3\,\mathbb{E}[F_i^2]^2}\right)^2+2\sum_{1\leq i<j\leq d}\sqrt{\mathbb{E}[F_i^4]}\sqrt{(\mathbb{E}[F_j^4]-3\mathbb{E}[F_j^2]^2)}\;.$$

One can rewrite $\sum_{1\leq i<j\leq d}$ as $\sum_{j=2}^{d}\sum_{i=1}^{j-1}$ and then the desired result follows. □

Remark 4.2. Note that in the same way, we can provide another proof of the quantitative Peccati-Tudor Theorem in the Gaussian setting. In particular, keeping the indicator functions in (4.4), we can obtain the bound in a similar form as in [25, Theorem 1.5].

End of the proof of Theorem 1.7. First we obtain from (4.2)-(4.3) and Lemma 4.1 that

$$\begin{aligned}\mathbb{E}[\|S\|_{\mathrm{H.S.}}] &\leq (2q_d-1)\sqrt{\mathbb{E}[\|F\|_2^4-\|N\|_2^4]}\\ &\leq \sqrt{2}(2q_d-1)\sum_{i=1}^{d}\sqrt{\mathbb{E}[F_i^4]-3\,\mathbb{E}[F_i^2]^2}\\ &\quad+\sqrt{2}(2q_d-1)\left(\sum_{i=1}^{d-1}\mathbb{E}[F_i^4]^{1/4}\right)\sum_{j=2}^{d}(\mathbb{E}[F_j^4]-3\,\mathbb{E}[F_j^2]^2)^{1/4}\,.\end{aligned}\tag{4.5}$$

If $g\in C^3(\mathbb{R}^d)$ and $g(F),g(N)$ are integrable, then by Proposition 3.5, we deduce

$$\begin{aligned}&\left|\mathbb{E}[g(F)-g(N)]\right|\\ &\leq \Theta_1(g)\,\mathbb{E}[\|S\|_{\mathrm{H.S.}}]+\Theta_2(g)\sqrt{\sum_{i=1}^{d}2\Lambda_{i,i}\Sigma_{i,i}+\mathbb{E}[S_{i,i}]}\sqrt{\sum_{i=1}^{d}\varrho_i(F)}\\ &\leq \sqrt{2}(2q_d-1)\Theta_1(g)\sum_{i=1}^{d}\sqrt{\mathbb{E}[F_i^4]-3\,\mathbb{E}[F_i^2]^2}\\ &\quad+4q_d\Theta_2(g)\sqrt{\mathrm{Tr}(\Sigma)}\sum_{i=1}^{d}\sqrt{\mathbb{E}[F_i^4]-3\mathbb{E}[F_i^2]^2}\\ &\quad+\sqrt{2}(2q_d-1)\Theta_1(g)\left(\sum_{i=1}^{d-1}\mathbb{E}[F_i^4]^{1/4}\right)\sum_{j=2}^{d}(\mathbb{E}[F_j^4]-3\,\mathbb{E}[F_j^2]^2)^{1/4}\,,\end{aligned}$$

where the last inequality follows from (4.5) and (4.1). It is easy to check that

$$\sqrt{2}\Theta_1(g)(2q_d-1)+4q_d\sqrt{\mathrm{Tr}(\Sigma)}\Theta_2(g)=B_3(g)\quad\text{and}\quad\sqrt{2}(2q_d-1)\Theta_1(g)=A_2(g)\,.$$

Assertion (i) of Theorem 1.7 follows immediately. Assertion (ii) can be proved in the same way, by using moreover the relations:

$$\sqrt{2}K_1(g)(2q_d-1)+4q_d\sqrt{\mathrm{Tr}(\Sigma)}K_2(g)=B_2(g)\quad\text{and}\quad\sqrt{2}(2q_d-1)K_1(g)=A_1(g)\,.$$

REMARK 4.3. With the notation and assumptions given as in Theorem 1.7, if in addition $q_1 = q_d$, that is, all the components of the random vector F belong to the same Poisson Wiener chaos, then we can obtain better bounds, namely:

(i) For every $g \in C^3(\mathbb{R}^d)$ such that $g(F), g(N) \in L^1(\mathbb{P})$, we have

$$\left|\mathbb{E}[g(F)] - \mathbb{E}[g(N)]\right| \le B_3(g) \sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\,.$$

(ii) If, in addition, Σ is positive definite, then for every $g \in C^2(\mathbb{R}^d)$ such that $g(F), g(N) \in L^1(\mathbb{P})$, we have

$$\left|\mathbb{E}[g(F)] - \mathbb{E}[g(N)]\right| \le B_2(g) \sum_{i=1}^{d} \sqrt{\mathbb{E}[F_i^4] - 3\,\mathbb{E}[F_i^2]^2}\,.$$

5 Proofs of technical and auxiliary results

In this section, we first provide the proofs of Lemma 2.2, Lemma 2.3. The following result from [31] will be helfpul.

LEMMA **5.1** (Lemma 2.2 of [31])**.** *Given $p, q \in \mathbb{N}$, $f \in L^2_s(\mu^p)$ and $g \in L^2_s(\mu^q)$, then*

$$(p+q)!\,\|f\widetilde{\otimes}g\|_2^2 = p!q!\sum_{r=0}^{p\wedge q}\binom{p}{r}\binom{q}{r}\|f\otimes_r g\|_2^2 \ge p!q!\,\|f\|_2^2\|g\|_2^2 + \delta_{p,q}p!q!\,\langle f, g\rangle_2^2\,,$$

and in the case of $p = q$, one has

$$(2p)!\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\rangle_2 = 2p!^2\,\langle f, g\rangle_2^2 + \sum_{r=1}^{p-1} p!^2\binom{p}{r}^2\langle f\otimes_r g, g\otimes_r f\rangle_2\ .$$

Here we follow the convention that $\sum_{r=1}^{0} = 0$.

5.1 Proof of Lemma 2.2

Without loss of generality, we assume $F = I_p^\eta(f)$ and $G = I_q^\eta(g)$ for some $f \in L^2_s(\mu^p)$ and $g \in L^2_s(\mu^q)$. It follows from Lemma 2.1 and the definition of Γ that $J_{p+q}(FG) = I_{p+q}^\eta(f\widetilde{\otimes}g)$ and

$$2\,\Gamma(F,G) = (p+q)\mathbb{E}[FG] + \sum_{k=1}^{p+q-1}(p+q-k)\,J_k(FG)\,. \tag{5.1}$$

By orthogonality,

$$\mathrm{Var}(\Gamma(F,G)) = \frac{1}{4}\sum_{k=1}^{p+q-1}(p+q-k)^2\,\mathrm{Var}(J_k(FG))$$

$$\leq \frac{(p+q-1)^2}{4} \sum_{k=1}^{p+q-1} \mathrm{Var}(J_k(FG)) \ .$$

Similarly, as $FG \in L^2(\mathbb{P})$, we have $FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q} J_k(FG)$ so that

$$\begin{aligned}\mathbb{E}[F^2G^2] &= \mathbb{E}[FG]^2 + \sum_{k=1}^{p+q-1} \mathrm{Var}(J_k(FG)) + \mathrm{Var}(J_{p+q}(FG)) \\ &= \mathbb{E}[FG]^2 + \sum_{k=1}^{p+q-1} \mathrm{Var}(J_k(FG)) + (p+q)!\|f\widetilde{\otimes}g\|_2^2 \ .\end{aligned}$$

It follows from Lemma 5.1 that

$$(p+q)!\|f\widetilde{\otimes}g\|_2^2 \geq p!q!\|f\|_2^2\|g\|_2^2 + \delta_{p,q}p!q!\langle f,g\rangle_2^2 = \mathrm{Var}(F)\,\mathrm{Var}(G) + \mathbb{E}[FG]^2 \,.$$

Hence

$$\begin{aligned}\mathrm{Var}(\Gamma(F,G)) &\leq \frac{(p+q-1)^2}{4} \sum_{k=1}^{p+q-1} \mathrm{Var}(J_k(FG)) \\ &= \frac{(p+q-1)^2}{4}\Big(\mathbb{E}[F^2G^2] - \mathbb{E}[FG]^2 - (p+q)!\|f\widetilde{\otimes}g\|_2^2\Big) \\ &\leq \frac{(p+q-1)^2}{4}\Big(\mathbb{E}[F^2G^2] - 2\,\mathbb{E}[FG]^2 - \mathrm{Var}(F)\mathrm{Var}(G)\Big) \ . \qquad (5.2)\end{aligned}$$

In particular, Lemma 5.1, applied to $p = q$ and $f = g$, gives us

$$(2p)!\|f\widetilde{\otimes}f\|_2^2 = 2p!^2\|f\|_2^4 + p!^2 \sum_{r=1}^{p-1} \binom{p}{r}^2 \|f \otimes_r f\|_2^2 \ ,$$

therefore implying

$$\begin{aligned}\mathrm{Var}(\Gamma(F,F)) &\leq \frac{(2p-1)^2}{4}\Big(\mathbb{E}[F^4] - \mathbb{E}[F^2]^2 - (2p)!\|f\widetilde{\otimes}f\|_2^2\Big) \\ &= \frac{(2p-1)^2}{4}\left(\mathbb{E}[F^4] - 3\,\mathbb{E}[F^2]^2 - p!^2 \sum_{r=1}^{p-1} \binom{p}{r}^2 \|f \otimes_r f\|_2^2\right) \ .\end{aligned}$$

This proves (2.6) and

$$p!^2 \sum_{r=1}^{p-1} \binom{p}{r}^2 \|f \otimes_r f\|_2^2 \leq \mathbb{E}[F^4] - 3\,\mathbb{E}[F^2]^2 \ . \qquad (5.3)$$

It is also clear from (5.2) that

$$\sum_{k=1}^{2p-1} \mathrm{Var}(J_k(F^2)) \leq \mathbb{E}[F^4] - 3\,\mathbb{E}[F^2]^2 \,. \qquad (5.4)$$

It remains to show (2.5) now: similarly, we write $F^2 = \mathbb{E}[F^2] + \sum_{k=1}^{2p} J_k(F^2)$ and by (5.1)

$$\Gamma(F,F) = p\,\mathbb{E}[F^2] + \frac{1}{2}\sum_{k=1}^{2p-1}(2p-k)J_k(F^2)\ . \tag{5.5}$$

So by orthogonality, we have

$$\begin{aligned}
-\mathbb{E}[F^4] + \frac{3}{p}\mathbb{E}[F^2\,\Gamma(F,F)] &= -\mathbb{E}[F^4] + 3\mathbb{E}[F^2]^2 + \frac{3}{2p}\sum_{k=1}^{2p-1}(2p-k)\,\mathrm{Var}(J_k(F^2)) \\
&\le -\mathbb{E}[F^4] + 3\mathbb{E}[F^2]^2 + \frac{3(2p-1)}{2p}\sum_{k=1}^{2p-1}\mathrm{Var}(J_k(F^2)) \\
&\le -\mathbb{E}[F^4] + 3\mathbb{E}[F^2]^2 + \frac{3(2p-1)}{2p}(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2) \\
&= \frac{4p-3}{2p}(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2)\ .
\end{aligned}$$

The other inequality in (2.5) is a trivial consequence of Proposition 3.2-(c). The proof of Lemma 2.2 is complete.

Remark 5.2. 1. Let $F \in C_p$ have nonzero variance, then we have that $\mathbb{E}[F^4] > 3\mathbb{E}[F^2]^2$. Indeed, we can always assume $F \in L^4(\mathbb{P})$. If $p = 1$, $F = I_1^\eta(f)$ for some $f \in L^2(\mu)$, then by product formula (see *e.g.* Proposition 6.1 in [19]), one has $\mathbb{E}[I_1^\eta(f)^4] = 3\,\|f\|_2^4 + \int_{\mathcal{Z}} f(z)^4\,d\mu > 3\mathbb{E}[F^2]^2$. For $p \ge 2$, $F = I_p^\eta(f)$ for some $f \in L_s^2(\mu^p)$, then according to (5.3), $\mathbb{E}[F^4] = 3\mathbb{E}[F^2]^2$ would imply $\|f \otimes_1 f\|_2 = 0$, which would further imply by standard arguments that $f = 0$ μ-almost everywhere, which is a contradiction to the fact that F is nonzero.

2. Let $F \in C_p \cap L^4(\mathbb{P})$, one has $p(\mathbb{E}[F^4]-3\mathbb{E}[F^2]^2) \le 6\,\mathrm{Var}\,(\Gamma(F,F))$, which shall be compared with (2.6). In fact, it follows first from (2.5) that $\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 \le 3\mathbb{E}[F^2(p^{-1}\Gamma(F,F) - \mathbb{E}[F^2])]$, and by (5.5) and orthogonality property, we have

$$\begin{aligned}
\mathbb{E}[F^2(\Gamma(F,F) - p\mathbb{E}[F^2])] &= \frac{1}{2}\sum_{k=1}^{2p-1}(2p-k)\,\mathrm{Var}\,(J_k(F^2)) \\
&\le \frac{1}{2}\sum_{k=1}^{2p-1}(2p-k)^2\,\mathrm{Var}\,(J_k(F^2)) = 2\,\mathrm{Var}\,(\Gamma(F,F))\,,
\end{aligned}$$

hence $p(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2) \le 6\,\mathrm{Var}\,(\Gamma(F,F))$.

3. Let F, N be given as in Theorem 1.7, then from (4.3) it follows that $\mathbb{E}[\|F\|_2^4] \ge \mathbb{E}[\|N\|_2^4]$. Moreover, if one of the components F_j in F has nonzero variance, it follows from the above two points and again (4.3) that $\mathbb{E}[\|F\|_2^4] > \mathbb{E}[\|N\|_2^4]$.

5.2 Proof of Lemma 2.3

Assume $F = I_p^\eta(f)$ and $G = I_q^\eta(g)$ are in $L^4(\mathbb{P})$ for some $f \in L_s^2(\mu^p)$, $g \in L_s^2(\mu^q)$. Then it follows from Lemma 2.1 that $J_{2p}(F^2) = I_{2p}^\eta(f\widetilde{\otimes}f)$ and $J_{2q}(G^2) = I_{2q}^\eta(g\widetilde{\otimes}g)$. Moreover, one has

$$\begin{aligned}\mathbb{E}[F^2G^2] &= \mathbb{E}\left[F^2\sum_{k=0}^{2q} J_k(G^2)\right]\\ &= \mathbb{E}[F^2J_0(G^2)] + \mathbb{E}[F^2J_{2q}(G^2)] + \mathbb{E}\left[F^2\sum_{k=1}^{2q-1} J_k(G^2)\right]\\ &= \operatorname{Var}(F)\operatorname{Var}(G) + \mathbb{E}[F^2\,J_{2q}(G^2)] + \mathbb{E}\left[F^2\sum_{k=1}^{2q-1} J_k(G^2)\right].\end{aligned}$$

If $p < q$, then $\mathbb{E}[F^2\,J_{2q}(G^2)] = 0$, so that

$$\operatorname{Cov}(F^2, G^2) = \mathbb{E}\left[F^2\sum_{k=1}^{2q-1} J_k(G^2)\right] \le \sqrt{\mathbb{E}[F^4]}\sqrt{\sum_{k=1}^{2q-1}\operatorname{Var}(J_k(G^2))}\,,$$

where the above inequality follows from Cauchy-Schwarz inequality and isometry property. The desired result (2.7) follows from (5.4).

