Vaasan yliopisto
UNIVERSITY OF VAASA

BENJAMIN ARRAS
EHSAN AZMOODEH
GUILLAUME POLY
YVIK SWAN

Stein's method on the second Wiener chaos

2-Wasserstein distance

PROCEEDINGS OF THE UNIVERSITY OF VAASA
WORKING PAPERS 9
MATHEMATICS 5

Publisher Vaasan yliopisto	**Date of publication** January 2016	
Authors Benjamin Arras, Ehsan Azmoodeh, Guillaume Poly and Yvik Swan	**Type of publication** Working papers	
	Name and number of series Proceedings of the University of Vaasa. Working Papers, 9 Proceedings of the University of Vaasa. Working papers, Mathematics, 5	
Contact information University of Vaasa Faculty of Technology Department of Mathematics and Statistics P.O. Box 700 FI-65101 Vaasa Finland	**ISBN** 978-952-476-666-1 (online)	
	ISSN 1799-7658 (Proceedings of the University of Vaasa. Working Papers 9, online)	
	Number of pages 34	**Language** English
Title of publication Stein's method on the second Wiener chaos :2-Wasserstein distance		

Abstract

In the first part of the paper we use a new Fourier technique to obtain a Stein characterization for random variables in the second Wiener chaos. We provide the connection between this result and similar conclusions that can be derived using Malliavin calculus. We also introduce a new form of discrepancy which we use, in the second part of the paper, to provide bounds on the 2-Wasserstein distance between linear combinations of independent centered random variables. Our method of proof is entirely original. In particular it does not rely on estimation of bounds on solutions of the so-called Stein equations at the heart of Stein's method. We provide several applications, and discuss comparison with recent similar results on the same topic.

Keywords

Stein's method, Stein discrepancy, Second Wiener chaos, Variance Gamma distribution, 2-Wasserstein distance, Malliavin Calculus

MSC 2010

60F05, 60G50, 60G15, 60H07

Contents

1 INTRODUCTION

1.1 Background

Stein's method is a popular and versatile probabilistic toolkit for stochastic approximation. Presented originally in the context of Gaussian CLTs with dependant summands (see [28]) it has now been extended to cater for a wide variety of quantitative asymptotic results, see [8] for a thorough overview of Gaussian approximation or https://sites.google.com/site/steinsmethod for an up-to-date list of references on non-Gaussian and non-Poisson Stein-type results. To this date one of the most active areas of application of the method is in Gaussian analysis, via Nourdin and Peccati's so-called Malliavin/Stein calculus on Wiener space, see [20] or Ivan Nourdin's dedicated webpage https://sites.google.com/site/malliavinstein.

Given two random objects F, F_∞, Stein's method allows to compute fine bounds on quantities of the form

$$\sup_{h\in\mathcal{H}} \left|\mathbb{E}\left[h(F)\right] - \mathbb{E}\left[h(F_\infty)\right]\right|.$$

The method rests on three pins :

A. a "Stein pair", i.e. a linear operator and a class of functions $(\mathcal{A}_\infty, \mathcal{F}(\mathcal{A}_\infty))$ such that $\mathbb{E}\left[\mathcal{A}_\infty(f(F_\infty))\right] = 0$ for all $f \in \mathcal{F}(\mathcal{A}_\infty)$;

B. a contractive inverse operator $\mathcal{A}_\infty^{-1}$ acting on the centered functions $\bar{h} = h - \mathbb{E}h(F_\infty)$ in $\mathcal{H}$ and contraction information, i.e. tight bounds on $\mathcal{A}_\infty^{-1}(\bar{h})$ and its derivatives;

C. handles on the structure of F (such as $F = F_n = T(X_1, \ldots, X_n)$ a U-statistic, $F = F(X)$ a functional of an isonormal Gaussian process, F a statistic on a random graph, etc.).

Given the conjunction of these three elements one can then apply some form of transfer principle :

$$\sup_{h\in\mathcal{H}} \left|\mathbb{E}\left[h(F)\right] - \mathbb{E}\left[h(F_\infty)\right]\right| = \sup_{h\in\mathcal{H}} \left|\mathbb{E}\left[\mathcal{A}_\infty\left(\mathcal{A}_\infty^{-1}(\bar{h}(F))\right)\right]\right|; \qquad (1.1)$$

remarkably the right-hand-side of the above is often much more amenable to computations than the left-hand-side, even in particularly unfavourable circumstances. This has resulted in Stein's method delivering several striking successes (see [6, 8, 20]) which have led the method to becoming the recognised and acclaimed tool it is today.

In general the identification of a Stein operator is the cornerstone of the method. While historically most practical implementations relied on adhoc arguments, several general tools exist, including Stein's *density approach* [29] and Barbour's *generator approach* [4]. A general theory for Stein operators is available in [17]. In many important cases, these are first order differential operators (see [9]) or difference operators (see [18]). Higher order differential operators have recently come into light (see [15, 25]).

Once an operator is $\mathcal{A}_\infty$ identified, the task is then to bound the resulting rhs of (1.1); there are many ways to perform this. In this paper we focus on Nourdin and Peccati's approach to the method. Let F_∞ be a standard Gaussian random variable. Then the appropriate operator is $\mathcal{A}_\infty f(x) = f'(x) - xf(x)$. Given a sufficiently regular centered random variable F with finite variance and smooth density, define its Stein kernel $\tau_F(F)$ through the integration by parts formula

$$\mathbb{E}[\tau_F(F)\phi'(F)] = \mathbb{E}\left[F\phi(F)\right] \text{ for all absolutely continuous } \phi. \qquad (1.2)$$

Then, for f a solution to $f'(x) - xf(x) = h(x) - \mathbb{E}[h(F_\infty)]$ write

$$\mathbb{E}[h(F)] - \mathbb{E}[h(F_\infty)] = \mathbb{E}\left[f_h'(F) - Ff_h(F)\right] = \mathbb{E}\left[(1-\tau_F(F))f_h'(F)\right]$$

so that

$$|\mathbb{E}[h(F)] - \mathbb{E}[h(F_\infty)]| \leq \|f_h'\|\sqrt{\mathbb{E}\left[(1-\tau_F(F))^2\right]}.$$

At this stage two good things happen : (i) the constant $\sup_{h\in\mathcal{H}}\|f_h'\|$ (which is intrinsically Gaussian and does not depend on the law of F) is bounded for wide and relevant classes $\mathcal{H}$; (ii) the quantity

$$S(F\,||\,F_\infty) = \mathbb{E}\left[(1-\tau_F(F))^2\right] \tag{1.3}$$

(called the *Stein discrepancy*) is tractable, via Malliavin calculus, as soon as F is a sufficiently regular functional of a Gaussian process. These two realizations opened a wide field of applications within the so-called "Malliavin Stein Fourth moment theorems", see [22, 20]. A similar approach holds also if F_∞ is centered Gamma, see [21, 2], and more generally if the law of the target random variable F_∞ belongs to the family of Variance Gamma distributions, see [12] for the method and [15] for the bounds on the corresponding solutions. See also [30, 11] for other generalizations. We stress that in the Gaussian case, Stein's method provides bounds e.g. in the Total Variation distance, whereas technicalities related to the Gamma and Variance Gamma targets impose that one must deal with smoother distances (i.e. integrated probability measures of the form (1.1) with $\mathcal{H}$ a class of smooth functions) in such cases.

1.2 Purpose of this paper

The primary purpose of this paper is to extend Nourdin and Peccati's "Stein discrepancy analysis" to provide meaningful bounds on $\mathrm{d}(F, F_\infty)$ for $\mathrm{d}(.,.)$ some appropriate probability metric and random variables F_∞ belonging to the second Wiener chaos, that is

$$F_\infty = \sum_{i=1}^{q} \alpha_{\infty,i}(N_i^2 - 1). \tag{1.4}$$

where $q \geq 2$, $\{N_i\}_{i=1}^q$ are i.i.d. $\mathscr{N}(0,1)$ random variables, and the coefficients $\{\alpha_{\infty,i}\}_{i=1}^q$ are distinct.

Such a generalization immediately runs into a series of obstacles which need to be dealt with. We single out three crucial questions : (Q1) what operator $\mathcal{A}_\infty$? (Q2) what quantity will play the role of the Stein discrepancy $S(F\,||\,F_\infty)$? (Q3) what kind of distances $\mathrm{d}(.,.)$ can we tackle through this approach?

In this paper we provide a complete answers to a more general version of (Q1), hereby opening the way for applications of Stein's method to a wide variety of new target distributions. We use results from [3] to answer (Q2) for chaotic random variables. We also provide an answer to (Q3) for $\mathrm{d}(.,.)$ the p-Wasserstein distances with $p \leq 2$, under specific assumptions on the structure of F. Such a result extends the scope of Stein's method to so far unchartered territories, because aside for the case $p = 1$, p-Wasserstein distances do not admit a representation of the form (1.1).

1.3 Overview of the results

In the first part of the paper, Section 2, we discuss Stein's method for target distributions of the form (1.4). In Section 2.1 we introduce an entirely new Fourier-based approach to prove a Stein-type characterization for a large family of F_∞ encompassing those of the form (1.4). The operator $\mathcal{A}_\infty$ we obtain is a differential operator of order q. In Section 2.3 we use recent results from [3] to derive a Malliavin-based justification for our $\mathcal{A}_\infty$ when F_∞ is of the form (1.4). We also introduce a new quantity $\Delta(F_n, F_\infty)$ for which we will provide a heuristic justification in Section 2.4 of the fact that $\Delta(F_n, F_\infty)$ generalizes the Stein discrepancy $S(F \,||\, F_\infty)$ in a natural way for chaotic random variables. Finally we argue that quantitative assessments for general targets of the form (1.4) are out of the scope of the current version of Stein's method.

In the second part of the paper, Section 3, we introduce an entirely new *polynomial* approach to Stein's method to provide bounds on the Wasserstein-2 distance (and hence the Wasserstein-1 distance) in terms of $\Delta(F_n, F_\infty)$. Our approach bypasses entirely the need for estimating bounds on solutions of Stein equations. More specifically we provide a tool for providing quantitative assessments on $\mathrm{d}_{W_2}(F_n, F_\infty)$ in terms of the generalized Stein discrepancy $\Delta(F_n, F_\infty)$ for F_∞ as in (1.4) and

$$F_n = \sum_{i=1}^{\infty} \alpha_{n,i}(N_i^2 - 1)$$

still with $\{N_i\}_{i\geq 1}$ i.i.d. standard Gaussian and now $\{\alpha_{n,i}\}_{i\geq 1}$ not necessarily distinct real numbers. As mentioned above, the fact that we bound the Wasserstein-2 distance is not anecdotal : this distance (which is useful in many important settings, see [31]) does not bear a dual representation of the form (1.1) and is thus entirely out of the scope of the traditional versions of Stein's method. In Section 3.2 we provide an intuitive explanation of the proof of our main results. In Section 3.3 we apply our bounds to particular cases and compare them to the only competitor bounds available in the literature which are due to [12], wherein only the case (1.4) with $q = 2$ and $\alpha_1 = -\alpha_2$ is covered. Finally in Section 3.3 we provide the proof.

2 STEIN'S METHOD FOR THE SECOND WIENER CHAOS

Here we set up Stein's method for target distributions in the second Wiener chaos of the form

$$F_\infty = \alpha_{\infty,1}(N_1^2 - 1) + \cdots + \alpha_{\infty,q}(N_q^2 - 1)$$

where $\{N_i\}_{i=1}^q$ is a family of i.i.d. $\mathscr{N}(0,1)$ random variables.

2.1 Overview of known results

In the special case when $\alpha_{\infty,i} = 1$ for all i, then $F_\infty = \sum_{i=1}^q (N_i^2 - 1) \sim \chi^2_{(q)}$ is a centered chi-squared random variable with q degree of freedom. Pickett [26] has shown that a Stein's equation for target distribution F_∞ is given by the first order differential equation

$$xf'(x) + \frac{1}{2}(q - x)f(x) = h(x) - \mathbb{E}[h(F_\infty)].$$

For more recent results in this direction consult [16] and references therein.

Another important contribution in our direction is given by Gaunt in [14] with $q = 2$ and $\alpha_{\infty,1} = -\alpha_{\infty,2} = \frac{1}{2}$. In this case

$$F_\infty \overset{\text{law}}{=} N_1 \times N_2$$

where N_1 and N_2 are two independent $\mathscr{N}(0,1)$ random variables. He has shown that a Stein's equation for F_∞ can be given by the following second order differential equation

$$xf''(x) + f'(x) - xf(x) = h(x) - \mathbb{E}[h(F_\infty)].$$

It is a well known fact that the density function of the target random variable $N_1 \times N_2$ is expressible in terms of the modified Bessel function of the second kind so that it is given by solution of a known second order differential equation and the Stein operator follows from some form of duality argument.