Now we consider the case where $p = q$,

$$\begin{aligned}\mathbb{E}\left[F^2\sum_{k=1}^{2p-1} J_k(G^2)\right] &= \sum_{k=1}^{2p-1}\mathbb{E}[J_k(F^2)J_k(G^2)]\\ &\le \sqrt{\sum_{k=1}^{2p-1}\operatorname{Var}\left(J_k(F^2)\right)}\sqrt{\sum_{k=1}^{2p-1}\operatorname{Var}\left(J_k(G^2)\right)} \text{ (by Cauchy-Schwarz)}\\ &\le \sqrt{(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2)(\mathbb{E}[G^4] - 3\mathbb{E}[G^2]^2)} \quad \text{due to (5.4).}\end{aligned}$$

By orthogonality property, we have

$$\begin{aligned}\mathbb{E}[J_{2p}(F^2)J_{2p}(G^2)] &= (2p)!\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\rangle_2\\ &= 2p!^2\langle f, g\rangle_2^2 + \sum_{r=1}^{p-1} p!^2\binom{p}{r}^2\langle f\otimes_r g, g\otimes_r f\rangle_2\,,\end{aligned}$$

where the last equality follows from Lemma 5.1.

As a consequence, one has

$$\mathbb{E}[F^2J_{2p}(G^2)] - 2\mathbb{E}[FG]^2 = \sum_{r=1}^{p-1} p!^2\binom{p}{r}^2\langle f\otimes_r g, g\otimes_r f\rangle_2 \le \sum_{r=1}^{p-1} p!^2\binom{p}{r}^2\left\|f\otimes_r g\right\|_2^2$$

by Cauchy-Schwarz. Note that, by definition of contractions and Fubini theorem, we have $\left\|f\otimes_r g\right\|_2^2 = \langle f\otimes_{p-r} f, g\otimes_{p-r} g\rangle_2$ for each $r = 1, \ldots, p-1$. Thus,

$$
\begin{aligned}
&\sum_{r=1}^{p-1} p!^2 \binom{p}{r}^2 \left\| f \otimes_r g \right\|_2^2 \\
&= \sum_{r=1}^{p-1} p!^2 \binom{p}{r}^2 \langle f \otimes_{p-r} f, g \otimes_{p-r} g \rangle_2 = \sum_{r=1}^{p-1} p!^2 \binom{p}{r}^2 \langle f \otimes_r f, g \otimes_r g \rangle_2 \\
&\leq \sum_{r=1}^{p-1} p!^2 \binom{p}{r}^2 \|f \otimes_r f\|_2 \|g \otimes_r g\|_2 \quad \text{(by Cauchy-Schwarz)} \\
&\leq \sqrt{\sum_{r=1}^{p-1} p!^2 \binom{p}{r}^2 \|f \otimes_r f\|_2^2} \sqrt{\sum_{r=1}^{p-1} p!^2 \binom{p}{r}^2 \|g \otimes_r g\|_2^2} \quad \text{(by Cauchy-Schwarz)} \\
&\leq \sqrt{\mathbb{E}[F^4] - 3\,\mathbb{E}[F^2]^2} \sqrt{\mathbb{E}[G^4] - 3\,\mathbb{E}[G^2]^2} \quad \text{due to (5.3).}
\end{aligned}
$$

Hence, we obtain

$$
\begin{aligned}
\operatorname{Cov}(F^2, G^2) - 2\,\mathbb{E}[FG]^2 &= \mathbb{E}[F^2 J_{2p}(G^2)] - 2\,\mathbb{E}[FG]^2 + \mathbb{E}\left[F^2 \sum_{k=1}^{2p-1} J_k(G^2) \right] \\
&\leq 2 \sqrt{\mathbb{E}[F^4] - 3\,\mathbb{E}[F^2]^2} \sqrt{\mathbb{E}[G^4] - 3\,\mathbb{E}[G^2]^2} \ .
\end{aligned}
$$

The proof is completed.

5.3 Proof of Proposition 1.5

It follows from (5.3) that

$$
p!^2 \sum_{r=1}^{p-1} \binom{p}{r}^2 \|f_n \otimes_r f_n\|_2^2 \leq \mathbb{E}[I_p^\eta(f_n)^4] - 3\,\mathbb{E}[I_p^\eta(f_n)^2]^2 \ .
$$

If $\mathbb{E}[I_p^\eta(f_n)^4] \to 3$ as $n \to +\infty$, then $\|f_n \otimes_r f_n\|_2 \to 0$ for each $r \in \{1, \ldots, p-1\}$. Therefore by Theorem 1.1, $\mathbb{E}[I_p^W(f_n)^4] \to 3$ and moreover by (1.1),

$$
\lim_{n \to +\infty} d_{\mathrm{TV}}(I_p^W(f_n), N) = 0 \ .
$$

This completes the proof of our transfer principle.

5.4 Proof of Theorem 1.11

The equivalence of (A_1) and (A_2) is the content of Theorem 7.5 in [29]. For each $i \in \mathbb{N}$, define

$$
g_i = \frac{1}{\sqrt{t_{i+1} - t_i}} \mathbf{I}_{[t_i, t_{i+1})},
$$

then the homogeneous sum $Q_d(f,N,\mathbf{P})$, defined according to (1.6), can be expressed as the d-th multiple integral $I_d^\eta(\widehat{f})$, where

$$\widehat{f} := \sum_{1\le i_1,\dots,i_d\le N} f(i_1,\dots,i_d) g_{i_1}\otimes\cdots\otimes g_{i_d}\ . \tag{5.6}$$

From now on, we identify $\widehat{f}$ with f in case of no confusion. Observe that the sequence $\mathbf{G}$ of *i.i.d* standard Gaussian random variables can be realised via the Brownian motion $(W_t, t\in\mathbb{R}_+)$. That is, for each $i\in\mathbb{N}$, we put $G_i = I_1^W(g_i)$. As a consequence, the homogeneous sum $Q_d(f,N,\mathbf{G})$ can be rewritten as $I_d^W(\widehat{f})$, with $\widehat{f}$ given in (5.6).

With these notions at hand and in view of our transfer principle, if (A_0) holds, then, for each $j\in\{1,\dots,d\}$ and every $r\in\{1,\dots,q_j-1\}$, $\left\|f_{n,j}\otimes_r f_{n,j}\right\|_2 \to 0$, as $n\to+\infty$. Then, (A_1) is an immediate consequence of Theorem 1.10.

Finally, it is known that the fourth *central* moment of a Poisson random variable with parameter $\lambda\in(0,+\infty)$ is given by $\lambda(1+3\lambda)$, then $\mathbb{E}[P_i^4] = 3 + (t_{i+1}-t_i)^{-1}$. If $\inf\{t_{i+1}-t_i : i\in\mathbb{N}\} > 0$, then Jensen's inequality implies

$$\sup_{i\in\mathbb{N}} \mathbb{E}[|P_i|^3] \le \sup_{i\in\mathbb{N}} \mathbb{E}[|P_i|^4]^{3/4} < +\infty\,.$$

Hence, we obtain the implication "$(A_2)\Rightarrow(A_3)$", while the implication "$(A_3)\Rightarrow(A_0)$" is a consequence of Theorem 3.4 in [37]. The proof of Theorem 1.11 is finished.

5.5 Proof of Proposition 3.3

Without loss of any generality, we may and will assume that $\mathrm{Var}(Y)=1$ and $N\sim\mathcal{N}(0,1)$. Let $f:\mathbb{R}\to\mathbb{R}$ be 1-Lipschitz function and consider

$$g(x) = e^{x^2/2}\int_{-\infty}^x (f(y)-\mathbb{E}[f(N)])e^{-y^2/2}\,dy\ ,\quad x\in\mathbb{R}\,,$$

which satisfies the *Stein's equation*

$$g'(x) - xg(x) = f(x) - \mathbb{E}[f(N)] \tag{5.7}$$

as well as $\|g'\|_\infty \le \sqrt{2/\pi}$, $\|g''\|_\infty \le 2$, see *e.g.* Section 2.3 in [44]. In what follows, we fix such a pair (f,g) of functions. Let $G:\mathbb{R}\to\mathbb{R}$ be a differentiable function such that $G'=g$. Then due to $Y_t \overset{law}{=} Y$ and $Y\in L^4(\mathbb{P})$, one has

$$0 = \mathbb{E}[G(Y_t)-G(Y)] = \mathbb{E}[g(Y)(Y_t-Y) + \frac{1}{2}g'(Y)(Y_t-Y)^2] + \mathbb{E}[R_g]$$

with $|R_g| \le \dfrac{1}{6}\|g''\|_\infty\,|Y_t-Y|^3$. It follows that

$$0 = \mathbb{E}\left[g(Y)\times\frac{1}{t}\,\mathbb{E}[Y_t-Y|\mathscr{G}]\right] + \frac{1}{2}\,\mathbb{E}\left[g'(Y)\times\frac{1}{t}\,\mathbb{E}[(Y_t-Y)^2|\mathscr{G}]\right] + \frac{1}{t}\,\mathbb{E}[R_g].$$

By assumption (c) and as $t \downarrow 0$,

$$\left|\frac{1}{t}\, E[R_g]\right| \le \frac{1}{3t}\, E[|Y_t - Y|^3] \le \frac{1}{3}\sqrt{\frac{1}{t}\mathbb{E}[(Y_t-Y)^2]}\sqrt{\frac{1}{t}\mathbb{E}[(Y_t-Y)^4]}$$
$$\to \frac{1}{3}\sqrt{2\lambda + \mathbb{E}[S]}\sqrt{\varrho(Y)}$$

Therefore as $t \downarrow 0$, assumptions (a) and (b) imply that

$$0 = \lambda\mathbb{E}[g'(Y) - Yg(Y)] + \frac{1}{2}\,\mathbb{E}[g'(Y)S] + \lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}[R_g]\ .$$

The above equation shall be understood as "the limit $\lim_{t\downarrow 0} t^{-1}\mathbb{E}[R_g]$ exists and is equal to $-\lambda\mathbb{E}[g'(Y) - Yg(Y)] - \frac{1}{2}\,\mathbb{E}[g'(Y)S]$, bounded by $\frac{1}{3}\sqrt{2\lambda + \mathbb{E}[S]}\sqrt{\varrho(Y)}$.

Plugging this into the Stein's equation (5.7), we deduce the desired conclusion, namely

$$\begin{aligned}
d_{\mathcal{W}}(Y,N) &= \sup_{f\in \mathrm{Lip}(1)}\left|\mathbb{E}[f(Y) - f(N)]\right| \le \sup_{\substack{\|g'\|_\infty \le \sqrt{2/\pi}\\ \|g''\|_\infty\le 2}}\left|\mathbb{E}[g'(Y) - Yg(Y)]\right| \\
&\le \sup_{\substack{\|g'\|_\infty \le \sqrt{2/\pi}\\ \|g''\|_\infty\le 2}}\left(\frac{\|g'\|_\infty}{2\lambda}\mathbb{E}[|S|] + \left|\frac{1}{\lambda}\lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}[R_g]\right|\right) \\
&\le \frac{1}{\lambda\sqrt{2\pi}}\mathbb{E}[|S|] + \frac{\sqrt{2\lambda + \mathbb{E}[S]}}{3\lambda}\sqrt{\varrho(Y)}\ .
\end{aligned}$$

The general case follows from the fact that $d_{\mathcal{W}}(Y,N) = \sigma\, d_{\mathcal{W}}(Y/\sigma, N/\sigma)$ for $\sigma > 0$.

5.6 Proof of Proposition 3.5

By the same argument as in the proof of Theorem 3 in [24], we can assume $g \in C^\infty(\mathbb{R}^d)$ and define

$$f(x) = \int_0^1 \frac{1}{2t}\Big(\mathbb{E}[g(\sqrt{t}\,x + \sqrt{1-t}\,N)] - \mathbb{E}[g(N)]\Big)\,dt\ ,$$

which is a solution to the following *Stein's equation*

$$\langle x, \nabla f(x)\rangle - \langle \mathrm{Hess} f(x), \Sigma\rangle_{\mathrm{H.S.}} = g(x) - \mathbb{E}[g(N)]\ . \tag{5.8}$$

It is known that $M_r(f) \le r^{-1} M_r(g)$ for $r = 1,2,3$ and $\widetilde{M}_2(f) \le \frac{1}{2}\widetilde{M}_2(g)$. In particular, if Σ is positive definite, then $\widetilde{M}_2(f) \le \sqrt{2/\pi}\,\|\Sigma^{-1/2}\|_{\mathrm{op}}\, M_1(g)$ and $M_3(f) \le \sqrt{2\pi}\,\|\Sigma^{-1/2}\|_{\mathrm{op}}\, M_2(g)/4$, see [24, Lemma 2].

Again, it follows from the same arguments as in [24] that

$$\begin{aligned}
0 = \frac{1}{t}\,\mathbb{E}\bigg[\,\frac{1}{2}\Big\langle \mathrm{Hess} f(X), \Lambda^{-1}(X_t - X)(X_t-X)^T\Big\rangle_{\mathrm{H.S.}}\bigg] \\
+ \frac{1}{t}\mathbb{E}\Big[\,\langle \Lambda^{-1}(X_t - X), \nabla f(X)\rangle\Big] + \frac{1}{2t}\mathbb{E}[R],
\end{aligned} \tag{5.9}$$

where R is the error in the Taylor approximation satisfying

$$|R| \le \frac{1}{3}\|\Lambda^{-1}\|_{\mathrm{op}}\|X_t - X\|_2^3\beta \le \frac{\sqrt{d}}{3}\|\Lambda^{-1}\|_{\mathrm{op}}\beta\sqrt{\sum_{i=1}^d (X_{i,t}-X_i)^2}\sqrt{\sum_{i=1}^d (X_{i,t}-X_i)^4}\ ,$$

where $\beta := \min\{M_3(g)/3,\ \sqrt{2\pi}\|\Sigma^{-1/2}\|_{\mathrm{op}}\, M_2(g)/4\}$, and the last inequality follows from the elementary inequality $\|x-y\|_2^2 \le \sqrt{d}\big(\sum_{i=1}^d (x_i-y_i)^4\big)^{1/2}$ for $x, y \in \mathbb{R}^d$.

Notice meanwhile that the assumptions (a) and (b) imply that the limit $t^{-1}\mathbb{E}[R]$, as $t \downarrow 0$, is well defined and

$$\begin{aligned} -\lim_{t\downarrow 0}\frac{1}{2t}\mathbb{E}[R] &= \mathbb{E}\Big[\langle \mathrm{Hess} f(X), \Sigma\rangle_{\mathrm{H.S.}} - \langle X, \nabla f(X)\rangle\Big] + \frac{1}{2}\mathbb{E}\Big[\langle \mathrm{Hess} f(X), \Lambda^{-1}S\rangle_{\mathrm{H.S.}}\Big] \\ &= \mathbb{E}[g(N) - g(X)] + \frac{1}{2}\mathbb{E}\Big[\langle \mathrm{Hess} f(X), \Lambda^{-1}S\rangle_{\mathrm{H.S.}}\Big]\ , \end{aligned}$$

where the last equality comes from the definition of Stein's equation. Moreover, by assumption (c) and the above inequality, we have

$$\begin{aligned} \left|\lim_{t\to 0}\frac{1}{t}\,\mathbb{E}[R]\right| &\le \frac{\sqrt{d}}{3}\|\Lambda^{-1}\|_{\mathrm{op}}\beta\sqrt{\lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}\sum_{i=1}^d (X_{i,t}-X_i)^2}\sqrt{\lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}\sum_{i=1}^d (X_{i,t}-X_i)^4} \\ &= \frac{\sqrt{d}}{3}\|\Lambda^{-1}\|_{\mathrm{op}}\beta\sqrt{\sum_{i=1}^d 2\Lambda_{i,i}\Sigma_{i,i} + \mathbb{E}[S_{i,i}]}\sqrt{\sum_{i=1}^d \varrho_i(X)}, \end{aligned}$$

where the last equality follows from assumptions (b) and (c). To conclude our proof, it suffices to notice that $\mathbb{E}[\langle \mathrm{Hess} f(X), \Lambda^{-1}S\rangle_{\mathrm{H.S.}}]$ is bounded by

$$\min\left\{\frac{1}{2}\widetilde{M}_2(g),\ \sqrt{\frac{2}{\pi}}\,\|\Sigma^{-1/2}\|_{\mathrm{op}}\, M_1(g)\right\}\|\Lambda^{-1}\|_{\mathrm{op}}\,\mathbb{E}[\|S\|_{\mathrm{H.S.}}]\ .$$

Acknowledgements. We thank Ivan Nourdin and Giovanni Peccati for their helpful comments and stimulating discussions.

References

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

[2] S. Bai and M. S. Taqqu. The universality of homogeneous polynomial forms and critical limits. *J. Theoret. Probab.*, 29(4):1710–1727, 2016.

[3] N. Bouleau and L. Denis. *Dirichlet forms methods for Poisson point measures and Lévy processes, With emphasis on the creation-annihilation techniques*, volume 76 of *Probability Theory and Stochastic Modelling*. Springer, 2015.

[4] S. Bourguin and G. Peccati. Semicircular limits on the free poisson chaos: counterexamples to a transfer principle. *J. Funct. Anal.*, 267(4):963–997, 2014.

[5] S. Campese, I. Nourdin, G. Peccati, and G. Poly. Multivariate Gaussian approximations on Markov chaoses. *Electron. Commun. Probab.*, 21:Paper No. 48, 9, 2016.

[6] S. Chatterjee, J. Fulman, and A. Röllin. Exponential approximation by Stein's method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.*, 8:197–223, 2011.

[7] S. Chatterjee and E. Meckes. Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:257–283, 2008.

[8] S. Chatterjee and Q.-M. Shao. Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model. *Ann. Appl. Probab.*, 21(2):464–483, 2011.

[9] P. Diaconis. The distribution of leading digits and uniform distribution mod 1. *Ann. Probability*, 5(1):72–81, 1977.

[10] C. Döbler. Stein's method of exchangeable pairs for the Beta distribution and generalizations. *Electron. J. Probab.*, 20:no. 109, 1–34, 2015.

[11] C. Döbler and K. Krokowski. On the fourth moment condition for Rademacher chaos. `arXiv:1706.00751`*, to appear in: Ann. Inst. Henri Poincaré Probab. Stat.*, 2017.

[12] C. Döbler and G. Peccati. Quantiative de Jong theorems in any dimension. *Electron. J. Probab.*, 22:no. 2, 1–35, 2017.

[13] C. Döbler and G. Peccati. The fourth moment theorem on the Poisson space. `arXiv:1701.03120`*, to appear in: Ann. Probab.*, 2017.

[14] C. Döbler and M. Stolz. Stein's method and the multivariate CLT for traces of powers on the classical compact groups. *Electron. J. Probab.*, 16:no. 86, 2375–2405, 2011.

[15] P. Eichelsbacher and M. Löwe. Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.*, 15:no. 30, 962–988, 2010.

[16] P. Eichelsbacher and C. Thäle. New Berry-Esseen bounds for non-linear functionals of Poisson random measures. *Electron. J. Probab.*, 19:no. 102, 25, 2014.

[17] K. Krokowski, A. Reichenbachs, and C. Thäle. Berry-Esseen bounds and multivariate limit theorems for functionals of Rademacher sequences. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(2):763–803, 2016.

[18] K. Krokowski, A. Reichenbachs, and C. Thäle. Discrete Malliavin–Stein method: Berry–Esseen bounds for random graphs and percolation. *Ann. Probab.*, 45(2):1071–1109, 2017.

[19] G. Last. Stochastic analysis for Poisson processes. In G. Peccati and M. Reitzner, editors, *Stochastic analysis for Poisson point processes*, Mathematics, Statistics, Finance and Economics, chapter 1, pages 1–36. Bocconi University Press and Springer, 2016.

[20] G. Last, G. Peccati, and M. Schulte. Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. *Probab. Theory Related Fields*, 165(3-4):667–723, 2016.

[21] G. Last and M. Penrose. *Lectures on the Poisson Process*. IMS Textbooks. Cambridge University Press, Cambridge, 2017.

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

[23] E. Meckes. An infinitesimal version of stein's method of exchangeable pairs. *Ph.d Dissertation, Stanford University*, 2006.

[24] E. Meckes. On Stein's method for multivariate normal approximation. In *High dimensional probability V: the Luminy volume*, volume 5 of *Inst. Math. Stat. Collect.*, pages 153–178. Inst. Math. Statist., Beachwood, OH, 2009.

[25] S. Noreddine and I. Nourdin. On the Gaussian approximation of vector-valued multiple integrals. *J. Multivariate Anal.*, 102(6):1008–1017, 2011.

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

[27] I. Nourdin and G. Peccati. *Normal approximations with Malliavin calculus*, volume 192 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2012. From Stein's method to universality.

[28] I. Nourdin, G. Peccati, G. Poly, and R. Simone. Classical and free fourth moment theorems: universality and thresholds. *J. Theoret. Probab.*, 29(2):653–680, 2016.

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

[30] I. Nourdin, G. Peccati, and G. Reinert. Stein's method and stochastic analysis of Rademacher functionals. *Electron. J. Probab.*, 15:no. 55, 1703–1742, 2010.

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

[32] I. Nourdin and G. Zheng. Exchangeable pairs on Wiener chaos. `arXiv:1704.02164`, 2017.

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

[34] G. Peccati, J. L. Solé, M. S. Taqqu, and F. Utzet. Stein's method and normal approximation of Poisson functionals. *Ann. Probab.*, 38(2):443–478, 2010.

[35] 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.

[36] G. Peccati and C. Zheng. Multi-dimensional Gaussian fluctuations on the Poisson space. *Electron. J. Probab.*, 15:no. 48, 1487–1527, 2010.

[37] G. Peccati and C. Zheng. Universal Gaussian fluctuations on the discrete Poisson chaos. *Bernoulli*, 20(2):697–715, 2014.

[38] G. Reinert and A. Röllin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.*, 37(6):2150–2173, 2009.

[39] Y. Rinott and V. Rotar. On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics. *Ann. Appl. Probab.*, 7(4):1080–1105, 1997.

[40] A. Röllin. A note on the exchangeability condition in Stein's method. *Statist. Probab. Lett.*, 78(13):1800–1806, 2008.

[41] M. Schulte. Normal approximation of Poisson functionals in Kolmogorov distance. *J. Theoret. Probab.*, 29(1):96–117, 2016.

[42] Q.-M. Shao and Z.-G. Su. The Berry-Esseen bound for character ratios. *Proc. Amer. Math. Soc.*, 134(7):2153–2159 (electronic), 2006.

[43] Ch. Stein. Approximate computation of expectations. In *Institute of Mathematical Statistics Lecture Notes - Monograph Series*, volume 7. Institute of Mathematical Statistics, 1986.

[44] G. Zheng. Normal approximation and almost sure central limit theorem for non-symmetric Rademacher functionals. *Stochastic Process. Appl.*, 127(5):1622–1636, 2017.

This page is left blank

Paper 5: A Peccati-Tudor type theorem for Rademacher chaoses

Guangqu Zheng*

submitted

Abstract

In this article, we prove that in the Rademacher setting, a random vector with chaotic components is close in distribution to a centred Gaussian vector, if both the *maximal influence* of the associated kernel and the fourth cumulant of each component is small. In particular, we recover the univariate case recently established in Döbler and Krokowski (2017).

Our main strategy consists in a novel adaption of the exchangeable pairs couplings initiated in Nourdin and Zheng (2017), as well as its combination with estimates via chaos decomposition.

1 Introduction

1.1 Motivation

Nualart and Peccati's fourth moment theorem states that a normalised sequence of fixed-order multiple Wiener-Itô integrals associated to a Brownian motion converges in law to the standard Gaussian if and only if the corresponding fourth moment converges to 3. It was proved in [21] using the Dambis-Dubins-Schwartz random-time change technique. Soon after the appearance of [21], several extensions have been made, among which the paper [23] by Peccati and Tudor provided a significant multivariate extension using the same techique. Roughly speaking, a sequence of chaotic random vectors on the Wiener space converges in distribution to a centred Gaussian vector with matched covariance matrix if and only if the asymptotic normality holds true for each component. Note that the necessary condition boils down to the convergence of the fourth moments due to the fourth moment theorem of Nualart and Peccati.

In 2009, Nourdin and Peccati [15] combined the Malliavin calculus and Stein's method of normal approximation so as to literally create a new field of research, known as the *Malliavin-Stein approach*. One of its many highlights is the obtention of the (quantitative) fourth moment theorem in the total-variation distance. Here is the bound quoted from the monograph [16]:

2010 *Mathematics Subject Classification*. Primary: 60F05, 60B12; Secondary: 47N30.

Key words and phrases. Fourth moment theorem; Rademacher chaos; Stein's method; exchangeable pairs; spectral decomposition; maximal influence.

*Email: guangqu.zheng@uni.lu

given a normalised q-th Wiener-Itô integral F associated to a Brownian motion, one has

$$d_{\mathrm{TV}}(F,Z) := \sup_{A\in\mathscr{B}(\mathbb{R})}\Big|\mathbb{P}(F\in A)-\mathbb{P}(Z\in A)\Big| \leq \frac{2}{\sqrt{3}}\sqrt{\frac{q-1}{q}(\mathbb{E}[F^4]-3)}\,,$$

where Z is a standard Gaussian random variable and $\mathscr{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$. As an immediate consequence, the fourth moment theorem of Nualart and Peccati follows.

The success of the Malliavin-Stein approach stems from the integration by parts on both sides, namely, the Stein's lemma within the Stein's method and the duality relation between Malliavin derivative and Skorohod divergence on a Gaussian space, see the monograph [16] for a comprehensive treatment. The only ingredients required from the Stein's method are the Stein's lemma, Stein's equation and the regularity properties of the Stein's solution, while "exchangeable pairs", another fundamental tool and notable cornerstone of Stein's method, had not been touched until the recent investigation [20] made by Nourdin and Zheng. They constructed infinitely many exchangeable pairs of Brownian motions and combined them with E. Meckes' abstract results [12, 13] on exchangeable pairs to recover the quantitative fourth moment theorem on a Gaussian space in any dimension. Such an elementary strategy was soon adapted by Döbler, Vidotto and Zheng in [7] for their investigation on the Poisson space, and they were able to obtain the quantitative fourth moment theorem in any dimension. In fact, the univariate fourth moment theorem on the Poisson space was established earlier in [6] under some *integrability* assumptions involving the difference operator, which are partially due to the inherent discreteness of the Poisson space. Remarkably, the authors of [7] were able to obtain the exact fourth moment theorem under the *weakest possible* assumption of finite fourth moment. This illustrates the power of the elementary exchangeable pairs approach.

In this work, under *suitable* assumptions, we establish a Peccati-Tudor type theorem in the Rademacher setting using the elementary exchangeable pairs approach.

1.2 Main result

We first fix a rich probability space $(\Omega,\mathcal{F},\mathbb{P})$, on which our random objects are defined. Let $\mathbb{E}$ be the associated expectation operator.

We write $\mathbb{N} := \{1,2,\dots\}$ and denote by $\mathbf{X}$ a sequence of independent Rademacher random variables $(X_k, k\in\mathbb{N})$ such that $\mathbb{P}(X_k=1)=p_k=1-q_k=1-\mathbb{P}(X_k=-1)\in(0,1)$. We call it the *symmetric* case, whenever $p_k=1/2$ for each $k\in\mathbb{N}$; otherwise, we call it the *general* case. We write $\mathbf{Y}=(Y_k,k\in\mathbb{N})$ for the *normalised version* of $\mathbf{X}$, that is,

$$Y_k = \frac{X_k - p_k + q_k}{2\sqrt{p_k q_k}}\,,\quad k\in\mathbb{N}\,. \tag{1.1}$$

We write $\mathfrak{H}=\ell^2(\mathbb{N})$, equipped with usual ℓ^2-norm and for $p\in\mathbb{N}$, $\mathfrak{H}^{\otimes p}$ means the p-th tensor product of $\mathfrak{H}$ and $\mathfrak{H}^{\odot p}$ its symmetric subspace. We denote $\mathfrak{H}_0^{\odot p} := \{f\in\mathfrak{H}^{\odot p} : f|_{\triangle_p^c}=0\}$ with $\triangle_p=\{(i_1,\dots,i_p)\in\mathbb{N}^p : i_k\neq i_j \text{ for different } k,j\}$. Clearly, $\mathfrak{H}_0^{\odot 0}=\mathfrak{H}^{\otimes 0}=\mathbb{R}$ and $\mathfrak{H}_0^{\odot 1}=\mathfrak{H}$.

Let $f\in\mathfrak{H}_0^{\odot d}$ with $d\in\mathbb{N}$ and $\Xi=(\xi_k,k\in\mathbb{N})$ be a generic sequence of independent normalised random variables. We define the following homogeneous sum with order d, based on the kernel

f, by setting,

$$Q_d(f;\Xi) := \sum_{i_1,\dots,i_d\in\mathbb{N}} f(i_1,\dots,i_d)\xi_{i_1}\cdots\xi_{i_d} \tag{1.2}$$

and in particular, $Q_d(f;\mathbf{Y})$ is called the (discrete) multiple integral of f. We write $C_d = \{Q_d(f;\mathbf{Y}) : f \in \mathfrak{H}_0^{\odot d}\}$ and call it the d-th Rademacher chaos, and as a convention, we put $C_0 = \mathbb{R}$. In case of no ambiguity, we will simple write $Q_d(f)$ for $Q_d(f;\mathbf{Y})$.

Let us introduce an important notion before we state our main result: for a given kernel $f \in \mathfrak{H}_0^{\odot d}$, we denote by $\mathcal{M}(f)$ the *maximal influence* of f, namely

(1.3)
$$\mathcal{M}(f) := \sup_{k\in\mathbb{N}} \sum_{i_1,\dots,i_{d-1}\in\mathbb{N}} f(i_1,\dots,i_{d-1},k)^2 \quad \text{for } d\geq 2 \quad \text{and} \quad \mathcal{M}(f) := \sup_{k\in\mathbb{N}} f(k)^2 \quad \text{for } d=1.$$

This notion is adapted from the boolean analysis (see *e.g.* [22]), in which the class of low-influence functions is often what is interesting or necessary in practice. It is also closely related to the invariance principle established in [14] and the universality phenomenon of Gaussian Wiener chaos [18]. See also Section 4 for more details.