The more relevant studies of target distributions having a second order Stein's differential equations include: Variance-Gamma distribution [15], Laplace distribution [27], or a family of probability distributions given by densities

$$f_s(x) = \Gamma(s)\sqrt{\frac{2}{s\pi}}\exp(-\frac{x^2}{2s})U(s-1, \frac{1}{2}, \frac{x^2}{2s}), \quad x > 0,\ s \geq \frac{1}{2}$$

appearing in preferential attachment random graphs, and $U(a, b; x)$ is the Kummer's confluent hypergeometric function, see [25]. We stress the fact that Stein's method is completely open even in the simple case when the target random variable F_∞ has only two non-zero eigenvalues $\alpha_{\infty,1}$ and $\alpha_{\infty,2}$, i.e.

$$F_\infty = \alpha_{\infty,1}(N_1^2 - 1) + \alpha_{\infty,2}(N_2^2 - 1)$$

such that $|\alpha_{\infty,1}| \neq |\alpha_{\infty,2}|$. It is worth mentioning that such distributions are beyond the Variance-Gamma class, and are appearing more and more in very recent and delicate limit theorems, see [5] for asymptotic behavior of generalized Rosenblatt process at extreme critical exponents and [19] for asymptotic nodal length distributions.

2.2 Fourier-based approach

Before stating the next theorem, we need to introduce some notations. For any d-tuple $(\lambda_1, ..., \lambda_d)$ of real numbers, we define the symmetric elementary polynomial of order $k \in \{1, ..., d\}$ evaluated at $(\lambda_1, ..., \lambda_d)$ by:

$$e_k(\lambda_1, ..., \lambda_d) = \sum_{1 \leq i_1 < i_2 < ... < i_k \leq d} \lambda_{i_1} ... \lambda_{i_k}.$$

We set, by convention, $e_0(\lambda_1, ..., \lambda_d) = 1$. Moreover, for any $(\mu_1, ..., \mu_d) \in \mathbb{R}^*$ and any $k \in \{1, ..., d\}$, we denote by $(\lambda/\mu)_k$ the $d-1$ tuple defined by:

$$(\frac{\lambda}{\mu})_k = \Big(\frac{\lambda_1}{\mu_1}, ..., \frac{\lambda_{k-1}}{\mu_{k-1}}, \frac{\lambda_{k+1}}{\mu_{k+1}}, ..., \frac{\lambda_d}{\mu_d}\Big).$$

For any $(\alpha, \mu) \in \mathbb{R}_+^*$, we denote by $\gamma(\alpha, \mu)$ a Gamma law with parameters (α, μ) whose density is:

$$\forall x \in \mathbb{R}_+^*, \ \gamma_{\alpha,\mu}(x) = \frac{\mu^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\big(-\mu x\big).$$

Theorem 2.1. *Let $d \geq 1$ and $(m_1, \dots, m_d) \in \mathbb{N}^d$. Let $((\alpha_1, \mu_1), ..., (\alpha_d, \mu_d)) \in (\mathbb{R}_+^*)^{2d}$ and $(\lambda_1, \dots, \lambda_d) \in \mathbb{R}^\star$ and consider:*

$$F = -\sum_{i=1}^d \lambda_i \frac{m_i \alpha_i}{\mu_i} + \sum_{i=1}^d \lambda_i \gamma_i(m_i \alpha_i, \mu_i),$$

where $\{\gamma_i(m_i\alpha_i, \mu_i)\}$ is a collection of independent gamma random variables with appropriate parameters. Let Y be a real valued random variable such that $\mathbb{E}[|Y|] < +\infty$. Then $Y \overset{law}{=} F$ if and only if

$$\mathbb{E}\Bigg[(Y + \sum_{i=1}^d \lambda_i \frac{m_i\alpha_i}{\mu_i})(-1)^d \Bigg(\prod_{j=1}^d \frac{\lambda_j}{\mu_j}\Bigg)\phi^{(d)}(Y) + \sum_{l=1}^{d-1} (-1)^l \Bigg(Y e_l(\frac{\lambda_1}{\mu_1}, ..., \frac{\lambda_d}{\mu_d}) + \sum_{k=1}^d \lambda_k \frac{m_k \alpha_k}{\mu_k} \Big(e_l(\frac{\lambda_1}{\mu_1}, ..., \frac{\lambda_d}{\mu_d}) - e_l((\frac{\lambda}{\mu})_k)\Big)\Bigg)\phi^{(l)}(Y) + Y\phi(Y)\Bigg] = 0, \tag{2.1}$$

for all $\phi \in S(\mathbb{R})$.

Proof. $(\Rightarrow)$. Let F be as in the statement of the theorem. We denote by $J_+ = \{j \in \{1, ..., d\} : \lambda_j > 0\}$ and $J_- = \{j \in \{1, ..., d\} : \lambda_j < 0\}$. Let us compute the characteristic function of F. For any $\xi \in \mathbb{R}$, we have:

$$\begin{aligned}
\phi_F(\xi) &= \mathbb{E}[\exp(i\xi F)], \\
&= \exp\Big(-i\xi\sum_{k=1}^{d}\lambda_k\frac{m_k\alpha_k}{\mu_k}\Big)\prod_{j=1}^{d}\mathbb{E}\Big[\exp\Big(i\xi\lambda_j\gamma_j(m_j\alpha_j,\mu_j)\Big)\Big], \\
&= \exp\Big(-i\xi < m\alpha;\lambda/\mu >\Big)\prod_{j\in J_+}\Big(\exp\Big(\int_0^{+\infty}\Big(e^{i\xi\lambda_j x}-1\Big)\Big(\frac{m_j\alpha_j}{x}e^{-\mu_j x}\Big)dx\Big)\Big) \\
&\times\prod_{j\in J_-}\Big(\exp\Big(\int_0^{+\infty}\Big(e^{i\xi\lambda_j x}-1\Big)\Big(\frac{m_j\alpha_j}{x}e^{-x\mu_j}\Big)dx\Big)\Big), \\
&= \exp\Big(-i\xi < m\alpha;\lambda/\mu >\Big)\exp\Big(\int_0^{+\infty}\Big(e^{i\xi x}-1\Big)\Big(\sum_{j\in J_+}\frac{m_j\alpha_j}{x}e^{-\frac{x\mu_j}{\lambda_j}}\Big)dx\Big) \\
&\times\exp\Big(\int_0^{+\infty}\Big(e^{-i\xi x}-1\Big)\Big(\sum_{j\in J_-}\frac{m_j\alpha_j}{x}e^{-\frac{x\mu_j}{(-\lambda_j)}}\Big)dx\Big),
\end{aligned}$$

where we have used the LÃl'vy-Khintchine representation of the Gamma distribution. We denote μ_j/λ_j by ν_j. Differentiating with respect to ξ together with standard computations, we obtain:

$$\prod_{k=1}^{d}(\nu_k-i\xi)\frac{d}{d\xi}\Big(\phi_F(\xi)\Big) = \Big[-i < m\alpha,\lambda/\mu > \prod_{k=1}^{d}(\nu_k-i\xi)+i\sum_{k=1}^{d}m_k\alpha_k\prod_{l=1,l\neq k}^{d}(\nu_l-i\xi)\Big]\phi_F(\xi).$$

Let us introduce two differential operators characterized by their symbols in Fourier domain. For smooth enough test functions, ϕ, we define:

$$\begin{aligned}
\mathcal{A}_{d,\nu}(\phi)(x) &= \frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}(\phi)(\xi)\Big(\prod_{k=1}^{d}(\nu_k-i\xi)\Big)\exp(ix\xi)d\xi, \\
\mathcal{B}_{d,m,\nu}(\phi)(x) &= \frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}(\phi)(\xi)\Big(\sum_{k=1}^{d}m_k\alpha_k\prod_{l=1,l\neq k}^{d}(\nu_l-i\xi)\Big)\exp(ix\xi)d\xi, \\
\mathcal{F}(\phi)(\xi) &= \int_{\mathbb{R}}\phi(x)\exp(-ix\xi)dx.
\end{aligned}$$

Integrating against smooth test functions the differential equation satistifed by the characteristic function ϕ_F, we have, for the left hand side:

$$\begin{aligned}
\int_{\mathbb{R}}\mathcal{F}(\phi)(\xi)\Big(\prod_{k=1}^{d}(\nu_k-i\xi)\Big)\frac{d}{d\xi}\Big(\phi_F(\xi)\Big)d\xi &= \int_{\mathbb{R}}\mathcal{F}(\mathcal{A}_{d,\nu}(\phi))(\xi)\frac{d}{d\xi}\Big(\phi_F(\xi)\Big)d\xi, \\
&= -\int_{\mathbb{R}}\frac{d}{d\xi}\Big(\mathcal{F}(\mathcal{A}_{d,\nu}(\phi))(\xi)\Big)\phi_F(\xi)d\xi, \\
&= i\int_{\mathbb{R}}\mathcal{F}(x\mathcal{A}_{d,\nu}(\phi))(\xi)\phi_F(\xi)d\xi,
\end{aligned}$$

where we have used the standard fact $d/d\xi(\mathcal{F}(f)(\xi)) = -i\mathcal{F}(xf)(\xi)$. Similarly, for the right hand side, we obtain:

$$\text{RHS} = \int_{\mathbb{R}} \mathcal{F}(\phi)(\xi)\Big[- i < m\alpha, \lambda/\mu > \prod_{k=1}^{d} (\nu_k - i\xi) + i \sum_{k=1}^{d} m_k \alpha_k \prod_{l=1, l\neq k}^{d} (\nu_l - i\xi)\Big] \phi_F(\xi) d\xi,$$
$$= i \int_{\mathbb{R}} \mathcal{F}(- < m\alpha, \lambda/\mu > \mathcal{A}_{d,\nu}(\phi) + \mathcal{B}_{d,m,\nu}(\phi))(\xi)\phi_F(\xi)d\xi.$$

Thus,

$$\int_{\mathbb{R}} \mathcal{F}((x+ < m\alpha, \lambda/\mu >)\mathcal{A}_{d,\nu}(\phi) - \mathcal{B}_{d,m,\nu}(\phi))(\xi)\phi_F(\xi)d\xi = 0$$

Going back in the space domain, we obtain the following Stein-type characterization formula:

$$\mathbb{E}[(F+ < m\alpha, \lambda/\mu >)\mathcal{A}_{d,\nu}(\phi)(F) - \mathcal{B}_{d,m,\nu}(\phi)(F)] = 0.$$

In order to conclude the first half of the proof, we need to compute explicitely the coefficients of the operators $\mathcal{A}_{d,\nu}$ and $\mathcal{B}_{d,m,\nu}$ in the following expansions:

$$\mathcal{A}_{d,\nu} = \sum_{k=0}^{d} a_k \frac{d^k}{dx^k},$$
$$\mathcal{B}_{d,m,\nu} = \sum_{k=0}^{d-1} b_k \frac{d^k}{dx^k}.$$

First of all, let us consider the following polynomial in $\mathbb{R}[X]$:

$$P(x) = \prod_{j=1}^{d} (\nu_j - x) = (-1)^d \prod_{j=1}^{d} (x - \nu_j).$$

We denote by $p_0, ..., p_d$ the coefficients of $\prod_{j=1}^{d}(X - \nu_j)$ in the basis $\{1, X, ..., X^d\}$. Vieta formula readily give:

$$\forall k \in \{0, ..., d\},\ p_k = (-1)^{d+k} e_{d-k}(\nu_1, ..., \nu_d),$$

It follows that the Fourier symbol of $\mathcal{A}_{d,\nu}$ is given by:

$$\prod_{k=1}^{d} (\nu_k - i\xi) = P(i\xi) = \sum_{k=0}^{d} (-1)^k e_{d-k}(\nu_1, ...\nu_d)(i\xi)^k.$$

Thus, we have, for ϕ smooth enough:

$$\mathcal{A}_{d,\nu}(\phi)(x) = \sum_{k=0}^{d} (-1)^k e_{d-k}(\nu_1, ..., \nu_d)\phi^{(k)}(x).$$

Let us proceed similarly for the operator $B_{d,m,\nu}$. We denote by P_k the following polynomial in $\mathbb{R}[X]$ (for any $k \in \{1, ..., d\}$):

$$P_k(x) = (-1)^{d-1} \prod_{l=1, l\neq k}^{d} (x - \nu_l).$$

A similar argument provides the following expression:

$$P_k(x) = \sum_{l=0}^{d-1} (-1)^l e_{d-1-l}(\underline{\nu}_k) x^l,$$

where $\underline{\nu}_k = (\nu_1, ..., \nu_{k-1}, \nu_{k+1}, ..., \nu_d)$. Thus, the symbol of the differential operator $B_{d,m,\nu}$ is given by:

$$\sum_{k=1}^{d} m_k \alpha_k \prod_{l=1, l\neq k}^{d} (\nu_l - i\xi) = \sum_{l=0}^{d-1} (-1)^l \bigg(\sum_{k=1}^{d} m_k \alpha_k e_{d-1-l}(\underline{\nu}_k) \bigg) (i\xi)^l.$$