In this work, we are mainly concerned with random variables in a Rademacher chaos and random vectors with components in Rademacher chaoses. More precisely, we establish the following result.

Theorem 1.1. Fix integers $d \geq 2$ and $1 \leq q_1 \leq \ldots \leq q_d$, and consider the sequence of random vectors

$$F^{(n)} = (F_1^{(n)},\dots,F_d^{(n)})^T := (Q_{q_1}(f_{1,n}),\dots,Q_{q_d}(f_{d,n}))^T$$

with kernels $f_{j,n}$ in $\mathfrak{H}_0^{\odot q_j}$ for each $n \in \mathbb{N}, j \in \{1,\dots,d\}$. Assume that the covariance matrix Σ_n of $F^{(n)}$ converges in Hilbert-Schmidt norm to a nonnegative definite symmetric matrix $\Sigma = (\Sigma_{i,j}, 1 \leq i,j \leq d)$, as $n \to +\infty$. Suppose that the following condition holds:

$$\lim_{n\to+\infty} \sum_{j=1}^d \mathcal{M}(f_{j,n}) = 0 \ .$$

If for each $j \in \{1,\dots,d\}$, $\mathbb{E}\Big[(F_j^{(n)})^4\Big]$ converges to $3\Sigma_{j,j}^2$, as $n \to +\infty$, then $F^{(n)}$ converges in distribution to $Z \sim \mathcal{N}(0,\Sigma)$, as $n \to +\infty$.

The above theorem is analogous to the Peccati-Tudor theorem on a Gaussian space [23], so we call it a Peccati-Tudor type theorem, which explains our title. One of the main tools we need for the proof is the following ingredient from Stein's method of exchangeable pairs. As one will see easily, we can obtain a quantitative version of Theorem 1.1, which will be an analogue to [7, Theorem 1.7] and left for interested readers.

Recall first that two random variables W and W', defined on a common probability space, are said to form an exchangeable pair, if (W, W') has the same distribution as (W', W).

Proposition 1.1 (Proposition 3.5 in [7])**.** For each $t > 0$, let (F, F_t) be an exchangeable pair of centred d-dimensional random vectors defined on a common probability space. Let $\mathscr{G}$ be a σ-algebra that contains $\sigma\{F\}$. Assume that $\Lambda \in \mathbb{R}^{d\times d}$ is an invertible deterministic matrix and Σ is a symmetric, non-negative definite deterministic matrix such that

(a) $\lim_{t\downarrow 0} \frac{1}{t}\,\mathbb{E}[F_t - F|\mathscr{G}] = -\Lambda F$ in $L^1(\Omega)$,

(b) $\lim_{t\downarrow 0} \frac{1}{t}\,\mathbb{E}[(F_t - F)(F_t - F)^T|\mathscr{G}] = 2\Lambda\Sigma + S$ in $L^1(\Omega, \|\cdot\|_{\mathrm{H.S.}})$ for some matrix $S = S(F)$, and with $\|\cdot\|_{\mathrm{H.S.}}$ the Hilbert-Schmidt norm,

(c) for each $i \in \{1, \ldots, d\}$, there exists some real number $\rho_i(F)$ such that $\lim_{t\downarrow 0} \frac{1}{t}\,\mathbb{E}[(F_{i,t} - F_i)^4] = \rho_i(F)$, where $F_{i,t}$ (resp. F_i) stands for the i-th coordinate of F_t (resp. F).

Then, for $g \in C^3(\mathbb{R}^d)$ such that $g(F), g(Z) \in L^1(\mathbb{P})$, we have, with $Z \sim \mathcal{N}(0, \Sigma)$,

$$\begin{aligned}&\left|\mathbb{E}[g(F)] - \mathbb{E}[g(Z)]\right|\\ &\leq \frac{\|\Lambda^{-1}\|_{\mathrm{op}}\sqrt{d}\,M_2(g)}{4}\mathbb{E}\left[\sqrt{\sum_{i,j=1}^d S_{i,j}^2}\right] + \frac{\sqrt{d}M_3(g)\|\Lambda^{-1}\|_{\mathrm{op}}}{18}\sqrt{\sum_{i=1}^d 2\Lambda_{i,i}\Sigma_{i,i} + \mathbb{E}[S_{i,i}]}\sqrt{\sum_{i=1}^d \rho_i(F)}\,,\end{aligned}$$

where $M_k(g) := \sup_{x\in\mathbb{R}^d}\left\|D^k g(x)\right\|_{\mathrm{op}}$ with $\|\cdot\|_{\mathrm{op}}$ the operator norm.

The rest of this paper is organised as follows: Section 1.3 is devoted to a brief overview of related results and we sketch our strategy of proving Theorem 1.1 in Section 1.4; in Section 2, we provide preliminary knowledge on Rademacher chaos and a crucial *exchangeable pairs coupling*. The proof of our main result will be given in Section 3 and some discussion about universality around Rademacher chaos will be presented in Section 4.

1.3 A brief overview of literature

Soon after the appearance of [15], Nourdin, Peccati and Reinert combined Stein's method and a discrete version of Malliavin calculus to study the Gaussian approximation of Rademacher functionals in the symmetric case. This analysis is known as the *discrete Malliavin-Stein approach*. It has been generalised by the authors of [9, 10] not only in the multivariate setting but also in the general case where functionals involving non-symmetric, non-homogeneous Rademacher random variables were investigated. Recently, Döbler and Krokowski [5] gave the following *fourth-moment-influence* bound and pointed out that it is *optimal* in the sense that there are examples, in which the fourth moment condition alone would not guarantee the asymptotic normality.

Theorem 1.2 (Theorem 1.1 in [5])**.** Fix $p \in \mathbb{N}$ and $f \in \mathfrak{H}_0^{\odot p}$ satisfying $p!\|f\|^2_{\mathfrak{H}^{\otimes p}} = 1$. Let Z be a standard Gaussian and $F = Q_p(f; \mathbf{Y}) \in L^4(\mathbb{P})$, then we have the following bound in Wasserstein distance:

$$d_{\mathrm{W}}(F, Z) := \sup_{\|h'\|_\infty \leq 1}\left|\mathbb{E}[h(F) - h(Z)]\right| \leq C_1\sqrt{\left|\mathbb{E}[F^4] - 3\right|} + C_2\sqrt{\mathcal{M}(f)}\,,$$

where C_1, C_2 are two numerical constants. This result echoes the remarkable de Jong's central limit theorem [4].

Besides the aforementioned references, Krokowski [8] derived a multiplication formula that generalises the one in [17], and applying as well the Chen-Stein's method, he studied the Poisson approximation of Rademacher functionals. Independently, Privault and Torrisi [26] also derived a multiplication formula and moreover, they obtained a generalisation of the approximate chain rule from [17], and applied them to study Gaussian and Poisson approximation of Rademacher functionals in the general case. Concerning the normal approximation in [17] or [26], the authors were only able to obtain the bounds in some "smooth-version" distance, due to regularity involving in their chain rules and Stein's solution. In a follow-up work, Zheng [28] obtained a neater chain rule that requires *minimal* regularity (see [28, Remark 2.3]), from which he obtained the bound in Wasserstein distance as well as an almost sure central limit theorem for Rademacher chaos. It is worthy pointing out that without using any chain rule, the authors of [9, 10] used carefully a representation of the discrete Malliavin gradient and the fundamental theorem of calculus to deduce the Berry-Esseen bound for normal approximation. Using similar ideas, Döbler and Krokowski [5] also provided the Berry-Esseen bound for their *fourth-moment-influence* theorem, which is of the same order as the above Wasserstein bound.

1.4 Strategy of proving Theorem 1.1

Stein's method of exchangeable pairs was first systematically presented in Charles Stein's 1986 monograph [27], which was subsequently developed and ramified by many authors. Concerning our work, we mention in particular E. Meckes' dissertation [12], in which she developed an infinitesimal version of this method to obtain total-variation bound in normal approximation. This infinitesimal version of Stein's method of exchangeable pairs was later generalised in [3, 13] for the multivariate normal approximation.

As announced, Proposition 1.1 is one of our main tools, and it can be seen as a generalisation of [13]. To use it, we need to construct a suitable family of random vectors F_t, $t \geq 0$ such that (F_t, F) is exchangeable for each t and satisfies several *asymptotic regression* conditions. In fact, we will first construct a family of Rademacher sequences $\mathbf{X}^t$ such that $(\mathbf{X}^t, \mathbf{X})$ is an exchangeable pair of $\{\pm 1\}^{\mathbb{N}}$-valued random variables for each $t \geq 0$. More precisely, let $\mathbf{X}'$ be an independent copy of $\mathbf{X}$ and $\Theta = (\theta_k, k \in \mathbb{N})$ be a sequence of *i.i.d.* standard exponential random variables such that $\mathbf{X}$, $\mathbf{X}'$ and Θ are independent. For each $t \in [0, +\infty)$, we define

$$X_k^t := X_k \mathbf{1}_{(\theta_k \geq t)} + X_k' \mathbf{1}_{(\theta_k < t)} \ .$$

It has been pointed out in [10] that $\mathbf{X}^t$ has the same distribution as $\mathbf{X}$, see also Remark 3.4 in [17] for the symmetric case. However, both of these two articles did not explicitly state the exchangeability of $\mathbf{X}^t$ and $\mathbf{X}$, which will be proved in Lemma 2.2. Assuming this and writing $F = \mathfrak{f}(\mathbf{X})$ for some representative $\mathfrak{f} : \{\pm 1\}^{\mathbb{N}} \to \mathbb{R}^d$, we can set $F_t = \mathfrak{f}(\mathbf{X}^t)$. It is easy to see that the exchangeability can be passed to (F, F_t) now. If $F = (Q_{p_1}(f_1; \mathbf{Y}), \ldots, Q_{p_d}(f_d; \mathbf{Y}))$, then we can write $F_t = (Q_{p_1}(f_1; \mathbf{Y}^t), \ldots, Q_{p_d}(f_d; \mathbf{Y}^t))$ with $\mathbf{Y}^t$ the normalised version of $\mathbf{X}^t$ in the sense of (1.1).

Moreover, this exchangeable pairs coupling fits well with the Mehler's formula, which gives a nice representation of the discrete *Ornstein-Uhlenbeck semigroup* $(P_t, t \geq 0)$: given

$F \in L^2(\Omega, \sigma\{\mathbf{X}\}, \mathbb{P})$, we can first write $F = \mathfrak{f}(\mathbf{X})$ for some $\mathfrak{f} : \{\pm 1\}^{\mathbb{N}} \to \mathbb{R}$, then the *Mehler formula* ([10, Proposition 3.1]) states that

$$P_t F = \mathbb{E}\Big[\mathfrak{f}(\mathbf{X}^t) \,|\, \sigma\{\mathbf{X}\}\Big]. \tag{1.4}$$

For $\xi \in C_p$, as we will see in Section 2, $P_t\xi = e^{-pt}\xi$, then the asymptotic linear regression (a) in Proposition 1.1 follows easily, and with slightly more effort, the higher order regressions can also be obtained, see Proposition 2.1.

Another important ingredient in our proof is Ledoux's spectral point-of-view for fourth moment theorem [11], which was later refined *e.g.* in [1, 2]. Such a spectral viewpoint helps one get rid of some computational deadlock that is usually caused by the complicated multiplication formula. In particular, our proof is motivated by some arguments in [2].

As a byproduct of our strategy, we will provide a short proof of Theorem 1.2 in the beginning of Section 3. Some estimate from this proof will also be helpful for our multivariate case.

Acknowledgement. Part of this work was done during a visit at National University of Singapore. I thank very much Professor Louis H. Y. Chen at NUS for his very generous support and kind hospitality. The gratitude also goes to Professor Giovanni Peccati for sharing his alternative proof of Lemma 2.4 in [6], which motived our proof of Lemma 2.1.

2 Preliminaires

Denote by $\sigma\{\mathbf{X}\}$ the σ-algebra generated by the sequence $\mathbf{X}$, and note that $\sigma\{\mathbf{X}\} = \sigma\{\mathbf{Y}\}$. The *Wiener-Itô-Wash chaos decomposition* asserts that any random variable $F \in L^2(\Omega, \sigma\{\mathbf{X}\}, \mathbb{P})$ admits a unique representation

$$F = \mathbb{E}[F] + \sum_{p\geq 1} Q_p(f_p) \quad \text{with } f_p \in \mathfrak{H}_0^{\odot p} \text{ for each } p \in \mathbb{N}, \tag{2.1}$$

where the above series converges in $L^2(\mathbb{P})$. We denote by $J_k(\cdot)$ the projection onto the k-th Rademacher chaos C_k: for F given in (2.1), $J_p(F) = Q_p(f_p)$ for each $p \in \mathbb{N}$, and $J_0(F) = \mathbb{E}[F]$. It is not difficult to check that for $f \in \mathfrak{H}_0^{\odot p}$ and $g \in \mathfrak{H}_0^{\odot q}$, it holds that

$$\mathbb{E}[Q_p(f)Q_q(g)] = \mathbf{1}_{\{p=q\}} p! \langle f, g\rangle_{\mathfrak{H}^{\otimes p}}.$$

This is known as the *orthogonality* property of the multiple integrals. One can refer to N. Privault's survey [25] for more details and relevant discrete Malliavin calculus.

The authors of [17] established a multiplication formula for discrete multiple integrals in the symmetric case: given $f \in \mathfrak{H}_0^{\odot p}$ and $g \in \mathfrak{H}_0^{\odot q}$, one has

$$Q_p(f)Q_q(g) = \sum_{r=0}^{p\wedge q} r! \binom{p}{r}\binom{q}{r} Q_{p+q-2r}\big(f\widetilde{\otimes}_r g \mathbf{1}_{\Delta_{p+q-2r}}\big), \tag{2.2}$$

where the *r-contraction* $f \otimes_r g$ of f and g is defined by

$$(f\otimes_r g)(i_1,\dots,i_{p-r},j_1,\dots,j_{q-r}) := \sum_{k_1,\dots,k_r\in\mathbb{N}} f(i_1,\dots,i_{p-r},k_1,\dots,k_r)\cdot g(j_1,\dots,j_{q-r},k_1,\dots,k_r)$$

and $f\widetilde{\otimes}_r g$ is the canonical symmetrisation of $f \otimes_r g$, *i.e.* for any $h \in \mathfrak{H}^{\otimes p}$, $\widetilde{h}$ is given by

$$\widetilde{h}(i_1, \ldots, i_p) = \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} h(i_{\sigma(1)}, \ldots, i_{\sigma(p)}) ,$$

with $\mathfrak{S}_p$ the permutation group over $\{1, \ldots, p\}$. We follow the convention that $\widetilde{c} = c$ for each $c \in \mathbb{R}$. Note it is easy to deduce from the Cauchy-Schwarz inequality that $\|\widetilde{h}\|_{\mathfrak{H}^{\otimes p}} \leq \|h\|_{\mathfrak{H}^{\otimes p}}$ for each $h \in \mathfrak{H}^{\otimes p}$, then applying the above orthogonality property and mathematical induction gives us a weak form of the *hypercontractivity* property in the symmetric case, namely, $\mathbb{E}[|F|^r] < +\infty$ for any $F \in C_p$, $p, r \in \mathbb{N}$.