Thus, we have:

$$B_{d,m,\nu}(\phi)(x) = \sum_{l=0}^{d-1} (-1)^l \bigg(\sum_{k=1}^{d} m_k \alpha_k e_{d-1-l}(\underline{\nu}_k) \bigg) \phi^{(k)}(x).$$

Consequently, we obtain:

$$\mathbb{E}[(F+ < m\alpha, \lambda/\mu >) \sum_{k=0}^{d} (-1)^k e_{d-k}(\nu_1, ..., \nu_d) \phi^{(k)}(F)$$
$$- \sum_{l=0}^{d-1} (-1)^l \bigg(\sum_{k=1}^{d} m_k \alpha_k e_{d-1-l}(\underline{\nu}_k) \bigg) \phi^{(k)}(F)] = 0.$$

Finally, there is a straightforward relationship between $e_k(\nu_1, ..., \nu_d)$ and $e_{d-k}(\lambda_1/\mu_1, ..., \lambda_d/\mu_d)$. Namely,

$$\forall k \in \{0, ..., d\},\ e_k(\nu_1, ..., \nu_d) = \frac{\prod_{j=1}^{d} \mu_j}{\prod_{j=1}^{d} \lambda_j} e_{d-k}(\frac{\lambda_1}{\mu_1}, ..., \frac{\lambda_d}{\mu_d}).$$

Thus, multiplying by $\prod_{j=1}^{d} \lambda_j / \prod_{j=1}^{d} \mu_j$, the previous Stein-type characterisation equation, we have:

$$\mathbb{E}\bigg[(F+ < m\alpha, \lambda/\mu >)(-1)^d \bigg(\prod_{j=1}^{d} \frac{\lambda_j}{\mu_j} \bigg) \phi^{(d)}(F) + \sum_{l=1}^{d-1} (-1)^l \bigg(F e_l(\frac{\lambda_1}{\mu_1}, ..., \frac{\lambda_d}{\mu_d})$$
$$+ \sum_{k=1}^{d} \lambda_k m_k \frac{\alpha_k}{\mu_k} (e_l(\frac{\lambda_1}{\mu_1}, ..., \frac{\lambda_d}{\mu_d}) - e_l((\frac{\lambda}{\mu})_k)) \bigg) \phi^{(l)}(F) + F\phi(F) \bigg] = 0.$$

($\Leftarrow$) Let Y be a real valued random variable such that $\mathbb{E}[|Y|] < +\infty$ and:

$$\forall \phi \in S(\mathbb{R}),\ \mathbb{E}\bigg[(Y+ < m\alpha, \lambda/\mu >)(-1)^d \bigg(\prod_{j=1}^{d} \frac{\lambda_j}{\mu_j} \bigg) \phi^{(d)}(Y) + \sum_{l=1}^{d-1} (-1)^l \bigg(Y e_l(\frac{\lambda_1}{\mu_1}, ..., \frac{\lambda_d}{\mu_d})$$
$$+ \sum_{k=1}^{d} \lambda_k m_k \frac{\alpha_k}{\mu_k} (e_l(\frac{\lambda_1}{\mu_1}, ..., \frac{\lambda_d}{\mu_d}) - e_l((\frac{\lambda}{\mu})_k)) \bigg) \phi^{(l)}(Y) + Y\phi(Y) \bigg] = 0.$$

By the previous step, this implies that:

$$\forall \phi \in S(\mathbb{R}),\ \int_{\mathbb{R}} \mathcal{F}((x+ < m\alpha, \lambda/\mu >)\mathcal{A}_{d,\nu}(\phi) - \mathcal{B}_{d,m,\nu}(\phi))(\xi) \phi_Y(\xi) d\xi = 0,$$
$$\Leftrightarrow \int_{\mathbb{R}} \mathcal{F}(x\mathcal{A}_{d,\nu}(\phi))(\xi)\phi_Y(\xi)d\xi = \int_{\mathbb{R}} \mathcal{F}(- < m\alpha, \lambda/\mu > \mathcal{A}_{d,\nu}(\phi) + \mathcal{B}_{d,m,\nu}(\phi))(\xi)\phi_Y(\xi)d\xi,$$
$$\Leftrightarrow \prod_{k=1}^{d} (\nu_k - i\xi) \frac{d}{d\xi} \bigg(\phi_Y \bigg)(.) = \bigg[-i < m\alpha, \lambda/\mu > \prod_{k=1}^{d} (\nu_k - i\xi) + i \sum_{k=1}^{d} \alpha_k m_k \prod_{l=1, l\neq k}^{d} (\nu_l - i\xi) \bigg] \phi_Y(.),$$

in $S'(\mathbb{R})$. Since $\mathbb{E}[|Y|] < +\infty$, the characteristic function of Y is differentiable on the whole real line so that:

$$\forall \xi \in \mathbb{R},\ \frac{d}{d\xi}\Big(\phi_Y\Big)(\xi) = \Big[- i < m\alpha, \lambda/\mu > +i \sum_{k=1}^{d} m_k \alpha_k \frac{1}{\nu_k - i\xi}\Big] \phi_Y(\xi)$$

Moreover, we have $\phi_Y(0) = 1$. Thus, by Cauchy-Lipschitz theorem, we have:

$$\forall \xi \in \mathbb{R},\ \phi_Y(\xi) = \phi_F(\xi).$$

This concludes the proof of the theorem. □

Taking $\alpha_k = \mu_k = 1/2$ in the previous theorem implies the following straightforward corollary:

Corollary 2.1. *Let $d \geq 1$, $q \geq 1$ and $(m_1, \ldots, m_d) \in \mathbb{N}^d$ such that $m_1 + ... + m_d = q$. Let $(\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^\star$ and consider:*

$$F = \sum_{i=1}^{m_1} \lambda_1 (N_i^2 - 1) + \sum_{i=m_1+1}^{m_1+m_2} \lambda_2 (N_i^2 - 1) + ... + \sum_{i=m_1+...+m_{d-1}+1}^{q} \lambda_d (N_i^2 - 1),$$

Let Y be a real valued random variable such that $\mathbb{E}[|Y|] < +\infty$. Then $Y \overset{law}{=} F$ if and only if

$$\mathbb{E}\Big[(Y + \sum_{i=1}^{d} \lambda_i m_i)(-1)^d 2^d \Big(\prod_{j=1}^{d} \lambda_j\Big) \phi^{(d)}(Y) + \sum_{l=1}^{d-1} 2^l (-1)^l \Big(Y e_l(\lambda_1, ..., \lambda_d)$$
$$+ \sum_{k=1}^{d} \lambda_k m_k \left(e_l(\lambda_1, ..., \lambda_d) - e_l((\underline{\lambda}_k)\right) \Big) \phi^{(l)}(Y) + Y \phi(Y)\Big] = 0, \tag{2.2}$$

for all $\phi \in S(\mathbb{R})$.

Proof. Let F be as in the statement of the theorem. Then, it is sufficient to observe that we have the following equality in law:

$$F \overset{\text{law}}{=} -\sum_{k=1}^{d} m_k \lambda_k + \sum_{i=1}^{d} \lambda_i \gamma_i (\frac{m_i}{2}, \frac{1}{2}).$$

To end the proof of the corollary, we apply the previous theorem with $\alpha_k = \mu_k = 1/2$ for every k. □

Example 2.1. Let $d = 1$, $m_1 = q \geq 1$ and $\lambda_1 = \lambda > 0$. The differential operator reduces to (on smooth test function ϕ):

$$-2\lambda(x + q\lambda)\phi^{(1)}(x) + x\phi(x).$$

This differential operator is similar to the one characterising the gamma distribution of parameters $(q/2, 1/(2\lambda))$. Indeed, we have, for $F \overset{\text{law}}{=} \gamma(q/2, 1/(2\lambda))$, on smooth test function, ϕ:

$$\mathbb{E}\Big[F\phi^{(1)}(F) + (\frac{q}{2} - \frac{F}{2\lambda})\phi(F)\Big] = 0$$

We can move from the first differential operator to the second one by performing a scaling of parameter $-1/(2\lambda)$ and the change of variable $x = y - q\lambda$.

Example 2.2. Let $d = 2,\ q = 2,\ \lambda_1 = -\lambda_2 = 1/2$ and $m_1 = m_2 = 1$. The differential operator reduces to (on smooth test function ϕ):

$$\begin{aligned} T(\phi)(x) &= 4(x+ < m, \lambda >)\lambda_1\lambda_2\phi^{(2)}(x) - 2[xe_1(\lambda_1, \lambda_2) + \lambda_1 m_1(e_1(\lambda_1, \lambda_2) - e_1(\lambda_2)) \\ &\quad + \lambda_2 m_2(e_1(\lambda_1, \lambda_2) - e_1(\lambda_1))]\phi^{(1)}(x) + x\phi(x), \\ &= -x\phi^{(2)}(x) - \phi^{(1)}(x) + x\phi(x), \end{aligned}$$

where we have used the fact that $e_1(\lambda_1, \lambda_2) = \lambda_1 + \lambda_2 = 0,\ e_1(\lambda_2) = \lambda_2 = -1/2,\ e_1(\lambda_1) = \lambda_1 = 1/2$. Therefore, up to a minus sign factor, we retrieve the differential operator associated with the random variable:

$$F = N_1 \times N_2.$$

2.3 Malliavin-based approach

In this section, we assume that the random objects we consider do live in the Wiener space. Let $\mathrm{X} = \{\mathrm{X(h)};\ \mathrm{h} \in \mathfrak{H}\}$ stand for an isonormal process over a separable Hilbert space $\mathfrak{H}$. The reader may consult [20, Chapter 2] for a detailed discussion on this topic. The main aim of this section is to use Malliavin calculus on the Wiener space to obtain a Stein characterization for target random variables of the form (1.4). The following definition includes the iterated Malliavin Γ-operators that lie at the core of this approach. The notation $\mathbb{D}^\infty$ stands for the class of infinitely many times Malliavin differentiable random variables.

Definition 2.1 (see [20]). *Let $F \in \mathbb{D}^\infty$. The sequence of random variables $\{\Gamma_i(F)\}_{i\geq 0} \subset \mathbb{D}^\infty$ is recursively defined as follows. Set $\Gamma_0(F) = F$ and, for every $i \geq 1$,*

$$\Gamma_i(F) = \langle DF, -DL^{-1}\Gamma_{i-1}(F)\rangle_{\mathfrak{H}}.$$

For instance, one has that $\Gamma_1(F) = \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}} = \tau_F(F)$ the Stein kernel of F.

For further use, we also recall that (see again [20]) the cumulants of the random element F and the iterated Malliavin Γ- operators are linked by the relation

$$\kappa_{r+1}(F) = r!\,\mathbb{E}[\Gamma_r(F)] \text{ for } r = 0, 1, \cdots.$$

Following [24, 3], we define two crucial polynomials P and Q as follows:

$$Q(x) = (P(x))^2 = \Big(x\prod_{i=1}^{q}(x - \alpha_{\infty,i})\Big)^2. \tag{2.3}$$

Finally, for any random element F, we define the following quantity (whose first appearance is in [3])

$$\Delta(F, F_\infty) := \sum_{r=2}^{\deg(Q)} \frac{Q^{(r)}(0)}{r!}\frac{\kappa_r(F)}{2^{r-1}(r-1)!}. \tag{2.4}$$

Proposition 2.1. *[3, Proposition 3.2] Let F be a centered random variable living in a finite sum of Wiener chaoses. Moreover, assume that*

(i) *$\kappa_r(F) = \kappa_r(F_\infty)$, for all $2 \leq r \leq k + 1 = deg(P)$, and*

(ii)

$$\mathbb{E}\left[\sum_{r=1}^{k+1} \frac{P^{(r)}(0)}{r!\ 2^{r-1}} \Big(\Gamma_{r-1}(F) - \mathbb{E}[\Gamma_{r-1}(F)]\Big)\right]^2 = 0.$$

Then, $F \overset{\text{law}}{=} F_\infty$, *and* F *belongs to the second Wiener chaos.*

In fact item **(ii)** of Proposition 2.1 can be used to derive a Stein equation for F_∞. To this end, set

$$a_l = \frac{P^{(l)}(0)}{l!\, 2^{l-1}} \quad 1 \leq l \leq q+1,$$
$$b_l = \sum_{r=l}^{q+1} a_r \mathbb{E}[\Gamma_{r-l+1}(F_\infty)] = \sum_{r=l}^{q+1} \frac{a_r}{(r-l+1)!} \kappa_{r-l+2}(F_\infty) \quad 2 \leq l \leq q+1$$

Now, we introduce the following differential operator of order q (acting on functions $f \in C^q(\mathbb{R})$) :

$$\mathcal{A}_\infty f(x) := \sum_{l=2}^{q+1} (b_l - a_{l-1}x) f^{(q+2-l)}(x) - a_{q+1} x f(x). \tag{2.5}$$

Then, we have the following result.