However, in the general case, one can *not* even guarantee the existence of finite fourth moment of a generic multiple integral. Such a phenomenon, due to the asymmetry, is also revealed in the corresponding multiplication formulae, see Proposition 2.2 in [8] and Proposition 5.1 in [26]. As already pointed out in [5], given $F \in C_p \cap L^4(\mathbb{P})$, one can *not* directly deduce from these multiplication formulae that F^2 admits a finite chaotic decomposition. Adapting the induction arguments from the proof of [6, Lemma 2.4], Döbler and Krokowski gave the following positive result.

Lemma 2.1 (Lemma 2.3 in [5])**.** Let $F = Q_p(f) \in L^4(\mathbb{P})$ and $G = Q_q(g) \in L^4(\mathbb{P})$ for some $f \in \mathfrak{H}_0^{\odot p}$ and $g \in \mathfrak{H}_0^{\odot q}$. Then $FG \in L^2(\mathbb{P})$ admits a finite chaos decomposition of the form

$$FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q-1} J_k(FG) + Q_{p+q}(f\widetilde{\otimes} g \mathbf{1}_{\Delta_{p+q}}) .$$

In particular, if $Q_1(h)$ belongs to $L^4(\mathbb{P})$ for some $h \in \mathfrak{H}$, then

$$Q_1(h)^2 = \|h\|_{\mathfrak{H}}^2 + Q_1(w) + Q_2(h\widetilde{\otimes} h \mathbf{1}_{\Delta_2}) \quad \text{with } w(k) = \frac{h(k)^2(q_k - p_k)}{\sqrt{p_k q_k}}, k \in \mathbb{N}.$$

(*As this lemma is crucial for our work and for the sake of completeness, we provide in Section 3.3 another and direct proof suggested by Giovanni Peccati.*)

2.1 Ornstein-Uhlenbeck Structure and carré du champs operator

Denote by $\mathrm{dom}(L)$ the set of those F in (2.1) verifying

$$\sum_{p=1}^{\infty} p^2 \mathbb{E}[Q_p(f_p)^2] = \sum_{p=1}^{\infty} p^2 p! \|f_p\|_{\mathfrak{H}^{\otimes p}}^2 < +\infty .$$

For such a $F \in \mathrm{dom}(L)$, we define $LF = -\sum_{p\geq 1} pQ_p(f_p)$. In particular, if $F \in C_p$, $LF = -pF$. In other words, $-L$ has pure spectrum $\mathbb{N} \cup \{0\}$ and each eigenvalue $p \in \{0\} \cup \mathbb{N}$ corresponds to the eigenspace C_p. And we call L the *Ornstein-Uhlenbeck operator*, equipped with its domain $\mathrm{dom}(L)$.

For $F, G \in \mathrm{dom}(L)$ such that $FG \in \mathrm{dom}(L)$, we define the *carré du champs* operator $\Gamma(F, G)$ by setting

$$\Gamma(F, G) := \frac{1}{2}(L(FG) - FLG - GLF) .$$

In particular, for F, G as in Lemma 2.1, one has $FG \in \mathrm{dom}(L)$ and

$$\Gamma(F,G) = \frac{1}{2}[(p+q)+L]\left(\sum_{k=0}^{p+q} J_k(FG)\right) = \frac{p+q}{2}\mathbb{E}[FG] + \sum_{k=1}^{p+q-1} \frac{p+q-k}{2} J_k(FG)\,, \tag{2.3}$$

and as a consequence of the orthogonality property, one deduces that

$$\mathrm{Var}(\Gamma(F,G)) = \sum_{k=1}^{p+q-1} \frac{(p+q-k)^2}{4}\mathrm{Var}(J_k(FG)) \le \max\{p^2,q^2\} \sum_{k=1}^{p+q-1} \mathrm{Var}(J_k(FG))\,, \tag{2.4}$$

which is all we need about the carré du champs.

For each $t \in [0,+\infty)$ and F as in (2.1), we define

$$P_t F := \mathbb{E}[F] + \sum_{p=1}^{\infty} e^{-pt} Q_p(f_p)\ .$$

$(P_t, t \ge 0)$ is called the Ornstein-Uhlenbeck semigroup, which can be represented alternatively by the Mehler formula (1.4). To verify (1.4), one can first consider $F = Q_p(f_p)$ in a Rademacher chaos with $f_p \in \mathfrak{H}_0^{\odot p}$ having finite support and then use the standard approximation argument. Note that for $F \in \mathrm{dom}(L)$, it is not difficult to check $t^{-1}(P_tF - F)$ converges in $L^2(\mathbb{P})$ to LF, as $t \downarrow 0$.

2.2 Exchangeable pairs of Rademacher sequences

Lemma 2.2. Let $\mathbf{X}^t$ and $\mathbf{X}$ be given as before, then $(\mathbf{X}, \mathbf{X}^t)$ has the same distribution as $(\mathbf{X}^t, \mathbf{X})$. In particular, for any $f_j \in \mathfrak{H}_0^{\odot p_j}$ with $p_j \in \mathbb{N}$, $j = 1, \ldots, d$,

$$(Q_{p_1}(f_1;\mathbf{Y}), \ldots, Q_{p_d}(f_d;\mathbf{Y})) \text{ and } (Q_{p_1}(f_1;\mathbf{Y}^t), \ldots, Q_{p_d}(f_d;\mathbf{Y}^t))$$

form an exchangeable pair, where $\mathbf{Y}^t$ stands for the normalised version of $\mathbf{X}^t$ in the sense of (1.1).

Proof. Note first that $\mathbf{X}^t$ is a sequence of independent Rademacher random variables for each $t \in [0,+\infty)$. For each $k \in \mathbb{N}$, it is easy to check that

$$\mathbb{P}(X_k^t = -1, X_k = 1) = \mathbb{P}(X_k^t = 1, X_k = -1) = (1-e^{-t})p_k q_k\ .$$

This gives us the exchangeability of (X_k, X_k^t) for each $k \in \mathbb{N}$. Let $\mathbf{a} = (a_i, i \in \mathbb{N}), \mathbf{b} = (b_i, i \in \mathbb{N}) \in \{\pm 1\}^{\mathbb{N}}$, then using the independence within those two sequences $\mathbf{X}, \mathbf{X}^t$, we obtain

$$\begin{aligned}
\mathbb{P}(\mathbf{X} = \mathbf{a}, \mathbf{X}^t = \mathbf{b}) &= \prod_{k\in\mathbb{N}} \mathbb{P}(X_k = a_k, X_k^t = b_k) \\
&= \prod_{k\in\mathbb{N}} \mathbb{P}(X_k = b_k, X_k^t = a_k) \quad \text{by exchangeability of } X_k, X_k^t \\
&= \mathbb{P}(\mathbf{X} = \mathbf{b}, \mathbf{X}^t = \mathbf{a})\ .
\end{aligned}$$

This proves the exchangeability of $\mathbf{X}, \mathbf{X}^t$. The rest follows from a standard approximation argument: it is clear that after truncation, (with $[N] := \{1, \dots, N\}$)

$$(Q_{p_1}(f_1 \mathbf{1}_{[N]^{p_1}}; \mathbf{Y}), \dots, Q_{p_d}(f_d \mathbf{1}_{[N]^{p_d}}; \mathbf{Y})) \text{ and } (Q_{p_1}(f_1 \mathbf{1}_{[N]^{p_1}}; \mathbf{Y}^t), \dots, Q_{p_d}(f_d \mathbf{1}_{[N]^{p_d}}; \mathbf{Y}^t))$$

form an exchangeable pair; letting $N \to +\infty$ and keeping in mind that the exchangeability is preserved in limit, we get the desired result. □

The following result brings more connections between our exchangeable pairs and Ornstein-Uhlenbeck operator.

Proposition 2.1. Let $F = Q_p(f; \mathbf{Y}) \in L^4(\mathbb{P})$ for some $f \in \mathfrak{H}_0^{\odot p}$ and define $F_t = Q_p(f; \mathbf{Y}^t)$. Then, (F, F_t) is an exchangeable pair for each $t \in \mathbb{R}_+$. Moreover,

(a) $\lim_{t \downarrow 0} \frac{1}{t} \mathbb{E}[F_t - F | \sigma\{\mathbf{X}\}] = LF = -pF$ in $L^4(\mathbb{P})$.

(b) If $G = Q_q(g; \mathbf{Y}) \in L^4(\mathbb{P})$ and $G_t = Q_q(g; \mathbf{Y}^t)$ for some $g \in \mathfrak{H}_0^{\odot q}$,

then we have $\lim_{t \downarrow 0} \frac{1}{t} \mathbb{E}[(F_t - F)(G_t - G) | \sigma\{\mathbf{X}\}] = 2\Gamma(F, G)$, with the convergence in $L^2(\mathbb{P})$.

(c) $\lim_{t \downarrow 0} \frac{1}{t} \mathbb{E}[(F_t - F)^4] = -4p\, \mathbb{E}[F^4] + 12\, \mathbb{E}[F^2 \Gamma(F, F)] \geq 0$.

Proof. By the Mehler formula (1.4), we have

$$\frac{1}{t} \mathbb{E}[F_t - F | \sigma\{\mathbf{X}\}] = \frac{P_t(F) - F}{t} = \frac{e^{-pt} - 1}{t} F,$$

converges in $L^4(\mathbb{P})$ to $-pF = LF$, as $t \downarrow 0$. As a consequence of Lemma 2.1, FG has a finite chaos expansion of the form $FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q} Q_k(h_k; \mathbf{Y})$ for some $h_k \in \mathfrak{H}_0^{\odot k}$. Therefore, $F_t G_t = \mathbb{E}[FG] + \sum_{k=1}^{p+q} Q_k(h_k; \mathbf{Y}^t)$, implying

$$\frac{1}{t} \mathbb{E}[F_t G_t - FG | \sigma\{\mathbf{X}\}] = \sum_{k=1}^{p+q} \frac{1}{t} \mathbb{E}[Q_k(h_k; \mathbf{Y}^t) - Q_k(h_k; \mathbf{Y}) | \sigma\{\mathbf{X}\}]$$

converges in $L^2(\mathbb{P})$ to $\sum_{k=1}^{p+q} -k\, J_k(FG) = L(FG)$, as $t \downarrow 0$. Hence, we infer that in $L^2(\mathbb{P})$ and as $t \downarrow 0$,

$$\begin{aligned}\frac{1}{t} \mathbb{E}[(F_t - F)(G_t - G) | \sigma\{\mathbf{X}\}] &= \frac{1}{t} \mathbb{E}[F_t G_t - FG | \sigma\{\mathbf{X}\}] - F \frac{\mathbb{E}[G_t - G | \sigma\{\mathbf{X}\}]}{t} - G \frac{\mathbb{E}[F_t - F | \sigma\{\mathbf{X}\}]}{t} \\ &\to L(FG) - FLG - GLF = 2\,\Gamma(F, G).\end{aligned}$$

Since the pair (F, F_t) is exchangeable, we can write

$$\begin{aligned}\mathbb{E}[(F_t - F)^4] &= \mathbb{E}[F_t^4 + F^4 - 4F_t^3 F - 4F^3 F_t + 6F_t^2 F^2] \\ &= 2\mathbb{E}[F^4] - 8\mathbb{E}[F^3 F_t] + 6\mathbb{E}[F^2 F_t^2] \qquad \text{(by exchangeability of } (F, F_t)) \\ &= 4\mathbb{E}[F^3(F_t - F)] + 6\mathbb{E}[F^2(F_t - F)^2] \quad \text{(after rearrangement)} \\ &= 4\mathbb{E}[F^3 \mathbb{E}[F_t - F | \sigma\{\mathbf{X}\}]] + 6\mathbb{E}[F^2 \mathbb{E}[(F_t - F)^2 | \sigma\{\mathbf{X}\}]].\end{aligned}$$

so (c) follows immediately from (a),(b) and the fact that $F \in L^4(\mathbb{P})$. □

3 Proofs

We begin with the following lemma, whose proof is postponed to Section 3.3.

Lemma 3.1. Given $F = Q_p(f)$ with $f \in \mathfrak{H}_0^{\odot p}$ and $G = Q_q(g)$ with $g \in \mathfrak{H}_0^{\odot q}$, we assume that $F, G \in L^4(\mathbb{P})$. Then we have the following estimates:

$$\sum_{k=1}^{p+q-1} \mathrm{Var}(J_k(FG)) \le \mathbb{E}[F^2G^2] - 2\mathbb{E}[FG]^2 - \mathrm{Var}(F)\mathrm{Var}(G) + (p+q)! \left\| f\widetilde{\otimes}g\mathbf{1}_{\Delta^c_{p+q}} \right\|^2_{\mathfrak{H}^{\otimes p+q}}, \tag{3.1}$$

and in particular,

$$\max\left\{ \sum_{k=1}^{2p-1} \mathrm{Var}(J_k(F^2)), p!^2 \sum_{r=1}^{p-1} \binom{p}{r}^2 \left\| f\otimes_r f \right\|^2_{\mathfrak{H}^{\otimes 2p-2r}} \right\} \le \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + (2p)! \left\| f\widetilde{\otimes}f\mathbf{1}_{\Delta^c_{2p}} \right\|^2_{\mathfrak{H}^{\otimes 2p}}, \tag{3.2}$$

with

$$\left\| f\widetilde{\otimes}g\mathbf{1}_{\Delta^c_{p+q}} \right\|^2_{\mathfrak{H}^{\otimes p+q}} \le \sum_{r=1}^{p\wedge q} r!\binom{p}{r}\binom{q}{r} \min\left\{ \|f\|^2_{\mathfrak{H}^{\otimes p}}\mathcal{M}(g), \|g\|^2_{\mathfrak{H}^{\otimes q}}\mathcal{M}(f) \right\}. \tag{3.3}$$

(As a convention, we put $\sum_{r=1}^{0} = 0$.)

Before we prove our multivariate limit theorem, we will give a short proof of the univariate case in Wasserstein distance, using our exchangeable pairs coupling.

3.1 Alternative proof of Theorem 1.2

We need the following result, which is the univariate analogue of Proposition 1.1.

Proposition 3.1. Let F and a family of real random variables $(F_t)_{t\ge 0}$ be defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $F_t \overset{law}{=} F$ for every $t \ge 0$. Assume that $F \in L^4(\Omega, \mathcal{G}, \mathbb{P})$ for some σ-algebra $\mathcal{G} \subset \mathcal{F}$ and that in $L^1(\mathbb{P})$,

(a) $\lim\limits_{t\downarrow 0} \frac{1}{t}\mathbb{E}[F_t - F|\mathcal{G}] = -\lambda F$ for some $\lambda > 0$,

(b) $\lim\limits_{t\downarrow 0} \frac{1}{t}\mathbb{E}[(F_t - F)^2|\mathcal{G}] = (2\lambda + S)\mathrm{Var}(F)$ for some random variable S;

(c) and $\lim\limits_{t\downarrow 0} \frac{1}{t}\mathbb{E}[(F_t - F)^4] = \rho(F)\mathrm{Var}(F)^2$ for some $\rho(F) \ge 0$.