Theorem 2.2. *Assume that* F *is a general centered random variable living in a finite sum of Wiener chaoses (and hence smooth in the sense of Malliavin calculus). Then*

$$F \overset{law}{=} F_\infty$$

if and only if $\mathbb{E}\left[\mathcal{A}_\infty f(F)\right] = 0$ *for all mappings* $f : \mathbb{R} \to \mathbb{R}$ *such that* $\mathbb{E}\left[|\mathcal{A}_\infty f(F)|\right] < \infty$, *and moreover* $\mathbb{E}[f^{(q)}(F)^2] < +\infty$.

Proof. Repeatedly using the Malliavin integration by parts formulae [20, Theorem 2.9.1], we obtain for any $2 \leq l \leq q+2$ that

$$\mathbb{E}\left[F f^{(q-l+2)}(F)\right] = \mathbb{E}\left[f^{(q)}(F)\Gamma_{l-2}(F)\right] + \sum_{r=q-l+3}^{q-1} \mathbb{E}\left[f^{(r)}(F)\right] \mathbb{E}\left[\Gamma_{r+l-q-2}(F)\right]. \tag{2.6}$$

For indices $l = 2, 3$, the second term in the right hand side of (2.6) is understood to be 0. Summing from $l = 2$ up to $l = q+2$, we obtain that

$$
\begin{aligned}
\sum_{l=2}^{q+2} a_{l-1}\, \mathbb{E}\left[F f^{(q-l+2)}(F)\right] &= \sum_{l=2}^{q+2} a_{l-1}\, \mathbb{E}\left[f^{(q)}(F)\Gamma_{l-2}(F)\right] \\
&\quad + \sum_{l=4}^{q+2} a_{l-1} \sum_{r=q-l+3}^{q-1} \mathbb{E}\left[f^{(r)}(F)\right] \mathbb{E}\left[\Gamma_{r+l-q-2}(F)\right] \\
&= \sum_{l=1}^{q+1} a_l\, \mathbb{E}\left[f^{(q)}(F)\Gamma_{l-1}(F)\right] \\
&\quad + \sum_{l=3}^{q+1} a_l \sum_{r=q-l+2}^{q-2} \mathbb{E}\left[f^{(r)}(F)\right] \mathbb{E}\left[\Gamma_{r+l-q-1}(F)\right] \\
&= \sum_{l=1}^{q+1} a_l\, \mathbb{E}\left[f^{(q)}(F)\Gamma_{l-1}(F)\right] \\
&\quad + \sum_{l=2}^{q+1} a_l \sum_{r=1}^{l-2} \mathbb{E}\left[f^{(q-r)}(F)\right] \mathbb{E}\left[\Gamma_{l-r-1}(F)\right].
\end{aligned} \tag{2.7}
$$

On the other hand,

$$
\begin{aligned}
\sum_{l=2}^{q+1} b_l\, \mathbb{E}\left[f^{(q+2-l)}(F)\right] &= \sum_{l=0}^{q-1} b_{l+2}\mathbb{E}\left[f^{(q-l)}(F)\right] \\
&= \sum_{l=0}^{q-1} \left[\sum_{r=l+2}^{q+1} a_r \mathbb{E}(\Gamma_{r-l-1}(F_\infty))\right] \mathbb{E}\left[f^{(q-l)}(F)\right] \\
&= \sum_{r=2}^{q+1} a_r \sum_{l=0}^{r-2} \mathbb{E}\left[\Gamma_{r-l-1}(F_\infty)\right] \times \mathbb{E}\left[f^{(q-l)}(F)\right].
\end{aligned} \tag{2.8}
$$

Wrapping up, we finally arrive at

$$
\begin{aligned}
\mathbb{E}\left[\mathcal{A}_\infty f(F)\right] &= -\mathbb{E}\left[f^{(q)}(F) \times \Big(\sum_{r=1}^{q+1} a_r \left[\Gamma_{r-1}(F) - \mathbb{E}[\Gamma_{r-1}(F)]\right]\Big)\right] \\
&\quad + \sum_{r=2}^{q+1} a_r \sum_{l=0}^{r-2} \Big\{\mathbb{E}[f^{(q-l)}(F)] \times \Big(\mathbb{E}\left[\Gamma_{r-l-1}(F_\infty)\right] - \mathbb{E}\left[\Gamma_{r-l-1}(F)\right]\Big)\Big\} \\
&= -\mathbb{E}\left[f^{(q)}(F) \times \Big(\sum_{r=1}^{q+1} a_r \left[\Gamma_{r-1}(F) - \mathbb{E}[\Gamma_{r-1}(F)]\right]\Big)\right] \\
&\quad + \sum_{r=2}^{q+1} a_r \sum_{l=0}^{r-2} \frac{\mathbb{E}[f^{(q-l)}(F)]}{(r-l-1)!} \times \Big(\kappa_{r-l}(F_\infty) - \kappa_{r-l}(F)\Big).
\end{aligned} \tag{2.9}
$$

We are now in a position to prove the claim. First we assume that $F \overset{\text{law}}{=} F_\infty$. Then obviously $\kappa_r(F) = \kappa_r(F_\infty)$ for $r = 2, \cdots, 2q+2$. Following the same arguments as in the proof of [3, Proposition 3.2], one can infer that, in fact, F belongs to the second Wiener chaos. Hence, according to [3, Lemma 3.1], and the Cauchy–Schwarz inequality, we obtain that

$$|\mathbb{E}[\mathcal{A}_\infty f(F)]| \leq \sqrt{\mathbb{E}\left[f^{(q)}(F)\right]^2} \times \sqrt{\mathbb{E}\Big[\sum_{r=1}^{q+1} a_r \left(\Gamma_{r-1}(F) - \mathbb{E}[\Gamma_{r-1}(F)]\right)\Big]^2}$$
$$= \sqrt{\mathbb{E}\left[f^{(q)}(F)\right]^2} \times \sqrt{\Delta(F, F_\infty)}$$
$$= \sqrt{\mathbb{E}\left[f^{(q)}(F)\right]^2} \times \sqrt{\Delta(F_\infty, F_\infty)} = 0.$$

Conversely, assume that $\mathbb{E}[\mathcal{A}_\infty f(F)] = 0$ for all the suitable functions f. Then relation (2.9) implies that, by choosing appropriate polynomials for function f, we have $\kappa_r(F) = \kappa_r(F_\infty)$ for $r = 2, \cdots, q+1$. Now, combining this observation together with relation (2.9), we infer that

$$\mathbb{E}\left[\sum_{r=1}^{q+1} a_r \Big(\Gamma_{r-1}(F) - \mathbb{E}[\Gamma_{r-1}(F)]\Big)\Big| F\right] = 0.$$

Using e.g. integrations by parts, the latter equation can be turned into a linear recurrent relation between the cumulants of F of order up to $q+1$. Combining this with the knowledge of the $q+1$ first cumulants characterize all the cumulants of F and hence the distribution F. Indeed, all the distributions in the second Wiener chaos are determined by their moments. □

Example 2.3. Consider the special case of only two non-zero distinct eigenvalues λ_1 and λ_2, i.e.

$$F_\infty = \lambda_1(N_1^2 - 1) + \lambda_2(N_2^2 - 1) \tag{2.10}$$

where $N_1, N_2 \sim \mathscr{N}(0,1)$ are independent. In this case, the polynomial P takes the form $P(x) = x(x-\alpha)(x-\beta)$. Simple calculations reveal that $P'(0) = \lambda_1\lambda_2, P''(0) = -2(\lambda_1+\lambda_2)$, and $P^{(3)}(0) = 3!$. Also, $\kappa_2(F_\infty) = \mathbb{E}[\Gamma_1(F_\infty)] = 2(\lambda_1^2 + \lambda_2^2)$, and $\kappa_3(F_\infty) = 2\mathbb{E}[\Gamma_2(F_\infty)] = 4(\lambda_1^3 + \lambda_2^3)$. Then, the Stein equation (2.5) reduces to that

$$\mathcal{A}_\infty f(x) = -4(\lambda_1\lambda_2 x + (\lambda_1+\lambda_2)\lambda_1\lambda_2)f''(x) + 2\left(\lambda_1^2 + \lambda_2^2 + (\lambda_1+\lambda_2)x\right)f'(x) - xf(x) \tag{2.11}$$

We also remark that when $\lambda_1 = -\lambda_2 = \frac{1}{2}$, and hence $F_\infty \overset{\text{law}}{=} N_1 \times N_2$, the Stein's equation (2.11) coincides with that in [14, equation (1.9)].

One can show that (2.5) is a specification of (2.2) in the particular setting considered in the current Section. The proof of this asserstion is quite technical and we do not include it. To convince the reader, let us inspect the particular case where $d = 2$, $q = 2$ and $\lambda_1 \neq \lambda_2$. Using the previous corollary, the differential operator, T, boils down to, (on smooth test function ϕ):

$$T(\phi)(x) = 4\lambda_1\lambda_2(x + \lambda_1 + \lambda_2)\phi^{(2)}(x) - 2[x(\lambda_1+\lambda_2) + \lambda_1^2 + \lambda_2^2]\phi^{(1)}(x) + x\phi(x),$$

which is, up to a minus sign factor, the differential operator appearing in Equation 2.11 with $\alpha = \lambda_1$ and $\beta = \lambda_2$. In the case of d different eigenvalues whose respective multiplicities are equal to 1, in order to switch from one operator to the other one, we must use the fundamental Newton-Girard identities linking the elementary symmetric polynomials valued at $(\lambda_1, ..., \lambda_d)$ together with the power sums valued at $(\lambda_1, ..., \lambda_d)$ (which reduce, up to a multiplicative factor, to the cumulants of F).

2.4 Cattywampus Stein's method

In this section, we first propose a heuristic to clarify why $\Delta(F, F_\infty)$ can be interpreted as a version of the Malliavin-Stein discrepancy, when F belongs to the second Wiener chaos. Let $\mathcal{H}$ be a family of test functions that at least characterizes the convergence in distribution, in the sense that

$$\mathrm{d}_{\mathcal{H}}(F, F_\infty) := \sup_{h\in\mathcal{H}} \Big| \mathbb{E}[h(F)] - \mathbb{E}[h(F_\infty)] \Big| \approx 0 \quad \text{if and only if} \quad F \overset{\text{law}}{\approx} F_\infty.$$

Let $h \in \mathcal{H}$ and take $\mathcal{A}_\infty$ as in (2.5). Consider the Stein equation

$$\mathcal{A}_\infty f(x) = h(x) - \mathbb{E}[h(F_\infty)] \qquad h \in \mathcal{H}. \tag{2.12}$$

Assumption 1. [Stein universality assumption] *For any $h \in \mathcal{H}$, equation (2.12) has a unique bounded solution f_h so that*

$$\|f_h^{(r)}\|_\infty < \infty \qquad \forall\, r = 1, \cdots, q \tag{2.13}$$

uniformly in h.

Now, we take a general centered random variable F, smooth in the sense of Malliavin differentiability, for example F belongs to a finite sum of Wiener chaoses. Then, under Assumption 1, the proof of Theorem 2.2 reveals that for some general constants C_1, C_2, using the Cauchy-Swartz inequality, one has

$$\begin{aligned}\sup_{h\in\mathcal{H}} \Big| \mathbb{E}[h(F)] - \mathbb{E}[h(F_\infty)] \Big| &\le C_1 \mathbb{E}\Big| \sum_{r=1}^{q+1} a_r \left(\Gamma_{r-1}(F) - \mathbb{E}[\Gamma_{r-1}(F)]\right) \Big| \\ &\quad + C_2 \sum_{r=2}^{q+1} |\kappa_r(F) - \kappa_r(F_\infty)| \\ &\le \sqrt{\mathbb{E}\Big[\sum_{r=1}^{q+1} a_r \left(\Gamma_{r-1}(F) - \mathbb{E}[\Gamma_{r-1}(F)]\right) \Big]^2} \\ &\quad + C_2 \sum_{r=2}^{q+1} |\kappa_r(F) - \kappa_r(F_\infty)|.\end{aligned} \tag{2.14}$$

However, taking into account [3, Lemma 3.1] when F itself belongs to the second Wiener chaos, then

$$\sup_{h\in\mathcal{H}} \Big| \mathbb{E}[h(F)] - \mathbb{E}[h(F_\infty)] \Big| \le C_1 \sqrt{\Delta(F, F_\infty)} + C_2 \sum_{r=2}^{q+1} |\kappa_r(F) - \kappa_r(F_\infty)|. \tag{2.15}$$

Starting from (2.15), to apply Stein's method in the second Wiener chaos it suffices, in a sense, to provide the estimates (2.13) required by Assumption 1.