Then, with $Z \sim \mathcal{N}(0, \mathrm{Var}(F))$, we have

$$d_{\mathrm{W}}(F, Z) \le \frac{\sqrt{\mathrm{Var}(F)}}{\lambda\sqrt{2\pi}}\mathbb{E}[|S|] + \frac{\sqrt{(2\lambda + \mathbb{E}[S])\mathrm{Var}(F)}}{3\lambda}\sqrt{\rho(F)}.$$

For the proof, one can refer to [7, Proposition 3.3]. One may also want to refer to Theorem 3.5 of [17] for a different coupling bound.

Now given $F = Q_p(f;\mathbf{Y}) \in L^4(\mathbb{P})$ (with $\mathbb{E}[F^2] = 1$), we can get by using (2.4) and (3.2) that

$$\mathrm{Var}\Big(p^{-1}\Gamma(F,F)\Big) \le \sum_{k=1}^{2p-1} \mathrm{Var}\Big(J_k(F^2)\Big) \le \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + (2p)!\Big\|f\widetilde{\otimes}f\mathbf{1}_{\Delta_{2p}^c}\Big\|^2_{\mathfrak{H}^{\otimes 2p}}$$

$$\le \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + \gamma_p\mathbb{E}[F^2]\mathcal{M}(f) \quad \text{with } \gamma_p := \frac{(2p)!}{p!}\sum_{r=1}^{p} r!\binom{p}{r}^2. \tag{3.4}$$

Also using the chaos expansion of F^2 and $\Gamma(F,F)$ as well as the orthogonality property, we have

$$3\mathbb{E}[F^2\Gamma(F,F)] - p\mathbb{E}[F^4] = 3\mathbb{E}\Big[F^2(\Gamma(F,F)-p)\Big] - p\Big(\mathbb{E}[F^4]-3\Big)$$

$$= 3\mathbb{E}\left[\left(\sum_{k=0}^{2p} J_k(F^2)\right)\left(\sum_{k=1}^{2p-1}\frac{2p-k}{2}J_k(F^2)\right)\right] - p\Big(\mathbb{E}[F^4]-3\Big) \le 3p\sum_{k=1}^{2p-1}\mathrm{Var}\Big(J_k(F^2)\Big) - p\Big(\mathbb{E}[F^4]-3\Big).$$

It follows from (3.4) that

$$3\mathbb{E}[F^2\Gamma(F,F)] - p\mathbb{E}[F^4] \le 2p\Big(\mathbb{E}[F^4]-3\Big) + 3p\gamma_p\mathcal{M}(f). \tag{3.5}$$

Now define $F_t = Q_p(f;\mathbf{Y}^t)$ for each $t \in [0,+\infty)$, then by Proposition 2.1, (F_t, F) is an exchangeable pair satisfying the conditions in Proposition 3.1 with $\mathcal{G} = \sigma\{\mathbf{X}\}$, $\lambda = p$, $S = 2\Gamma(F,F) - 2p$ and $\rho(F) = -4p\,\mathbb{E}[F^4] + 12\,\mathbb{E}[F^2\Gamma(F,F)]$. Therefore,

$$\begin{aligned}
d_{\mathrm{W}}(F,N) &\le \frac{1}{p\sqrt{2\pi}}\mathbb{E}[|2\Gamma(F,F)-2p|] + \frac{\sqrt{2p}}{3p}\sqrt{-4p\,\mathbb{E}[F^4] + 12\,\mathbb{E}[F^2\Gamma(F,F)]}\\
&\le \frac{2}{\sqrt{2\pi}}\sqrt{\mathrm{Var}\Big(p^{-1}\Gamma(F,F)\Big)} + \frac{\sqrt{2p}}{3p}\sqrt{-4p\,\mathbb{E}[F^4] + 12\,\mathbb{E}[F^2\Gamma(F,F)]} \quad (\text{as } \mathbb{E}[\Gamma(F,F)] = p)\\
&\le \sqrt{2/\pi}\sqrt{\mathbb{E}[F^4] - 3 + \gamma_p\mathcal{M}(f)} + \frac{2\sqrt{2}}{3}\sqrt{2(\mathbb{E}[F^4]-3) + 3\gamma_p\mathcal{M}(f)}\\
&\le \Big(\sqrt{2/\pi} + \frac{4}{3}\Big)\sqrt{|\mathbb{E}[F^4]-3|} + \Big(\sqrt{2/\pi} + \frac{2\sqrt{6}}{3}\Big)\sqrt{\gamma_p}\sqrt{\mathcal{M}(f)}
\end{aligned}$$

This proves Theorem 1.2 with $C_1 = \sqrt{2/\pi} + \frac{4}{3}$ and $C_2 = \Big(\sqrt{2/\pi} + \frac{2\sqrt{6}}{3}\Big)\sqrt{\frac{(2p)!}{p!}\sum_{r=1}^{p} r!\binom{p}{r}^2}$.

Remark 3.1. (1) For F in the first Rademacher chaos, one can directly prove Theorem 1.2 without using the exchangeable pairs. Indeed, if $F = Q_1(h) \in L^4(\mathbb{P})$ for some $h \in \mathfrak{H}$ with $\|h\|_{\mathfrak{H}} = 1$ and $Z \sim \mathcal{N}(0,1)$, then by [28, Theorem 3.1],

$$d_{\mathrm{W}}(F,Z) \le \sqrt{\sum_{k=1}^{\infty}\frac{1}{p_kq_k}h(k)^4}.$$

By Lemma 2.1, $F^2 = 1 + Q_1(w) + Q_2(h\otimes h\mathbf{1}_{\Delta_2})$ with $w(k) = \frac{h(k)^2(q_k-p_k)}{\sqrt{p_kq_k}}$, $k \in \mathbb{N}$. This implies

$$\mathbb{E}[F^4] = 1 + \sum_{k=1}^{\infty} h(k)^4\frac{(q_k-p_k)^2}{p_kq_k} + 2\|h\otimes h\|^2_{\mathfrak{H}^{\otimes 2}} - 2\|h\otimes h\mathbf{1}_{\Delta_2^c}\|^2_{\mathfrak{H}^{\otimes 2}}$$

$$= 3 + \sum_{k=1}^{\infty} h(k)^4 \frac{(q_k - p_k)^2}{p_k q_k} - 2\sum_{k=1}^{\infty} h(k)^4 = 3 + \sum_{k=1}^{\infty} h(k)^4 \frac{q_k^2 + p_k^2}{p_k q_k} - 4\sum_{k=1}^{\infty} h(k)^4 .$$

Noticing $p_k^2 + q_k^2 \geq 1/2$ for each $k \in \mathbb{N}$, we have

$$\frac{1}{2}\sum_{k=1}^{\infty} \frac{1}{p_k q_k} h(k)^4 \leq 4\sum_{k=1}^{\infty} h(k)^4 + \mathbb{E}[F^4] - 3 \leq 4\mathcal{M}(h) + \mathbb{E}[F^4] - 3 .$$

Hence, $d_{\mathrm{W}}(F, Z) \leq \sqrt{2}\sqrt{\left|\mathbb{E}[F^4] - 3\right|} + 2\sqrt{2}\sqrt{\mathcal{M}(h)}$. Moreover, using the so-called second-order Poincaré inequality in [10, Theorem 4.1], we can have the Berry-Esseen bound

$$d_{\mathrm{Kol}}(F, Z) := \sup_{z\in\mathbb{R}} \left|\mathbb{P}(F \leq z) - \mathbb{P}(Z \leq z)\right| \leq 2\sqrt{\sum_{k=1}^{\infty} \frac{1}{p_k q_k} h(k)^4} \leq 3\sqrt{\left|\mathbb{E}[F^4] - 3\right|} + 6\sqrt{\mathcal{M}(h)} .$$

(2) Continuing the discussion in previous point and assuming $p_k = p = 1 - q = 1 - q_k$ for each k, we have

$$\mathbb{E}[F^4] - 3 = \frac{p^2 + q^4 - 4pq}{pq} \sum_{k=1}^{\infty} h(k)^4 . \tag{3.6}$$

If $p \in (0, 1) \setminus \{\frac{1}{2} \pm \frac{1}{2\sqrt{3}}\}$, then we have the exact fourth moment bounds:

$$d_{\mathrm{W}}(F, Z) \leq \sqrt{\frac{1}{pq}\sum_{k=1}^{\infty} h(k)^4} \leq \left(\frac{\mathbb{E}[F^4] - 3}{p^2 + q^2 - 4pq}\right)^{1/2} \text{ and } d_{\mathrm{Kol}}(F, Z) \leq 2\left(\frac{\mathbb{E}[F^4] - 3}{p^2 + q^2 - 4pq}\right)^{1/2} ,$$

see also Corollary 1.4 in [5].

3.2 Proof of Theorem 1.2

Without losing any generality, we assume that $\Sigma_n = \Sigma$ and each component of $F^{(n)}$ belongs to $L^4(\mathbb{P})$. Recall that $F^{(n)} = (F_1^{(n)}, \ldots, F_d^{(n)})^T := \left(Q_{q_1}(f_{1,n}; \mathbf{Y}), \ldots, Q_{q_d}(f_{d,n}; \mathbf{Y})\right)^T$ and we define $F_t^{(n)} = (F_{1,t}^{(n)}, \ldots, F_{d,t}^{(n)})^T$ with $F_{i,t}^{(n)} := Q_{q_i}(f_{i,n}; \mathbf{Y}^t)$ so that by Lemma 2.2 and Proposition 2.1, $(F, F_t) := (F^{(n)}, F_t^{(n)})$ form an exchangeable pair satisfying the conditions in Proposition 1.1 with $\mathscr{G} = \sigma\{\mathbf{X}\}$, $\Lambda = \mathrm{diag}(q_1, \ldots, q_d)$ and

$$S = \left(2\Gamma(F_i^{(n)}, F_j^{(n)}) - 2q_j\Sigma_{i,j}\right)_{1\leq i,j\leq d}, \quad \rho_i(F^{(n)}) = -4q_i\,\mathbb{E}\Big[(F_i^{(n)})^4\Big] + 12\,\mathbb{E}\Big[(F_i^{(n)})^2\Gamma(F_i^{(n)}, F_i^{(n)})\Big].$$

Indeed, the condition (c) in Proposition 1.1 follows from the relation (c) in Proposition 2.1, and for each $i, j \in \{1, \ldots, d\}$, we have

$$\lim_{t\downarrow 0} \frac{1}{t}\mathbb{E}[F_{i,t}^{(n)} - F_i^{(n)} | \sigma\{\mathbf{X}\}] = -q_i F_i^{(n)} \quad \text{in } L^4(\mathbb{P}),$$

and

$$\lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}\Big[(F^{(n)}_{i,t}-F^{(n)}_{i})(F^{(n)}_{j,t}-F^{(n)}_{j})|\sigma\{\mathbf{X}\}\Big]=2q_j\Sigma_{i,j}+\Big[2\Gamma(F^{(n)}_{i},F^{(n)}_{j})-2q_j\Sigma_{i,j}\Big]\quad\text{in } L^2(\mathbb{P}).$$

It follows that

$$\Big\|\frac{1}{t}\mathbb{E}[F^{(n)}_{t}-F^{(n)}|\sigma\{\mathbf{X}\}]+\Lambda F^{(n)}\Big\|^2_{\mathbb{R}^d}=\sum_{i=1}^{d}\Big(\frac{1}{t}\mathbb{E}[F^{(n)}_{i,t}-F^{(n)}_{i}|\sigma\{\mathbf{X}\}]+q_iF^{(n)}_{i}\Big)^2$$

converges to zero in $L^2(\mathbb{P})$, as $t\downarrow 0$; and

$$\begin{aligned}&\Big\|\frac{1}{t}\mathbb{E}\Big[(F^{(n)}_{t}-F^{(n)})(F^{(n)}_{t}-F^{(n)})^T|\sigma\{\mathbf{X}\}\Big]-2\Lambda\Sigma-S\Big\|^2_{\mathrm{H.S.}}\\&=\sum_{i,j=1}^{d}\Big(\frac{1}{t}\mathbb{E}\Big[(F^{(n)}_{i,t}-F^{(n)}_{i})(F^{(n)}_{j,t}-F^{(n)}_{j})|\sigma\{\mathbf{X}\}\Big]-2\Gamma(F^{(n)}_{i},F^{(n)}_{j})\Big)^2\end{aligned}$$

converges to zero in $L^1(\mathbb{P})$, as $t\downarrow 0$.

Hence we can apply Proposition 1.1 and consequently, it suffices to show

$$\mathbb{E}[\|S\|_{\mathrm{H.S.}}]+\sqrt{\sum_{i=1}^{d}\rho_i(F^{(n)})}\le\Big(\sum_{i,j=1}^{d}\mathrm{Var}\Big(\Gamma(F^{(n)}_{i},F^{(n)}_{j})\Big)\Big)^{1/2}+\sqrt{\sum_{i=1}^{d}\rho_i(F^{(n)})}\to 0\ ,\quad\text{as } n\to+\infty.$$

In view of (3.4) and (3.5), it reduces to prove $\lim_{n\to+\infty}\mathrm{Var}\Big(\Gamma(F^{(n)}_{i},F^{(n)}_{j})\Big)=0$ for $i<j$. We split this part into two steps.

Step 1. Suppose F,G are two real random variables given as in Lemma 2.1 with $p\le q$, then we have

$$\mathbb{E}\Big[F^2G^2\Big]=\mathbb{E}[FG]^2+\sum_{k=1}^{p+q-1}\mathrm{Var}\Big(J_k(FG)\Big)+(p+q)!\Big\|f\widetilde{\otimes}g\mathbf{1}_{\triangle_{p+q}}\Big\|^2_{\mathfrak{H}^{\otimes p+q}}$$

and by (2.4) and Lemma 3.1, we get

$$\begin{aligned}&\frac{1}{q^2}\mathrm{Var}(\Gamma(F,G))\\&\le\sum_{k=1}^{p+q-1}\mathrm{Var}(J_k(FG))\le\mathrm{Cov}(F^2,G^2)-2\mathbb{E}[FG]^2+(2q)!\sum_{r=1}^{p}r!\binom{p}{r}\binom{q}{r}\min\Big\{\|f\|^2_{\mathfrak{H}^{\otimes p}}\mathcal{M}(g),\|g\|^2_{\mathfrak{H}^{\otimes q}}\mathcal{M}(f)\Big\}.\end{aligned}$$

Thus, we can further reduce our problem to show

$$\lim_{n\to+\infty}\Big(\mathrm{Cov}((F^{(n)}_{i})^2,(F^{(n)}_{j})^2)-2\mathbb{E}[F^{(n)}_{i}F^{(n)}_{j}]^2\Big)=0\quad\text{for any } 1\le i<j\le d,\tag{3.7}$$

which will be carried out in the next step.

Step 2. Let F, G be given as in previous step, we have

$$\mathbb{E}[F^2G^2] = \mathbb{E}\left(F^2\left\{\mathbb{E}[G^2] + \sum_{k=1}^{2q-1} J_k(G^2) + J_{2q}(G^2)\right\}\right)$$
$$= \mathrm{Var}(F)\mathrm{Var}(G) + \mathbb{E}\left(F^2\sum_{k=1}^{2q-1} J_k(G^2)\right) + \mathbf{1}_{(p=q)}\mathbb{E}[J_{2q}(F^2)J_{2q}(G^2)]\,.$$

If $p < q$, then $\mathbb{E}[FG] = 0$ and

$$\left|\mathrm{Cov}(F^2, G^2)\right| \leq \sqrt{\mathbb{E}[F^4]}\sqrt{\sum_{k=1}^{2q-1}\mathrm{Var}(J_k(G^2))} \leq \sqrt{\mathbb{E}[F^4]}\sqrt{\mathbb{E}[G^4] - 3\mathbb{E}[G^2]^2 + \gamma_p\mathbb{E}[G^2]\mathcal{M}(g)}\,,$$

where the second inequality follows from (3.4) and the constant γ_p is given therein.

If $p = q$, then

$$\mathbb{E}[J_{2q}(F^2)J_{2q}(G^2)] = (2q)!\left\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\mathbf{1}_{\triangle_{2q}}\right\rangle_{\mathfrak{H}^{\otimes 2q}} = (2q)!\left\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\right\rangle_{\mathfrak{H}^{\otimes 2q}} - (2q)!\left\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\mathbf{1}_{\triangle^c_{2q}}\right\rangle_{\mathfrak{H}^{\otimes 2q}}$$
$$= 2q!^2\langle f, g\rangle^2_{\mathfrak{H}^{\otimes q}} + \sum_{r=1}^{q-1} q!^2\binom{q}{r}^2\langle f\otimes_r g, g\otimes_r f\rangle_{\mathfrak{H}^{\otimes 2q-2r}} - (2q)!\left\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\mathbf{1}_{\triangle^c_{2q}}\right\rangle_{\mathfrak{H}^{\otimes 2q}}\,,$$

where the last equality follows from Lemma 2.2 in [19]. Consequently, $\mathrm{Cov}(F^2, G^2) - 2\mathbb{E}[FG]^2$ is equal to

$$\mathbb{E}\left(F^2\sum_{k=1}^{2q-1} J_k(G^2)\right) + \sum_{r=1}^{q-1} q!^2\binom{q}{r}^2\langle f\otimes_r g, g\otimes_r f\rangle_{\mathfrak{H}^{\otimes 2q-2r}} - (2q)!\left\langle f\widetilde{\otimes}f, g\widetilde{\otimes}g\mathbf{1}_{\triangle^c_{2q}}\right\rangle_{\mathfrak{H}^{\otimes 2q}}\,. \tag{3.8}$$

The first term in (3.8) can be rewritten as $\mathbb{E}\left[\sum_{k=1}^{2q-1} J_k(F^2)J_k(G^2)\right]$, which can be bounded by

$$\sqrt{\sum_{k=1}^{2q-1}\mathrm{Var}(J_k(F^2))}\sqrt{\sum_{k=1}^{2q-1}\mathrm{Var}(J_k(F^2))}$$
$$\leq \sqrt{\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + \gamma_q\mathbb{E}[F^2]\mathcal{M}(f)}\sqrt{\mathbb{E}[G^4] - 3\mathbb{E}[G^2]^2 + \gamma_q\mathbb{E}[G^2]\mathcal{M}(g)}\;;$$

and the second term in (3.8) can be bounded by

$$\sum_{r=1}^{q-1} q!^2\binom{q}{r}^2\left\|f\otimes_r g\right\|^2_{\mathfrak{H}^{\otimes 2q-2r}} = \sum_{r=1}^{q-1} q!^2\binom{q}{r}^2\langle f\otimes_{q-r} f, g\otimes_{q-r} g\rangle_{\mathfrak{H}^{\otimes 2r}} \tag{3.9}$$
$$\leq \sum_{r=1}^{q-1} q!^2\binom{q}{r}^2\left\|f\otimes_{q-r} f\right\|_{\mathfrak{H}^{\otimes 2r}}\cdot\left\|g\otimes_{q-r} g\right\|_{\mathfrak{H}^{\otimes 2r}}$$

$$= \sum_{r=1}^{q-1} q!^2 \binom{q}{r}^2 \left\| f \otimes_r f \right\|_{\mathfrak{H}^{\otimes 2q-2r}} \cdot \left\| g \otimes_r g \right\|_{\mathfrak{H}^{\otimes 2q-2r}}$$

$$\leq \sqrt{\sum_{r=1}^{q-1} q!^2 \binom{q}{r}^2 \left\| f \otimes_r f \right\|^2_{\mathfrak{H}^{\otimes 2q-2r}}} \sqrt{\sum_{r=1}^{q-1} q!^2 \binom{q}{r}^2 \left\| g \otimes_r g \right\|^2_{\mathfrak{H}^{\otimes 2q-2r}}} \tag{3.10}$$

$$\leq \sqrt{\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + \gamma_q \mathcal{M}(f)\mathbb{E}[F^2]} \sqrt{\mathbb{E}[G^4] - 3\mathbb{E}[G^2]^2 + \gamma_q \mathcal{M}(g)\mathbb{E}[G^2]} \,, \tag{3.11}$$

where (3.9) follows from the easy fact that $\|f \otimes_r g\|^2_{\mathfrak{H}^{\otimes 2q-2r}} = \langle f \otimes_{q-r} f, g \otimes_{q-r} g \rangle_{\mathfrak{H}^{\otimes 2r}}$, and we used Cauchy-Schwarz inequality in (3.10), while (3.11) can be deduced from Lemma 3.1 and (3.4); finally, the third term in (3.8) can be bounded by $\|f\|^2_{\mathfrak{H}^{\otimes q}} (2q)! \left\| g \widetilde{\otimes} g \mathbf{1}_{\Delta^c_{2q}} \right\|_{\mathfrak{H}^{\otimes 2q}} \leq \|f\|^2_{\mathfrak{H}^{\otimes q}} \sqrt{(2q)! \gamma_q \mathbb{E}[G^2] \mathcal{M}(g)}$. To conclude this case, we obtain

$$\begin{aligned} &\left| \mathrm{Cov}(F^2, G^2) - 2\mathbb{E}[FG]^2 \right| \\ &\leq 2 \sqrt{\left(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 + \gamma_q \mathcal{M}(f)\mathbb{E}[F^2] \right)\left(\mathbb{E}[G^4] - 3\mathbb{E}[G^2]^2 + \gamma_q \mathcal{M}(g)\mathbb{E}[G^2] \right)} \\ &\qquad + \|f\|^2_{\mathfrak{H}^{\otimes q}} \sqrt{(2q)! \gamma_q \mathbb{E}[G^2] \mathcal{M}(g)} \,. \end{aligned}$$

Combining the above two cases, we get immediately the relation (3.7), and hence we finish the proof of Theorem 1.1.

3.3 Proofs of technical lemmas

Proof of Lemma 2.1 Let us first introduce some notation: if $F = \mathfrak{f}(\mathbf{X})$, we write

$$F^{\oplus k} = \mathfrak{f}(X_1, \dots, X_{k-1}, +1, X_{k+1}, \dots) \text{ and } F^{\ominus k} = \mathfrak{f}(X_1, \dots, X_{k-1}, -1, X_{k+1}, \dots),$$

we define the *discrete gradient* $D_k F = \sqrt{p_k q_k}(F^{\oplus k} - F^{\ominus k})$, in particular, $D_k Y_k = 1$. We can define the iterated gradients $D^{(m)}_{k_1,\dots,k_m} = D_{k_1} \circ D^{(m-1)}_{k_2,\dots,k_m}$ with $D^{(1)}_k = D_k$. For example, $D_k Q_d(f) = d Q_{d-1}(f(k,\cdot))$ and $D^{(2)}_{k,\ell} Q_d(f) = d(d-1) Q_{d-1}(f(k,\ell,\cdot))$ for $d \geq 2$ and $f \in \mathfrak{H}_0^{\odot d}$, see [10] for more details.

Proof. It is clear that $FG \in L^2(\mathbb{P})$ has the chaotic expansion

$$FG = \mathbb{E}[FG] + \sum_{m \geq 1} Q_m(h_m) \,,$$

where for each $m \in \mathbb{N}$, the kernel $h_m \in \mathfrak{H}_0^{\odot m}$ is given by $h_m(k_1, \dots, k_m) := \frac{1}{m!} \mathbb{E}\left[D^{(m)}_{k_1,\dots,k_m}(FG) \right]$, due to the Stroock's formula (Proposition 2.1 in [10]). So it suffices to show that

$$D^{(p+q)}_{k_1,\dots,k_{p+q}}(FG) = (p+q)!(f \widetilde{\otimes} g)(k_1, \dots, k_{p+q}) \mathbf{1}_{\Delta_{p+q}}(k_1, \dots, k_{p+q}) \quad \text{and} \quad D^{(s)}_{k_1,\dots,k_s}(FG) = 0 \tag{3.12}$$

for any $s > p + q$. Note that the second part follows immediately from the first one.

Recall the product formula (see *e.g.* [10, (2.4)]) for the discrete gradient D_k: for $F, G \in L^2(\mathbb{P})$,

$$D_k(FG) = (D_k F)G + F(D_k G) - \frac{X_k}{\sqrt{p_k q_k}} (D_k F)(D_k G) =: D^L_k(FG) + D^R_k(FG) + D^M_k(FG) \,,$$

that is, we decompose D_k into three operations D_k^L, D_k^R and D_k^M. Therefore, we can write for $k_1 < \ldots < k_{p+q}$,

$$D^{(p+q)}_{k_1,\ldots,k_{p+q}}(FG) = \sum_{A_1,\ldots,A_{p+q}\in\{L,M,R\}} D^{A_1}_{k_1}\circ\cdots\circ D^{A_{p+q}}_{k_{p+q}}(FG) = \sum_{A_1,\ldots,A_{p+q}\in\{L,R\}} D^{A_1}_{k_1}\circ\cdots\circ D^{A_{p+q}}_{k_{p+q}}(FG)\,,$$

where the last equality follows from the fact that for $k\neq\ell$, $D_\ell(X_kF) = X_kD_\ell F$. Moreover, $D^{A_1}_{k_1}\circ\cdots\circ D^{A_{p+q}}_{k_{p+q}}(FG) = 0$ unless L appears exactly p times and R appears exactly q times in the words $A_1,\ldots,A_{p+q}$, so that one can further rewrite $D^{(p+q)}_{k_1,\ldots,k_{p+q}}(FG)$ as

$$\sum_{\substack{\sigma\in\mathfrak{S}_{p+q}:\\ \sigma(1)<\ldots<\sigma(p)\\ \sigma(p+1)<\ldots<\sigma(p+q)}} \Big(D^{(p)}_{k_{\sigma(1)},\ldots,k_{\sigma(p)}}F\Big)\Big(D^{(q)}_{k_{\sigma(p+1)},\ldots,k_{\sigma(p+q)}}G\Big) = \sum_{\sigma\in\mathfrak{S}_{p+q}} f(k_{\sigma(1)},\ldots,k_{\sigma(p)})g(k_{\sigma(p+1)},\ldots,k_{\sigma(p+q)})\,,$$

where the last equality follows from the symmetry of f and g, and it gives us $D^{(p+q)}_{k_1,\ldots,k_{p+q}}(FG) = (p+q)!(f\widetilde{\otimes}g)(k_1,\ldots,k_{p+q})$. This proves (3.12), while the particular case follows from again the Stroock's formula. More precisely, one can first deduce from the previous discussion that $Q_1(h)^2 = \|h\|^2_{\mathfrak{H}} + Q_1(w) + Q_2(h\otimes h\mathbf{1}_{\Delta_2})$ for some $w\in\mathfrak{H}$ given by $w(k) := \mathbb{E}[D_k(Q_1(h)^2)]$. By the definition of discrete gradient, one has

$$\begin{aligned} D_k(Q_1(h)^2) &= \sqrt{p_kq_k}\left\{\left(\sum_{j\neq k} h(j)Y_j + h(k)\frac{1-p_k+q_k}{2\sqrt{p_kq_k}}\right)^2 - \left(\sum_{j\neq k} h(j)Y_j + h(k)\frac{-1-p_k+q_k}{2\sqrt{p_kq_k}}\right)^2\right\}\\ &= h(k)^2\frac{q_k-p_k}{\sqrt{p_kq_k}} + 2h(k)\sum_{j\neq k} h(j)Y_j\ , \end{aligned}$$

which concludes our proof of Lemma 2.1. □

Proof of Lemma 3.1: It follows from Lemma 2.1 that

$$FG = \mathbb{E}[FG] + \sum_{k=1}^{p+q-1} J_k(FG) + Q_{p+q}\Big(f\widetilde{\otimes}g\mathbf{1}_{\Delta_{p+q}}\Big)\,,$$

therefore, by orthogonality property, one has

$$\begin{aligned} \mathbb{E}[F^2G^2] &= \mathbb{E}[FG]^2 + \sum_{k=1}^{p+q-1}\mathrm{Var}(J_k(FG)) + (p+q)!\,\Big\|f\widetilde{\otimes}g\mathbf{1}_{\Delta_{p+q}}\Big\|^2_{\mathfrak{H}^{\otimes p+q}}\\ &= \mathbb{E}[FG]^2 + \sum_{k=1}^{p+q-1}\mathrm{Var}(J_k(FG)) + (p+q)!\,\Big\|f\widetilde{\otimes}g\Big\|^2_{\mathfrak{H}^{\otimes p+q}} - (p+q)!\Big\|f\widetilde{\otimes}g\mathbf{1}_{\Delta^c_{p+q}}\Big\|^2_{\mathfrak{H}^{\otimes p+q}}\,. \end{aligned}$$

Recall from [19, Lemma 2.2] that

(3.13)

$$(p+q)!\,\Big\|f\widetilde{\otimes}g\Big\|^2_{\mathfrak{H}^{\otimes p+q}} = p!q!\,\sum_{r=0}^{p\wedge q}\binom{p}{r}\binom{q}{r}\Big\|f\otimes_r g\Big\|^2_{\mathfrak{H}^{\otimes p+q-2r}} \geq p!q!\|f\|^2_{\mathfrak{H}^{\otimes p}}\|g\|^2_{\mathfrak{H}^{\otimes q}} + \mathbf{1}_{(p=q)}p!^2\langle f,g\rangle^2_{\mathfrak{H}^{\otimes p}}\,,$$

thus (3.1) follows by noticing that $\mathbb{E}[FG] = \mathbf{1}_{(p=q)}p!\langle f, g\rangle_{\mathfrak{H}^{\otimes p}}$ and $\mathrm{Var}(F)\mathrm{Var}(G) = p!q!\|f\|^2_{\mathfrak{H}^{\otimes p}}\|g\|^2_{\mathfrak{H}^{\otimes q}}$.

Using (3.13) again, we have

(3.14)

$$\sum_{k=1}^{p+q-1}\mathrm{Var}(J_k(F^2)) = \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 - p!^2\sum_{r=1}^{p-1}\binom{p}{r}^2\left\|f\otimes_r f\right\|^2_{\mathfrak{H}^{\otimes 2p-2r}} + (2p)!\left\|f\widetilde{\otimes}f\mathbf{1}_{\triangle^c_{2p}}\right\|^2_{\mathfrak{H}^{\otimes 2p}} ,$$

which implies (3.2).