Remark 2.1. *This idea is at the heart e.g. of [12] where the authors could make use of the estimates provided by [15] to apply the above plan to targets as in Example 2.1 (and, more generally, to variance-gamma distribution).*

There are two major flaws to this approach and, therefore, to the classical take on Stein's method as adapted to the second Wiener chaos. First the bounds required in (2.13) can only be obtained by solving q-th degree inhomogeneous equations and it is unlikely that a unified approach will allow to deal with all targets of the form (1.4) in one sweep. Indeed different ranges of α_i's imply very different properties for the corresponding F_∞ and hence each application of Stein's method to these targets will necessitate ad-hoc target specific solutions, which in the current state of knowledge we do not even know to be bounded. Second, the bounds on higher order derivatives of f_h will necessarily depend on smoothness assumptions on the test functions h; hence several important distances of the form (1.1) with non-smooth h cannot be tackled via Stein's method.

Remark 2.2. *Such a flaw was already noted for multivariate Gaussian approximation, see [7], where estimates in total variation distance were realized to be beyond the scope of the classical approaches to the method. Recently an original solution was proposed in [23] wherein a class of random vectors F was identified to which one could apply what we will coin an* information theoretic approach to Stein's method*; this resulted in a general fourth moment bound on Total Variation distance for multivariate normal approximation for random vectors F satisfying a very particular integrability constraint (see [23, Condition (2.53)]).*

3 BYPASSING THE STEIN'S METHOD

As mentioned in the conclusion to the previous section, bounding the solutions of higher order Stein's equations is an extremely hard task. However, if the appoximating sequence has the shape of a weighted sum of i.i.d. random variables, we provide a new strategy to bound the two Wasserstein distance between the sequence and the target. Not only we completely bypass the major difficulty of bounding the stein's solution but we can deal with distances which cannot be reached by usual Stein's tools.

Definition 3.1. *Fix $p \geq 1$. Given two probability measures ν and μ on the Borel sets of $\mathbb{R}^d$ whose marginals have finite absolute moments of order p, define the Wasserstein distance (of order p) between ν and μ as the quantity*

$$\mathrm{W}_p(\nu,\mu) \;=\; \inf_{\pi}\left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p d\pi(x,y)\right)^{1/p}$$

where the infimum runs over all probability measures π on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals ν and μ.

Remark 3.1. *The Wasserstein metric may be equivalently defined by*

$$\mathrm{W}_p(\nu,\mu) \;=\; (\inf \mathbb{E}\|X-Y\|_d^p)^{1/p}$$

where the infimum is taken over all joint distributions of the random variables X and Y with marginals μ and ν respectively, and $\|\,\|_d$ stands for the Euclidean norm on $\mathbb{R}^d$. It is also well-known that convergence with respect to W_p is equivalent to the usual weak convergence of measures plus convergence of the first pth moments. Also, a direct application of Hölder inequality implies that if $1 \leq p \leq q$ then $\mathrm{W}_p \leq \mathrm{W}_q$. Relevant information about Wasserstein distances can be found, e.g. in [31].

3.1 Main result

Throughout this section, we assume that $\{W_k\}_{k\geq 1}$ is a general sequence of independent and identically distributed random variables having **finite moments** of any order such that $\mathbb{E}[W_1] = 0$ and $\mathbb{E}[W_1^2] = 1$, and moreover $\kappa_r(W_1) \neq 0$ for all $r = 2, \cdots, q+1$. For a given sequence $\{\alpha_{n,k}\}_{n,k\geq 1} \subset \mathbb{R}$, we assume that each element F_n of the approximating sequence is of the form

$$F_n = \sum_{k\geq 1} \alpha_{n,k} W_k \qquad n \in \mathbb{N}. \tag{3.1}$$

Similarly, we assume that the target random variable may be written in the following way

$$F_\infty := \sum_{k=1}^{q} \alpha_{\infty,k} W_k. \tag{3.2}$$

where $q \geq 2$ and all the coefficients $\{\alpha_{\infty,k}\}_{1\leq k\leq q} \subset \mathbb{R}$ are both non-zero and pairwise distinct real numbers. Also, without loss of generality, we assume that we are dealing with normalized random variables, meaning that

$$\mathbb{E}[F_n^2] = \sum_{k\geq 1} \alpha_{n,k}^2 = 1 \quad \text{and} \quad \mathbb{E}[F_\infty^2] = \sum_{k=1}^{q} \alpha_{\infty,k}^2 = 1, \quad \forall\, n \in \mathbb{N}. \tag{3.3}$$

For any $n \in \mathbb{N}$, we introduce, in the shape of a lemma, a crucial quantity that can be also written as a finite linear combination of the first $2q+2$ cumulants of the random variable F_n of the approximating sequence.

Lemma 3.1. *For any $n \in \mathbb{N}$ we have*

$$\begin{aligned} \Delta(F_n, F_\infty) = \Delta(F_n) :&= \sum_{k\geq 1} \alpha_{n,k}^2 \prod_{r=1}^{q} (\alpha_{n,k} - \alpha_{\infty,r})^2 \\ &= \sum_{r=2}^{2q+2} \Theta_r \sum_{k\geq 1} \alpha_{n,k}^r \\ &= \sum_{r=2}^{2q+2} \frac{\Theta_r}{\kappa_r(W_1)} \kappa_r(F_n). \end{aligned} \tag{3.4}$$

where the coefficients Θ_r are the coefficients of the polynomial

$$Q(x) = (P(x))^2 = (x \prod_{i=1}^{q} (x - \alpha_{\infty,i}))^2. \tag{3.5}$$

Now, we are ready to sate our main results.

Theorem 3.1. *Let all notations and assumption in above are prevail. Then for some constant $C > 0$ depending only on target random variable F_∞ (and hence independent of n) so that*

$$d_{\mathrm{W}_2}(F_n, F_\infty) \leq C \left(\sqrt{\Delta(F_n)} + \sum_{r=2}^{q+1} |\kappa_r(F_n) - \kappa_r(F_\infty)| \right) \qquad \forall n \geq 1. \tag{3.6}$$

More precisely, there exists a threshold $U_\infty > 0$ and a constant $C > 0$ depending only on the target random variable F_∞ (and hence independent of n) such that the next bound

$$\sqrt{\Delta(F_n)} + \sum_{r=2}^{q+1} |\kappa_r(F_n) - \kappa_r(F_\infty)| \leq U_\infty, \tag{3.7}$$

implies

$$d_{\mathrm{W}_2}(F_n, F_\infty) \leq C \sqrt{\Delta(F_n)}. \tag{3.8}$$

In particular, if $\Delta(F_n) \to 0$ and moreover $\kappa_r(F_n) \to \kappa_r(F_\infty)$ for all $r = 2, \cdots, q+1$, then the threshold requirement (3.7) *takes place and therefore the sequence F_n converges in distribution towards target random variable F_∞ at rate $\sqrt{\Delta(F_n)}$.*

Theorem 3.2. *Let all the notations and assumptions of Theorem 3.1 prevail. Assume further that $\dim_{\mathbb{Q}}$ span $\{\alpha_{\infty,1}^2, \cdots, \alpha_{\infty,q}^2\} = q$. Then there exists a constant $C > 0$ depending only on the target random variable F_∞ such that*

$$d_{\mathrm{W}_2}(F_n, F_\infty) \leq C \sqrt{\Delta(F_n)}. \tag{3.9}$$

Remark 3.2. We remark that the condition $\dim_{\mathbb{Q}}$ span $\{\alpha_{\infty,1}^2, \cdots, \alpha_{\infty,q}^2\} = q$ implies that if $\sum_{i=1}^{q+1} n_i \alpha_{\infty,i}^2 = 1$ for some $n_i \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, then $n_1 = \cdots = n_{q+1} = 1$ according to normalization assumption (3.3). We will use this useful observation in the proof of Theorem 3.2.

Remark 3.3. In light of Theorem 3.1, it appears that if $\dim_{\mathbb{Q}} \operatorname{span}\left\{\alpha^2_{\infty,1}, \cdots, \alpha^2_{\infty,q}\right\} \neq q$ then even a standard application of Stein's method (that is, using Stein equations and hypothetical bounds on the solutions) would require the control of more moments than only the first two in order to bound the Wasserstein-1 distance.

Remark 3.4. According to Theorem 3.2 when $\dim_{\mathbb{Q}} \operatorname{span}\left\{\alpha^2_{\infty,1}, \cdots, \alpha^2_{\infty,q}\right\} = q$, one can drop the assumption of the separate convergences of the first $q+1$ cumulants, $\kappa_r(F_n) \to \kappa_r(F_\infty)$ for all $r = 2, \cdots, q+1$, in Theorem 3.1 and hence the only requirement $\Delta(F_n) \to 0$ implies the convergence in distribution of the sequence F_n towards the target random variable F_∞. For example, let $q = 2$ and

(1) $\alpha_{\infty,1} \in \mathbb{Q}$. Then, obviously $\alpha_{\infty,2} \in \mathbb{Q}$, and therefore $\dim_{\mathbb{Q}} \operatorname{span}\{\alpha^2_{\infty,1}, \alpha^2_{\infty,2}\} = 1 \neq q = 2$. Hence, for convergence in distribution of the sequence F_n towards the target random variable F_∞ in addition to convergence $\Delta(F_n) \to 0$ one needs also the convergence $\kappa_3(F_n) \to \kappa_3(F_\infty)$, see also Remark 3.5.

(2) $\alpha_{\infty,1} \in \mathbb{R} - \mathbb{Q}$ be a irrational number, and therefore according to normalization assumption (3.3) the coefficient $\alpha_{\infty,2}$ will be also an irrational number. In this case, we have $\dim_{\mathbb{Q}} \operatorname{span}\left\{\alpha^2_{\infty,1}, \alpha^2_{\infty,2}\right\} = q = 2$. Hence, the sole requirement $\Delta(F_n) \to 0$ is enough for convergence in distribution of the sequence F_n towards the target random variable F_∞.

3.2 Idea behind the proof

The main idea leading the proof relies on the following non trivial observation: $F_n = \sum_k \alpha_{n,k} W_k$ converges in distribution towards $F_\infty = \sum_{k=1}^q \alpha_{\infty,k} W_k$ if and only if one can find q coefficients among the sequence $\{\alpha_{n,k}\}_{k\geq 1}$ which are very close to the corresponding terms $\{\alpha_{\infty,k}\}_{1\leq k\leq q}$ and if the l^2-norm of the remaining terms is small. The main difficulty is to quantify this phenomenon by only using the Stein discrepancy $\Delta(F_n, F_\infty)$. To reach this goal, we proceed in several steps which are sketched below:

- Step 1: Without loss of generality, we can assume that, for each n, the sequence $|\alpha_{n,k}|$ decreases with k. Indeed, this will be useful to identify which coefficients have a non-zero limit and which coefficients go to zero.
- Step 2: Here we prove several inequalities which, roughly speaking, express the fact that if for some k the coefficient $|\alpha_{n,k}|$ is bounded from below by a fixed constant, then it is necessarily close to one of the limiting coefficients $\alpha_{\infty,k}$.
- Step 3: Here we prove that a certain number of the coefficients $\alpha_{n,k}$ (say $\alpha_{n,1}, \cdots, \alpha_{n,l_n}$) are all close to one element (say $\alpha_\infty(n,k)$) of the set $\{\alpha_{\infty,1}, \cdots, \alpha_{\infty,q}\}$ with possible repetitions. The difficult part consists in showing that $\sum_{k>l_n} \alpha^2_{n,k}$ is small. To do so, it is equivalent to prove that $\sum_{k=1}^{l_n} \alpha^2_{n,k}$ is close to one. Having in mind that $\sum_{k=1}^{l_n} |\alpha_{n,k} - \alpha_\infty(n,k)|$ is small, the latter claim is in turn equivalent to $\sum_{k=1}^{l_n} \alpha_\infty(n,k)^2 = 1$. We argue by contradiction using a maximality argument.
- Step 4: As we said in step 3, there might be repetitions among the coefficients $\alpha_\infty(n,k)$ and they may also vary with n. However, if the coefficients $\{\alpha^2_{\infty,k}\}_{1\leq k\leq q}$ are rationally independent, then there is only one way to chose $\alpha_\infty(n,k)$ in the set $\{\alpha_{\infty,1}, \cdots, \alpha_{\infty,q}\}$ such that $\sum_{k=1}^{l_n} \alpha^2_\infty(n,k) = 1$. This is the idea behind Theorem 3.2. However, when we do not assume rational independence of the coefficients, we need to use the assumption of the convergence of the cumulants through a Vandermonde argument to proceed.

3.3 Applications: second Wiener chaos

In this section, we apply our main results in a desirable framework when the approximating sequence F_n are elements of the second Wiener chaos of the isonormal process $\mathrm{X} = \{\mathrm{X(h)};\ \mathrm{h} \in \mathfrak{H}\}$ over a separable Hilbert space $\mathfrak{H}$. We refer the reader to [20] Chapter 2 for a detailed discussion on this topic. Recall that the elements in the second Wiener chaos are random variables having the general form $F = I_2(f)$, with $f \in \mathfrak{H}^{\odot 2}$. Notice that, if $f = h \otimes h$, where $h \in \mathfrak{H}$ is such that $\|h\|_{\mathfrak{H}} = 1$, then using the multiplication formula one has $I_2(f) = \mathrm{X(h)}^2 - 1 \overset{\text{law}}{=} \mathrm{N}^2 - 1$, where $N \sim \mathscr{N}(0,1)$. To any kernel $f \in \mathfrak{H}^{\odot 2}$, we associate the following *Hilbert-Schmidt* operator

$$A_f : \mathfrak{H} \mapsto \mathfrak{H}; \quad g \mapsto f \otimes_1 g.$$

We also write $\{\alpha_{f,j}\}_{j\geq 1}$ and $\{e_{f,j}\}_{j\geq 1}$, respectively, to indicate the (not necessarily distinct) eigenvalues of A_f and the corresponding eigenvectors. The next proposition gathers some relevant properties of the elements of the second Wiener chaos associated to X.