It remains to prove (3.3) and we'll use the same arguments as in the proof of [5, Lemma 3.3]:

$$\begin{aligned}\left\|f\widetilde{\otimes}g\mathbf{1}_{\triangle^c_{p+q}}\right\|^2_{\mathfrak{H}^{\otimes p+q}} \leq \left\|f\otimes g\mathbf{1}_{\triangle^c_{p+q}}\right\|^2_{\mathfrak{H}^{\otimes p+q}} &= \sum_{(i_1,\ldots,i_p,j_1,\ldots,j_q)\in\triangle^c_{p+q}} f(i_1,\ldots,i_p)^2 g(j_1,\ldots,j_q)^2\\ &= \sum_{r=1}^{p\wedge q} r!\binom{p}{r}\binom{q}{r}\sum_{\substack{(i_1,\ldots,i_p)\in\triangle_p\\(j_1,\ldots,j_q)\in\triangle_q\\ \mathrm{card}(\{i_1,\ldots,i_p\}\cap\{j_1,\ldots,j_q\})=r}} f(i_1,\ldots,i_p)^2 g(j_1,\ldots,j_q)^2 ,\end{aligned}\tag{3.15}$$

where $\mathrm{card}(A)$ means the cardinality of the set A, and the combinatorial constant $r!\binom{p}{r}\binom{q}{r}$ is the number of ways one can build r pairs of identical indices out of $(i_1,\ldots,i_p)\in\triangle_p$ and $(j_1,\ldots,j_q)\in\triangle_q$.

Therefore, it is enough to notice that for each $r\in\{1,\ldots,p\wedge q\}$, the inner sum in (3.15) is bounded by

$$\begin{aligned}&\sum_{\substack{(i_1,\ldots,i_{p-r},k_1,\ldots,k_r)\in\triangle_p\\(j_1,\ldots,j_{q-r},k_1,\ldots,k_r)\in\triangle_q}} f(i_1,\ldots,i_{p-r},k_1,\ldots,k_r)^2 g(j_1,\ldots,j_{q-r},k_1,\ldots,k_r)^2\\ \leq &\sum_{\substack{(i_1,\ldots,i_{p-1},k)\in\triangle_p\\(j_1,\ldots,j_{q-1},k)\in\triangle_q}} f(i_1,\ldots,i_{p-1},k)^2 g(j_1,\ldots,j_{q-1},k)^2 \leq \min\left\{\|f\|^2_{\mathfrak{H}^{\otimes p}}\mathcal{M}(g), \|g\|^2_{\mathfrak{H}^{\otimes q}}\mathcal{M}(f)\right\} .\end{aligned}$$

The proof of Lemma 3.1 is complete.

4 Universality of Homogeneous sums

Fix $d\geq 2$ and a *divergent* sequence $(N_n, n\geq 1)$ of natural numbers. Consider the kernels $f_n : \{1,\ldots,N_n\}^d \to \mathbb{R}$ symmetric and vanishing on diagonals and $d!\|f_n\|^2_{\mathfrak{H}^{\otimes d}} = 1$, then according to (1.2),

$$Q_d(f_n;\Xi) = \sum_{i_1,\ldots,i_d\leq N_n} f_n(i_1,\ldots,i_d)\xi_{i_1}\cdots\xi_{i_d} .$$

The following central limit theorem due to de Jong [4] gave sufficient conditions for asymptotic normality of $Q_d(f_n;\Xi)$.

Theorem 4.1. Under the above setting, let $\Xi = (\xi_i, i\geq 1)$ be a sequence of independent centred random variables with unit variance and *finite fourth moments*. If $\mathbb{E}[Q_d(f_n;\Xi)^4]\to 3$ and the maximal influence $\mathcal{M}(f_n)\to 0$ as $n\to+\infty$, then $Q_d(f_n;\Xi)$ converges in law to a standard Gaussian.

The above result exhibits the universality phenomenon as well as the importance of the notion "maximal influence". Another striking result with similar nature is the invariance principle established in [14], in which the authors were able to control distributional distance between homogeneous sums over different sequences of independent random variables in terms of maximal influence, see *e.g.* Theorem 2.1 therein.

Let us restrict ourselves to the Gaussian setting for a while: when $\mathbf{G}$ is a sequence of *i.i.d.* standard Gaussians, $Q_d(f_n;\mathbf{G})$ belongs to the d-th Gaussian Wiener chaos, and the fourth moment theorem [21] implies that if $Q_d(f_n;\mathbf{G})$ converges in law to a standard Gaussian (or equivalently $\mathbb{E}[Q_d(f_n;\mathbf{G})^4]\to 3$), then $\|f_n\otimes_{d-1}f_n\|_{\mathfrak{H}^{\otimes 2}}\to 0$. While $\mathcal{M}(f_n)\leq\|f_n\otimes_{d-1}f_n\|_{\mathfrak{H}^{\otimes 2}}$ due to [17, Lemma 2.4], so that $\mathcal{M}(f_n)\to 0$. This hints the universality of the Gaussian Wiener chaos, see [18] for more details.

The following result is (slightly) adapted from Theorem 7.5 in [18].

Theorem 4.2. Fix integers $d\geq 2$ and $q_d\geq\ldots\geq q_1\geq 2$. For each $j\in\{1,\ldots,d\}$, let $(N_{j,n},n\geq 1)$ be a sequence of natural numbers diverging to infinity, and let $f_{j,n}:\{1,\ldots,N_{j,n}\}^{q_j}\to\mathbb{R}$ be symmetric and vanishing on diagonals (*i.e.* $f_{j,n}\in\mathfrak{H}_0^{\odot q_j}$ with support contained in $\{1,\ldots,N_{j,n}\}^{q_j}$) such that

$$\lim_{n\to+\infty}\mathbf{1}_{(q_k=q_l)}q_k!\sum_{i_1,\ldots,i_{q_k}\leq N_{k,n}}f_{k,n}(i_1,\ldots,i_{q_k})f_{l,n}(i_1,\ldots,i_{q_k})=\Sigma_{k,l}\,,$$

where $\Sigma=(\Sigma_{i,j},1\leq i,j\leq d)$ is a symmetric nonnegative definite d by d matrix. Then the following statements are equivalent:

(A_1) Given a sequence $\mathbf{G}$ of i.i.d. standard Gaussians, $\left(Q_{q_1}(f_{1,n};\mathbf{G}),\ldots,Q_{q_d}(f_{d,n};\mathbf{G})\right)^T$ converges in distribution to $\mathcal{N}(0,\Sigma)$, as $n\to+\infty$.

(A_2) For every sequence $\Xi=(\xi_i,i\in\mathbb{N})$ of independent centred random variables with unit variance and $\sup_{i\in\mathbb{N}}\mathbb{E}[|\xi_i|^3]<+\infty$, the sequence of d-dimensional random vectors $\left(Q_{q_1}(f_{1,n};\Xi),\ldots,Q_{q_d}(f_{d,n};\Xi)\right)^T$ converges in distribution to $\mathcal{N}(0,\Sigma)$, as $n\to+\infty$.

Similar universality result for Poisson chaos was first established in [24] and refined recently in [7]. It was pointed out in [24] and [18] that homogeneous sums inside the Rademacher chaos are not universal with respect to normal approximation and a counterexample is available *e.g.* in [24, Proposition 1.7]:

A Counterexample: Let $\mathbf{Y}$ be a sequence of *i.i.d.* random variables with $\mathbb{P}(Y_1=1)=\mathbb{P}(Y_1=-1)=1/2$ (that is, in the symmetric setting). Fix $q\geq 2$ and for each $N\geq q$, we set

$$f_N(i_1,\ldots,i_q)=\begin{cases}\dfrac{1}{q!\,\sqrt{N-q+1}}\,, & \text{if } \{i_1,\ldots,i_q\}=\{1,2,\ldots,q-1,s\} \text{ for } q\leq s\leq N;\\ 0\,, & \text{otherwise.}\end{cases}$$

Then in the symmetric case,

$$Q_q(f_N;\mathbf{Y})=Y_1Y_2\cdots Y_{q-1}\sum_{i=q}^{N}\frac{Y_i}{\sqrt{N-q+1}}$$

converges in law to the standard Gaussian, while if $\mathbf{G}$ is a sequence of i.i.d. standard Gaussians, then for every $N \geq 2$, $Q_q(f_N; \mathbf{G}) \overset{law}{=} G_1 G_2 \cdots G_q$ fails to be Gaussian. It is easy to check that the maximal influence $\mathcal{M}(f_N)$ of the kernel f_N is equal to $1/(qq!)$ for every $N \geq 2$, which is consistent with de Jong's theorem.

In the end of this section, we provide a (partially) universal result for Rademacher chaos that complements [7, 18, 24].

Proposition 4.1. Let the assumptions in Theorem 4.2 prevail. Then, the following statement is equivalent to (A_1) and (A_2) in Theorem 4.2:

(A_3) in the symmetric case, as $n \to +\infty$, $\left(Q_{q_1}(f_{1,n}; \mathbf{Y}), \ldots, Q_{q_d}(f_{d,n}; \mathbf{Y})\right)^T$ converges in distribution to $\mathcal{N}(0, \Sigma)$, and $\mathcal{M}(f_{j,n}) \to 0$ for each $j \in \{1, \ldots, d\}$.

Proof. Suppose (A_1) holds true, then $\left(Q_{q_1}(f_{1,n}; \mathbf{Y}), \ldots, Q_{q_d}(f_{d,n}; \mathbf{Y})\right)^T$ converges in distribution to $\mathcal{N}(0, \Sigma)$ by "$(A_2) \Leftrightarrow (A_1)$"; and by the fourth moment theorem on a Gaussian space [21], (A_1) implies that $\|f_{j,n} \otimes_{q_j - 1} f_{j,n}\|_{\mathfrak{H}^{\otimes 2}} \to 0$, as $n \to +\infty$. Recall from [17, Lemma 2.4] that $\mathcal{M}(f) \leq \|f \otimes_{d-1} f\|_{\mathfrak{H}^{\otimes 2}}$ for each $f \in \mathfrak{H}_0^{\odot d}$, therefore $\mathcal{M}(f_{j,n}) \to 0$ for each $j \in \{1, \ldots, d\}$. This proves the implication "$(A_1) \Rightarrow (A_3)$".

It remains to show "$(A_3) \Rightarrow (A_1)$". Now we assume that (A_3) is true, then by a weak form of the *hypercontractivity* property (see Section 2), we have $\lim_{n\to+\infty} \mathbb{E}[Q_{q_j}(f_{n,j}; \mathbf{Y})^4] = 3\Sigma_{j,j}^2$ for each $j = 1, \ldots, d$. It follows from Lemma 3.1 that $\left\|f_{j,n} \otimes_r f_{j,n}\right\|_{\mathfrak{H}^{\otimes 2q_j - 2r}} \to 0$ for each $r = 1, \ldots, q_j - 1$, and any $j = 1, \ldots, d$. Hence, (A_1) follows immediately from the Peccati-Tudor theorem [23]. This concludes our proof. □

References

[1] E. Azmoodeh, S. Campese and G. Poly. Fourth moment theorems for Markov diffusion generators. *J. Funct. Anal.* 266 (4), 2341-2359

[2] S. Campese, I. Nourdin, G. Peccati and G. Poly. Multivariate Gaussian approximations on Markov chaoses. *Electron. Commun. Probab.* Volume 21 (2016), no. 48, 1-9.

[3] S. Chatterjee and E. Meckes. Multivariate normal approximation using exchangeable pairs. *ALEA* 4, 257-283, 2008.

[4] P. de Jong. A central limit theorem for generalized multilinear forms. *J. Multivariate Anal.* **34** 275-289.

[5] C. Döbler and K. Krokowski. On the fourth moment condition for Rademacher chaos. To appear in *Ann. Inst. Henri Poincaré Probab. Stat.* (2017+)

[6] C. Döbler and G. Peccati. The fourth moment theorem on the Poisson space. *Ann. Probab.* to appear (2017+)

[7] C. Döbler, A. Vidotto and G. Zheng. Fourth moment theorems on the Poisson space in any dimension. arXiv preprint, (2017)

[8] K. Krokowski. Poisson approximation of Rademacher functionals by the Chen-Stein method and Malliavin calculus. *Commun. Stoch. Anal.* 11 (2017), no. 2, 195-222

[9] K. Krokowski, A. Reichenbachs and Ch. Thäle. Berry-Esseen bounds and multivariate limit theorems for functionals of Rademacher sequences. *Ann. Inst. Henri Poincaré Probab. Stat.* 52 (2016), no. 2, 763-803

[10] K. Krokowski, A. Reichenbachs and Ch. Thäle. Discrete Malliavin-Stein method: Berry-Esseen bounds for random graphs and percolation. *Ann. Probab.* 45 (2017), no. 2, 1071-1109

[11] M. Ledoux. Chaos of a Markov operator and the fourth moment condition. *Ann. Probab.* Volume 40, Number 6 (2012), 2439-2459.

[12] E. Meckes. An Infinitesimal Version of Stein's Method of Exchangeable Pairs. *Ph.D dissertation*, Stanford University (2006)

[13] E. Meckes. On Stein's method for multivariate normal approximation. *IMS collections, High dimensional Probability* V: The Luminy Volume, Vol. 5 (2009) 153-178.

[14] E. Mossel, R. O'Donnell and K. Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. of Math.* (2) (2010) 171 295-341

[15] I. Nourdin and G. Peccati. Stein's method on Wiener chaos, *Probab. Theory Relat. Fields* (2009), Vol. 145, Issue 1, p. 75-118.

[16] I. Nourdin and G. Peccati. Normal approximations with Malliavin calculus: from Stein's method to universality, *Cambridge tracts in Mathematics*, Vol. 192, 2012, Cambridge University Press.

[17] I. Nourdin, G. Peccati and G. Reinert. Stein's Method and Stochastic Analysis of Rademacher Functionals. *Electron. J. Probab.* Volume 15 (2010), no. 55, 1703-1742.

[18] I. Nourdin, G. Peccati and G. Reinert. Invariance principles for homogeneous sums: Universality of Gaussian Wiener chaos. *Ann. Probab.* Volume 38, Number 5 (2010), 1947-1985.

[19] I. Nourdin and J. Rosiński. Asymptotic independence of multiple Wiener-Itô integrals and the resulting limit laws. *Ann. Probab.* 42, no. 2 (2014), 497-526

[20] I. Nourdin and G. Zheng. Exchangeable pairs on Wiener chaos. arXiv preprint, (2017)

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

[22] R. O'Donnell. Analysis of Boolean functions. *Cambridge University Press*, 2014

[23] G. Peccati and C.A. Tudor. Gaussian limits for vector-valued multiple stochastic integrals, *Séminaire de Probabilités* XXXVIII, 2005

[24] G. Peccati and C. Zheng. Universal Gaussian fluctuations on the discrete Poisson chaos. *Bernoulli*, 20(2):697-715, 2014

[25] N. Privault. Stochastic analysis of Bernoulli processes, *Probab. Surv.* 5 (2008) 435-483.

[26] N. Privault and G. L. Torrisi. The Stein and Chen-Stein methods for functionals of non-symmetric Bernoulli processes. *ALEA* 12 (2015) 309 -356.

[27] Ch. Stein. Approximate computation of expectations. In *Institute of Mathematical Statistics Lecture Notes - Monograph Series*, volume 7. Institute of Mathematical Statistics, 1986.

[28] G. Zheng. Normal approximation and almost sure central limit theorem for non-symmetric Rademacher functionals. *Stochastic Process. Appl.*, Volume 127, Issue 5, 2017, page1622–1636.

This is the last page.