Proposition 3.1 (See Section 2.7.4 in [20] and Lemma 3.1 in [3])**.** *Let $F = I_2(f)$, $f \in \mathfrak{H}^{\odot 2}$, be a generic element of the second Wiener chaos of* X*, and write $\{\alpha_{f,k}\}_{k\geq 1}$ for the set of the eigenvalues of the associated Hilbert-Schmidt operator A_f.*

1. *The following equality holds: $F = \sum_{k\geq 1} \alpha_{f,k}(N_k^2 - 1)$, where $\{N_k\}_{k\geq 1}$ is a sequence of i.i.d. $\mathscr{N}(0,1)$ random variables that are elements of the isonormal process* X*, and the series converges in L^2 and almost surely.*
2. *For any $r \geq 2$,*
$$\kappa_r(F) = 2^{r-1}(r-1)! \sum_{k\geq 1} \alpha_{f,k}^r.$$
3. *For polynomial Q as in* (3.5) *we have $\Delta(F) = \sum_{k\geq 1} Q(\alpha_{f,k})$. In particular $\Delta(F_\infty) = 0$.*

The next corollary is a direct application of our main findings, namely Theorem 3.1 and Theorem 3.2 and provides quantitative bounds for the main results in [24, 3].

Corollary 3.1. *Assume that the normalized sequence $F_n = \sum_{k\geq 1} \alpha_{n,k}(N_k^2 - 1)$ belongs to the second Wiener chaos associated to the isonormal process* X*, and the target random variable F_∞ as in* (3.2)*. Then there exists a constant $C > 0$ depending only on target random variable F_∞ (and hence independent of n) such that*

(a)
$$d_{\mathrm{W}_2}(F_n, F_\infty) \leq C \left(\sqrt{\Delta(F_n)} + \sum_{r=2}^{q+1} |\kappa_r(F_n) - \kappa_r(F_\infty)| \right).$$

(b) if moreover $\dim_{\mathbb{Q}}$ span$\{\alpha_{\infty,1}^2, \cdots, \alpha_{\infty,q}^2\} = q$, then $d_{\mathrm{W}_2}(F_n, F_\infty) \leq C\sqrt{\Delta(F_n)}$. This implies that the sole convergence $\Delta(F_n) \to \Delta(F_\infty) = 0$ is sufficient for convergence in distribution towards the target random variable F_∞.

Remark 3.5. The upper bound in Corollary 3.1, part (a) requires the separate convergences of the first $q+1$ cumulants for the convergence in distribution towards the target random variable F_∞ as soon as $\dim_{\mathbb{Q}} \operatorname{span}\{\alpha_{\infty,1}^2, \cdots, \alpha_{\infty,q}^2\} < q$. This is in very consistent with a quantitative result in [12]. In fact, when $q = 2$ and $\alpha_{\infty,1} = -\alpha_{\infty,2} = 1/2$, then the target random variable F_∞ $(= N_1 \times N_2$, where $N_1, N_2 \sim \mathscr{N}(0,1)$ are independent and equality holds in law) belongs to the class of *Variance–Gamma* distributions $VG_c(r, \theta, \sigma)$ with parameters $r = \sigma = 1$ and $\theta = 0$. Then, [12, Theorem 5.10, part (a)] reads

$$d_{W_1}(F_n, F_\infty) \leq C \sqrt{\Delta(F_n) + 1/4\,\kappa_3^2(F_n)}. \tag{3.10}$$

Therefore, for the convergence in distribution of the sequence F_n towards the target random variable F_∞ in addition to convergence $\Delta(F_n) \to \Delta(F_\infty) = 0$ one needs also the convergence of the third cumulant $\kappa_3(F_n) \to \kappa_3(F_\infty) = 0$. Also note that in this case we have $\dim_{\mathbb{Q}} \operatorname{span}\{\alpha^2_{\infty,1}, \alpha^2_{\infty,2}\} = 1 < q = 2$.

Example 3.1. Again assume that $q = 2$ and $\alpha_{\infty,1} = -\alpha_{\infty,2} = 1/2$. Consider the fixed sequence

$$F_n = F = \alpha_{\infty,1}(N_1^2 - 1) - \alpha_{\infty,2}(N_2^2 - 1) \qquad n \geq 1.$$

Then $\kappa_{2r}(F_n) = \kappa_{2r}(F_\infty)$ for all $r \geq 1$, in particular $\kappa_2(F_n) = \kappa_2(F_\infty) = 1$, and $\Delta(F_n) = \Delta(F_\infty) = 0$. However, it is easy to see that the sequence F_n does not converges in distribution towards the target random variable F_∞, because $2 = \kappa_3(F_n) \nrightarrow \kappa_3(F_\infty) = 0$. Again, we would like to stress that in this case we have $\dim_{\mathbb{Q}} \operatorname{span}\{\alpha^2_{\infty,1}, \alpha^2_{\infty,2}\} = 1 < q = 2$. Therefore the requirement of separate convergences of the first $q+1$ cumulants is essential in Theorem 3.1 as soon as $\dim_{\mathbb{Q}} \operatorname{span}\{\alpha^2_{\infty,1}, \cdots, \alpha^2_{\infty,q}\} < q$.

Example 3.2. [Bai–Taqqu Theorem, 2015] We conclude this section with a more ambitious example, providing rates of convergence in a recent result given by [5, Theorem 2.4]. We stress that many more examples and situations could be tackled by our method. Let Z_{γ_1,γ_2} be the random variable defined by:

$$Z_{\gamma_1,\gamma_2} = \int_{\mathbb{R}^2} \left(\int_0^1 (s - x_1)_+^{\gamma_1} (s - x_2)_+^{\gamma_2} ds \right) dB_{x_1} dB_{x_2},$$

with $\gamma_i \in (-1, -1/2)$ and $\gamma_1 + \gamma_2 > -3/2$. By Proposition 3.1 of [5], we have the following formula for the cumulants of Z_{γ_1,γ_2}:

$$\kappa_m(Z_{\gamma_1,\gamma_2}) = \frac{1}{2}(m-1)!\, A(\gamma_1,\gamma_2)^m C_m(\gamma_1,\gamma_2,1,1)$$

where,

$$\begin{aligned}
A(\gamma_1,\gamma_2) &= [(\gamma_1+\gamma_2+2)(2(\gamma_1+\gamma_2)+3)]^{\frac{1}{2}} \\
&\quad \times [B(\gamma_1+1, -\gamma_1-\gamma_2-1)B(\gamma_2+1, -\gamma_1-\gamma_2-1) \\
&\quad + B(\gamma_1+1, -2\gamma_1-1)B(\gamma_2+1, -2\gamma_2-1)]^{-\frac{1}{2}}, \\
C_m(\gamma_1,\gamma_2,1,1) &= \sum_{\sigma\in\{1,2\}^m} \int_{(0,1)^m} \prod_{j=1}^m [(s_j - s_{j-1})_+^{\gamma_{\sigma_j}+\gamma_{\sigma'_{j-1}}+1} B(\gamma_{\sigma'_{j-1}}+1, -\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1) \\
&\quad + (s_{j-1} - s_j)_+^{\gamma_{\sigma_j}+\gamma_{\sigma'_{j-1}}+1} B(\gamma_{\sigma_j}+1, -\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1)] ds_1 \ldots ds_m, \\
B(\alpha,\beta) &= \int_0^1 u^{\alpha-1}(1-u)^{\beta-1} du.
\end{aligned}$$

Let $\rho \in (0,1)$ and Y_ρ be the random variable defined by:

$$Y_\rho = \frac{a_\rho}{\sqrt{2}}(Z_1^2 - 1) + \frac{b_\rho}{\sqrt{2}}(Z_2^2 - 1),$$

with Z_i independent standard normal random variables and a_ρ and b_ρ defined by:

$$a_\rho = \frac{(\rho+1)^{-1} + (2\sqrt{\rho})^{-1}}{\sqrt{2(\rho+1)^{-2} + (2\rho)^{-1}}},$$
$$b_\rho = \frac{(\rho+1)^{-1} - (2\sqrt{\rho})^{-1}}{\sqrt{2(\rho+1)^{-2} + (2\rho)^{-1}}}.$$

For simplicity, we assume that $\gamma_1 \geq \gamma_2$ and $\gamma_2 = (\gamma_1 + 1/2)/\rho - 1/2$. Then [5, Theorem 2.4] implies that as γ_1 tends to $-1/2$:

$$Z_{\gamma_1,\gamma_2} \overset{\text{law}}{\to} Y_\rho. \tag{3.11}$$

Note that, in this case, γ_2 automatically tends to $-1/2$ as well. To prove the previous result, the authors of [5] prove the following convergence result:

$$\forall m \geq 2,\ \kappa_m(Z_{\gamma_1,\gamma_2}) \to \kappa_m(Y_\rho) = 2^{\frac{m}{2}-1}(a_\rho^m + b_\rho^m)(m-1)!\,.$$

Now, using Corollary 3.1 and applying Lemma 3.2, we can present the following quantative bound for convergence (3.11), namely as γ_1 tends to $-1/2$:

$$\mathrm{d}_{W_2}(Z_{\gamma_1,\gamma_2}, Y_\rho) \leq C_\rho \sqrt{-\gamma_1 - \frac{1}{2}},$$

where C_ρ is some strictly positive constant depending on ρ uniquely.

In order to apply Corollary 3.1 to obtain an explicit rate for convergence (3.11), we need to know at which speed $\kappa_m(Z_{\gamma_1,\gamma_2})$ converges towards $\kappa_m(Y_\rho)$. For this purpose, we have the following lemma:

Lemma 3.2. *Under the above assumptions, for any $m \geq 3$, we have, as γ_1 tends to $-1/2$:*

$$\kappa_m(Z_{\gamma_1,\gamma_2}) = \kappa_m(Y_\rho) + O(-\gamma_1 - \frac{1}{2})$$

Proof. First of all, we note that, as γ_1 tends to $-1/2$:

$$\begin{aligned}
A(\gamma_1,\gamma_2) = {} & [(\gamma_1 + \frac{1}{\rho}(\gamma_1 + \frac{1}{2}) + \frac{3}{2})(2\gamma_1 + \frac{2}{\rho}(\gamma_1 + \frac{1}{2}) + 2)]^{\frac{1}{2}} \\
& \times [B(\gamma_1 + 1, -(1 + \frac{1}{\rho})(\gamma_1 + \frac{1}{2}))B(\frac{1}{\rho}(\gamma_1 + \frac{1}{2}) + \frac{1}{2}, -(1 + \frac{1}{\rho})(\gamma_1 + \frac{1}{2})) \\
& + B(\gamma_1 + 1, -2\gamma_1 - 1)B(\frac{1}{\rho}(\gamma_1 + \frac{1}{2}) + \frac{1}{2}, -\frac{2}{\rho}(\gamma_1 + \frac{1}{2}))]^{-\frac{1}{2}}, \\
\approx {} & \frac{(-\gamma_1 - 1/2)}{\sqrt{(1 + \frac{1}{\rho})^{-2} + (\frac{4}{\rho})^{-1}}} - C_\rho(-3 + 2\gamma + 2\psi(\frac{1}{2}))(\gamma_1 + \frac{1}{2})^2 \\
& + o((-\gamma_1 - 1/2)^2),
\end{aligned}$$

where γ is the Euler constant, $\psi(.)$ is the Digamma function and C_ρ some strictly positive constant depending on ρ uniquely. Note that $-3+2\gamma+2\psi(1/2)<0$. Moreover, we have:

$$C_m(\gamma_1,\gamma_2,1,1) \approx \sum_{\sigma\in\{1,2\}^m}\int_{(0,1)^m}\prod_{j=1}^{m}\Big\{\mathbb{I}_{s_j>s_{j-1}}\Big[(-\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1)^{-1}-\log(s_j-s_{j-1})+(-\gamma-\psi(\frac{1}{2}))$$
$$+o(1)\Big]$$
$$+\mathbb{I}_{s_j<s_{j-1}}\Big[(-\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1)^{-1}-\log(s_{j-1}-s_j)+(-\gamma-\psi(\frac{1}{2}))+o(1)\Big]\Big\}ds_1...ds_m \quad (3.12)$$

$$\approx \sum_{\sigma\in\{1,2\}^m}\int_{(0,1)^m}\prod_{j=1}^{m}\Big[(-\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1)^{-1}+\mathbb{I}_{s_j>s_{j-1}}\log((s_j-s_{j-1})^{-1})$$
$$+\mathbb{I}_{s_j<s_{j-1}}\log((s_{j-1}-s_j)^{-1})+(-\gamma-\psi(\frac{1}{2}))+o(1)\Big]ds_1...ds_m. \quad (3.13)$$

Note that $-\gamma-\psi(\frac{1}{2})>0$. The diverging terms in $C_m(\gamma_1,\gamma_2,1,1)$ are $B(\gamma_{\sigma'_{j-1}}+1,-\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1)$ and $B(\gamma_{\sigma_j}+1,-\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1)$. At σ and j fixed, the only possible values are:

$$B(\gamma_1+1,-\gamma_1-\gamma_2-1)=B(\gamma_1+1,-(\gamma_1+\frac{1}{2})(1+\frac{1}{\rho})),$$
$$\approx -\frac{1}{(1+\frac{1}{\rho})(\gamma_1+\frac{1}{2})}+(-\gamma-\psi(\frac{1}{2}))+o(1),$$
$$B(\gamma_2+1,-\gamma_1-\gamma_2-1)=B(\frac{1}{\rho}(\gamma_1+\frac{1}{2})+\frac{1}{2},-(\gamma_1+\frac{1}{2})(1+\frac{1}{\rho})),$$
$$\approx -\frac{1}{(1+\frac{1}{\rho})(\gamma_1+\frac{1}{2})}+(-\gamma-\psi(\frac{1}{2}))+o(1),$$
$$B(\gamma_1+1,-2\gamma_1-1)\approx -\frac{1}{2(\gamma_1+\frac{1}{2})}+(-\gamma-\psi(\frac{1}{2}))+o(1),$$
$$B(\gamma_2+1,-2\gamma_2-1)=B(\frac{1}{\rho}(\gamma_1+\frac{1}{2})+\frac{1}{2},-\frac{2}{\rho}(\gamma_1+\frac{1}{2})),$$
$$\approx -\frac{\rho}{2(\gamma_1+\frac{1}{2})}+(-\gamma-\psi(\frac{1}{2}))+o(1).$$

Moreover, we have, for j fixed:

$$(s_j-s_{j-1})_+^{\gamma_{\sigma_j}+\gamma_{\sigma'_{j-1}}+1}=\mathbb{I}_{s_j>s_{j-1}}(s_j-s_{j-1})^{\gamma_{\sigma_j}+\gamma_{\sigma'_{j-1}}+1}$$
$$\approx \mathbb{I}_{s_j>s_{j-1}}[1+\log(s_j-s_{j-1})(\gamma_{\sigma_j}+\gamma_{\sigma'_{j-1}}+1)+o((\gamma_{\sigma_j}+\gamma_{\sigma'_{j-1}}+1))].$$

Developing the product in the right hand side of (3.12), we obtain:

$$
\begin{aligned}
C_m(\gamma_1,\gamma_2,1,1) \approx & \sum_{\sigma\in\{1,2\}^m}\prod_{j=1}^{m}(-\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1)^{-1} \\
& + (-\gamma-\psi(\frac{1}{2})) \sum_{\sigma\in\{1,2\}^m}\sum_{j=1}^{m}\prod_{k=1,\ k\neq j}^{m}(-\gamma_{\sigma_k}-\gamma_{\sigma'_{k-1}}-1)^{-1} \\
& + \sum_{\sigma\in\{1,2\}^m}\sum_{j=1}^{m}\prod_{k=1,\ k\neq j}^{m}(-\gamma_{\sigma_k}-\gamma_{\sigma'_{k-1}}-1)^{-1}\int_{(0,1)^m}\Big[\mathbb{I}_{s_j>s_{j-1}}\log((s_j-s_{j-1})^{-1}) \\
& + \mathbb{I}_{s_j<s_{j-1}}\log((s_{j-1}-s_j)^{-1})\Big]ds_1...ds_m + o((-\gamma_1-\frac{1}{2})^{-m+1}) \\
\approx & \sum_{\sigma\in\{1,2\}^m}\prod_{j=1}^{m}(-\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1)^{-1} \\
& + (-\gamma-\psi(\frac{1}{2})+\frac{3}{2}) \sum_{\sigma\in\{1,2\}^m}\sum_{j=1}^{m}\prod_{k=1,\ k\neq j}^{m}(-\gamma_{\sigma_k}-\gamma_{\sigma'_{k-1}}-1)^{-1} \\
& + o((-\gamma_1-\frac{1}{2})^{-m+1})
\end{aligned}
$$

This leads to the following asymptotic for the cumulants of Z_{γ_1,γ_2},

$$
\begin{aligned}
\kappa_m(Z_{\gamma_1,\gamma_2}) \approx & \frac{(m-1)!}{2}\Bigg[\frac{(-\gamma_1-1/2)}{\sqrt{(1+\frac{1}{\rho})^{-2}+(\frac{4}{\rho})^{-1}}} - C_\rho(-3+2\gamma+2\psi(\frac{1}{2}))(\gamma_1+\frac{1}{2})^2 \\
& + o((-\gamma_1-1/2)^2)\Bigg]^m\Bigg[\sum_{\sigma\in\{1,2\}^m}\prod_{j=1}^{m}(-\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1)^{-1} \\
& + (-\gamma-\psi(\frac{1}{2})+\frac{3}{2}) \sum_{\sigma\in\{1,2\}^m}\sum_{j=1}^{m}\prod_{k=1,\ k\neq j}^{m}(-\gamma_{\sigma_k}-\gamma_{\sigma'_{k-1}}-1)^{-1} \\
& + o((-\gamma_1-\frac{1}{2})^{-m+1})\Bigg], \\
\approx & \frac{(m-1)!}{2}\frac{(-\gamma_1-1/2)^m}{\left(\sqrt{(1+\frac{1}{\rho})^{-2}+(\frac{4}{\rho})^{-1}}\right)^m}\sum_{\sigma\in\{1,2\}^m}\prod_{j=1}^{m}(-\gamma_{\sigma_j}-\gamma_{\sigma'_{j-1}}-1)^{-1} \\
& + O((-\gamma_1-\frac{1}{2})) \\
\approx & \kappa_m(Y_\rho) + O((-\gamma_1-\frac{1}{2})),
\end{aligned}
$$

where we have used similar computations as in the proof of Theorem 2.4 of [5] for the last equality. □

3.4 Proof of Theorem 3.1

In what follows, we will need the following useful lemma.

Lemma 3.3. *For the vector* $\boldsymbol{\alpha_\infty} = (\alpha_{\infty,1}, \cdots, \alpha_{\infty,q}) \in \mathbb{R}^q$ *where* $\alpha_{\infty,i}$ *are non-zero and distinct, we denote*

$$d(x, \boldsymbol{\alpha_\infty}) := \min_{i=1,\cdots,q} |x - \alpha_{\infty,i}|, \qquad \forall\, x \in \mathbb{R}.$$

Then, there exists a constant M *such that*

$$d(x, \boldsymbol{\alpha_\infty})^2 \leq M \prod_{i=1}^{q} (x - \alpha_{\infty,i})^2.$$

Proof. Consider the function $f : \mathbb{R} - \{\alpha_{\infty,1}, \cdots, \alpha_{\infty,q}\} \to \mathbb{R}$ given by

$$f(x) := \frac{\prod_{i=1}^{q} (x - \alpha_{\infty,i})^2}{d(x, \boldsymbol{\alpha_\infty})^2}.$$

Then, obviously f is a continuous function on $\mathbb{R} - \{\alpha_{\infty,1}, \cdots, \alpha_{\infty,q}\}$ and can be extended to a continuous function on whole real line $\mathbb{R}$ by setting $f(\alpha_{\infty,i}) := \prod_{j \neq i} (\alpha_{\infty,j} - \alpha_{\infty,i})^2 \neq 0$ at each point $\alpha_{\infty,i}$ for $i = 1, \cdots, q$. On the other hand, note that we have $f(x) \to \infty$ as $|x|$ tends to infinity. Hence, f is bounded from below by a positive constant, say M. □

We split the proofs of Theorems 3.1 and 3.2 in several steps. Throughout, C stands for a generic constant that is independent of n but may differ from line to line.

Step 1: (Re-ordering the coefficients) Under the second moment conditions (3.3), we know that for any fixed $n \geq 1$, we have $\lim_{k \to \infty} \alpha_{n,k} = 0$. Therefore, $\max\{|\alpha_{n,k}| : k \geq 1\}$ is attained in, at least one value, say $|\alpha_{n,k_1}|$. Similarly, $\max\{|\alpha_{n,k}| : k \neq k_1\}$ is attained in some value $|\alpha_{n,k_2}|$. We repeat this procedure by induction and we may build a decreasing sequence $|\alpha_{n,k_1}| \geq |\alpha_{n,k_2}| \geq \cdots |\alpha_{n,k_p}| \geq \cdots$ such that for all $i \geq 1$, we have

$$|\alpha_{n,k_i}| = \max\{|\alpha_{n,k}| : k \neq k_1, k_2, \cdots, k_{i-1}\}.$$

Also, for all $n, p \geq 1$ we have $1 \geq \sum_{i=1}^{p} \alpha_{n,k_i}^2 \geq \sum_{i=1}^{p} \alpha_{n,k}^2 \to 1$ as p tends to infinity. Therefore

$$F_n \overset{\text{law}}{=} \sum_{i=1}^{\infty} \alpha_{n,k_i} W_i, \qquad \forall\, n \geq 1. \tag{3.13}$$

To emphasize the maximality property of α_{n,k_i}, we denote α_{n,k_i} by $\alpha_{\max}(n,i)$, and in the rest of the proof we assume that for each $n \geq 1$, F_n is given by the right hand side of (3.13).

Step 2: (Bounding the $\alpha_{\max}(n,i)$'s from below) For any $p \geq 1$, we introduce the quantity

$$\Delta_p(F_n) := \sum_{k=p}^{\infty} \alpha_{\max}^2(n,k) \prod_{r=1}^{q} \left(\alpha_{\max}(n,k) - \alpha_{\infty,r}\right)^2. \tag{3.14}$$

Next, we observe that

$$\begin{aligned}\Delta_p(F_n) &= \sum_{k=p}^{\infty} Q(\alpha_{\max}(n,k)) = \sum_{r=2}^{2q+2} \Theta_r \sum_{k=p}^{\infty} \alpha_{\max}^r(n,k) \\ &= \Theta_2 \Big(1 - \sum_{k=1}^{p-1} \alpha_{\max}^2(n,k)\Big) + \sum_{r=3}^{2q+2} \Theta_r \sum_{k=p}^{\infty} \alpha_{\max}^r(n,k).\end{aligned} \tag{3.15}$$

Besides, for any $r \geq 3$ and the maximality property of the coefficients $\alpha_{\max}(n,k)$ together with the normalization assumption (3.3) we have the following estimate (which is valid for all $r \geq 3$):

$$\left|\sum_{k=p}^{\infty} \alpha_{\max}^{r}(n,k)\right| \leq \left|\alpha_{\max}^{r-2}(n,p)\right| \times \left(1 - \sum_{k=1}^{p-1} \alpha_{\max}^{2}(n,k)\right).$$

Also note that, since $|\alpha_{\max}(n,p)| \leq 1$, we always have $|\alpha_{\max}(n,p)|^{r-2} \leq |\alpha_{\max}(n,p)|$ (still for all $r \geq 3$). We may deduce

$$\left|\Delta_p(F_n) - \Theta_2\Big(1 - \sum_{k=1}^{p-1} \alpha_{\max}^{2}(n,k)\Big)\right| \leq |\alpha_{\max}(n,p)|\Big(1 - \sum_{k=1}^{p-1} \alpha_{\max}^{2}(n,k)\Big) \sum_{r=3}^{2q+2} |\Theta_r|,$$

leading in turn to the lower bound

$$\left|\alpha_{\max}(n,p)\right| \geq \frac{|\Theta_2|}{\sum_{r=3}^{2q+2}|\Theta_r|} - \frac{\Delta_p(F_n)}{\Big(1 - \sum_{k=1}^{p-1} \alpha_{\max}^{2}(n,k)\Big) \times \sum_{r=3}^{2q+2}|\Theta_r|}. \tag{3.16}$$

Note that in the right hand side of (3.16), the first summand depends only on the limiting law. In order to deal with the second summand, we need control on $1 - \sum_{k=1}^{p-1} \alpha_{\max}^{2}(n,k)$. To this end we introduce the following useful quantities:

$$\vartheta = \min\Big\{1 - \sum_{i=1}^{q} n_i \alpha_{\infty,i}^{2} \,\Big|\, (n_1,\cdots,n_q) \in \mathbb{N}_0^q, \quad \text{and} \quad 1 - \sum_{i=1}^{q} n_i \alpha_{\infty,i}^{2} > 0\Big\}. \tag{3.17}$$

$$\varkappa = \max\Big\{1 - \sum_{i=1}^{q} n_i \alpha_{\infty,i}^{2} \,\Big|\, (n_1,\cdots,n_q) \in \mathbb{N}_0^q, \quad \text{and} \quad 1 - \sum_{i=1}^{q} n_i \alpha_{\infty,i}^{2} < 0\Big\}. \tag{3.18}$$

Note that, for any vector $(n_1,\cdots,n_q) \in \mathbb{N}_0^q$ such that $1 - \sum_{i=1}^{q} n_i \alpha_{\infty,i}^{2} > 0$ we have

$$\vartheta < 1 - \sum_{i=1}^{q} n_i \alpha_{\infty,i}^{2} \leq 1 - \alpha_{\min}^{2}(\infty)\left(\sum_{i=1}^{q} n_i\right)$$

with $\alpha_{\min}^{2}(\infty) = \min\{\alpha_{\infty,i}^{2} : i = 1,\cdots,q\}$. This leads to the important upper bound estimate

$$\sum_{i=1}^{q} n_i \leq \frac{1-\vartheta}{\alpha_{\min}^{2}(\infty)}, \tag{3.19}$$

which is finite because our assumption on the coefficients of the target random variable F_∞ implies that $\alpha_{\min}^{2}(\infty) \neq 0$. Finally we set

$$L := \lfloor \frac{1-\vartheta}{\alpha_{\min}^{2}(\infty)} \rfloor \text{ and } \varrho := \min\{\vartheta, |\varkappa|\}.$$

Step 3: (Induction procedure) We now aim to use Step 2 to control the distance between the eigenvalues $\alpha_{\max}(n,p)$ and the eigenvalues $\alpha_{\infty,p}$ of the target random variable F_∞ in terms of $\Delta(F_n)$. Taking into account the assumption $\Delta(F_n) \to 0$ as n tends to infinity, one can find an $N_0 \in \mathbb{N}$ such that

$$\Delta(F_n) \leq \min\{\frac{|\Theta_2|}{2}, \frac{\varrho|\Theta_2|}{4}\} \text{ and } \sqrt{L+1} \times \sqrt{M \times \left(\frac{2\sum_{r=3}^{2q+2}|\Theta_r|}{|\Theta_2|}\right)^2 \times \Delta(F_n)} < \frac{\varrho}{2}. \tag{3.20}$$

Next, for any $n \geq N_0$ we define

$$\mathscr{A}_n := \Big\{1 \leq p \leq L : \text{ for any } 1 \leq l \leq p \text{ we have } |\alpha_{\max}(n,l)| \geq \frac{|\Theta_2|}{2\sum_{r=3}^{2q+2}|\Theta_r|}\Big\}.$$

The collection $\mathscr{A}_n$ is not empty. In fact, using the estimate (3.16) with $p = 1$, one can immediately get

$$\Big|\alpha_{\max}(n,1)\Big| \geq \frac{|\Theta_2|}{\sum_{r=3}^{2q+2}|\Theta_r|} - \frac{\Delta_1(F_n)}{\sum_{r=3}^{2q+2}|\Theta_r|} \geq \frac{|\Theta_2|}{2\sum_{r=3}^{2q+2}|\Theta_r|} \tag{3.21}$$

because, for $n \geq N_0$, we know that $\Delta_1(F_n) = \Delta(F_n) \leq \frac{|\Theta_2|}{2}$ thanks to the first inequality in (3.20). Note that for any $p \geq 1$, we have $\Delta_p(F_n) \leq \Delta(F_n) \to 0$ as n tends to infinity. On the other hand, $\mathscr{A}_n$ is set bounded by L, and therefore has a maximal element which we denote by L_n. By the definitions of $\Delta(F_n)$ and of the set $\mathscr{A}_n$ we infer that

$$\sum_{k=1}^{L_n}\prod_{r=1}^{q}(\alpha_{\max}(n,k) - \alpha_{\infty,r})^2 \leq \left(\frac{2\sum_{r=3}^{2q+2}|\Theta_r|}{|\Theta_2|}\right)^2 \Delta(F_n).$$

Then, in virtue of Lemma 3.3 this reads

$$\sum_{k=1}^{L_n} d(\alpha_{\max}(n,k), \boldsymbol{\alpha_\infty})^2 \leq M \times \left(\frac{2\sum_{r=3}^{2q+2}|\Theta_r|}{|\Theta_2|}\right)^2 \Delta(F_n). \tag{3.22}$$

On the other hand, for any $1 \leq k \leq L_n$ there exists some $\alpha_\infty(n,k) \in \{\alpha_{\infty,1}, \cdots, \alpha_{\infty,q}\}$ realizing the minimum in definition of $d(\alpha_{\max}(n,k), \boldsymbol{\alpha_\infty})$. Here, one has to note that the coefficients $\alpha_\infty(n,k)$ in general can be repeated. So, we can rewrite (3.22) as

$$\sum_{k=1}^{L_n} |\alpha_{\max}(n,k) - \alpha_\infty(n,k)|^2 \leq M \times \left(\frac{2\sum_{r=3}^{2q+2}|\Theta_r|}{|\Theta_2|}\right)^2 \Delta(F_n), \tag{3.23}$$

which gives us part of the control we seek. It still remains to show that the remainder is well-behaved. To this end we will show that $\forall n \geq N_0$ there exists $l_n \in \{1, \ldots, L+1\}$ such that

$$\sum_{k=1}^{\ell_n} |\alpha_{\max}(n,k) - \alpha_\infty(n,k)|^2 \leq M \times \left(\frac{2\sum_{r=3}^{2q+2}|\Theta_r|}{|\Theta_2|}\right)^2 \Delta(F_n) \tag{3.24}$$

and

$$\sum_{k=\ell_n+1}^{\infty} \alpha_{\max}^2(n,k) \leq C\sqrt{\Delta(F_n)}. \tag{3.25}$$

First, taking into account the estimate (3.23), one can infer that

$$
\begin{aligned}
\Big|(1-\sum_{k=1}^{L_n}\alpha^2_{\max}(n,k)) \; &-(1-\sum_{k=1}^{L_n}\alpha^2_{\infty}(n,k))\Big| \\
&= \Big|\sum_{k=1}^{L_n}\big(\alpha^2_{\infty}(n,k)-\alpha^2_{\max}(n,k)\big)\Big| \\
&\le 2\sum_{k=1}^{L_n}|\alpha_{\max}(n,k)-\alpha_{\infty}(n,k)| \\
&\le \sqrt{L_n}\times\sqrt{M\times\left(\frac{2\sum_{r=3}^{2q+2}|\Theta_r|}{|\Theta_2|}\right)^2\times\Delta(F_n)} \\
&\le \sqrt{L}\times\sqrt{M\times\left(\frac{2\sum_{r=3}^{2q+2}|\Theta_r|}{|\Theta_2|}\right)^2\times\Delta(F_n)}. \qquad (3.26)
\end{aligned}
$$

In order to conclude we now seek for an index l_n such that

$$\sum_{k=1}^{\ell_n}\alpha^2_{\infty}(n,k)=1. \qquad (3.27)$$

Given $n\ge N_0$ we have three possibilities.

- If $1-\sum_{k=1}^{L_n}\alpha^2_{\infty}(n,k)=0$ then we can take $\ell_n=L_n$ and, by (3.26) and (3.23), we are done.
- If $1-\sum_{k=1}^{L_n}\alpha^2_{\infty}(n,k)>0$ then $1-\sum_{k=1}^{L+1}\alpha^2_{\infty}(n,k)=0$ and we can take $\ell_n=L+1$. Indeed we necessarily have $1-\sum_{k=1}^{L_n}\alpha^2_{\infty}(n,k)>\varrho$ by definition of ϱ. Using the second inequality in (3.20) as well as the estimate given in (3.26) one can infer that

  $$1-\sum_{k=1}^{L_n}\alpha^2_{\max}(n,k)\ge\frac{\varrho}{2}. \qquad (3.28)$$

  Now, using estimate (3.16) with $p=L_n+1$ together with the first estimate in (3.20) we obtain that

  $$|\alpha_{\max}(n,L_n+1)|\ge\frac{|\Theta_2|}{2\sum_{r=3}^{2q+2}|\Theta_r|}. \qquad (3.29)$$

  If L_n were strictly less than L, then it would contradict the fact that L_n is the maximal element of the set $\mathscr{A}_n$, and therefore $L_n=L$. Now, following exactly the same lines as in the beginning of this step and using (3.29) one can infer that

  (i) $\sum_{k=1}^{L+1}|\alpha_{\max}(n,k)-\alpha_{\infty}(n,k)|^2\le M\times\left(\frac{2\sum_{r=3}^{2q+2}|\Theta_r|}{|\Theta_2|}\right)^2\times\Delta(F_n)$.

  (ii) $\Big|(1-\sum_{k=1}^{L+1}\alpha^2_{\max}(n,k))-(1-\sum_{k=1}^{L+1}\alpha^2_{\infty}(n,k))\Big|<\frac{\varrho}{2}$.

  Now, we are left to show that $1-\sum_{k=1}^{L+1}\alpha^2_{\infty}(n,k)=0$. First, note that according to definition of L, if $1-\sum_{k=1}^{L+1}\alpha^2_{\infty}(n,k)\ne0$, then we have to have that $1-\sum_{k=1}^{L+1}\alpha^2_{\infty}(n,k)<0$. Now, again using definition of $\varkappa$ and ϱ, this implies that $1 - \sum_{k=1}^{L+1} \alpha_\infty^2(n,k) \le -\varrho$. Now, taking into account (ii) and the fact that $1 - \sum_{k=1}^{L+1} \alpha_{\max}^2(n,k) \ge 0$, we arrive to

$$-\frac{\varrho}{2} \le 1 - \sum_{k=1}^{L+1} \alpha_\infty^2(n,k) \le -\varrho.$$

That is obviously a contradiction and therefore $1 - \sum_{k=1}^{L+1} \alpha_\infty^2(n,k) = 0$. Finally, employing the same estimates as in (3.26) we get

$$\begin{aligned}\sum_{k \ge \ell_n+1} \alpha_{\max}^2(n,k) &= \Big|1 - \sum_{k \le \ell_n} \alpha_{\max}^2(n,k)\Big| \\ &= \Big|(1 - \sum_{k \le \ell_n} \alpha_{\max}^2(n,k)) - (1 - \sum_{k \le \ell_n} \alpha_\infty^2(n,k))\Big| \\ &\le C\Big(\sum_{k \le \ell_n} |\alpha_{\max}(n,k) - \alpha_\infty(n,k)|^2\Big)^{\frac{1}{2}} \\ &\le C\sqrt{\Delta(F_n)}.\end{aligned}$$

- The case $1 - \sum_{k=1}^{L_n} \alpha_\infty^2(n,k) < 0$ can be also discussed in the same way, this time leading to some $\ell_n < L$.

The square root in (3.25) is clearly not sharp and we need to remove it. To do this, observe that an obvious consequence of (3.25) is

$$\alpha_{\max}^2(n,k) \le C\sqrt{\Delta(F_n)} \qquad \forall\, k \ge \ell_n + 1.$$

Thus, if $\Delta(F_n)$ is small enough in the sense that $C\sqrt{\Delta(F_n)} \le U_\infty < \min_{1 \le i \le q+1} \alpha_{\infty,i}$, since $\alpha_{\infty,i} \neq 0$ by assumption, it follows trivially that we may find a universal positive constant C such that for any $n \in \mathbb{N}$ and for any $k \ge \ell_n + 1$ it holds that

$$\prod_{i=1}^{q} (\alpha_{\max}(n,k) - \alpha_{\infty,i})^2 > C.$$

Hence

$$\Delta(F_n) \ge \Delta_{\ell_n+1}(F_n) \ge C \sum_{k=\ell_n+1}^{\infty} \alpha_{\max}^2(n,k),$$

leading to the final estimate

$$\sum_{k=\ell_n+1}^{\infty} \alpha_{\max}^2(n,k) \le \frac{1}{C}\,\Delta(F_n). \tag{3.30}$$

Step 4: (An algebraic argument) We have showed that $\forall n \ge N_0$ there exists $l_n \in \{1, \ldots, L+1\}$ such that

$$\sum_{k=1}^{\ell_n} \alpha_\infty^2(n,k) = 1 \text{ and } \sum_{k=1}^{\ell_n} |\alpha_{\max}(n,k) - \alpha_\infty(n,k)|^2 + \sum_{k=\ell_n+1}^{\infty} \alpha_{\max}^2(n,k) \le C\Delta(F_n). \tag{3.31}$$

For every n, let $\nu(n,k), k = 1, \ldots, q$ stand for the multiplicity of the coefficient $\alpha_{\infty,k}$ realizing the minimum in the definition of the d-distance (the numbers $\nu(n,k)$ can, a priori, be equal or take value zero). Clearly $\sum_{k=1}^{q} \nu(n,k)\alpha_{\infty,k}^2 = 1$. In order to reap the conclusion in 2-Wasserstein distance, we are only left to show that, in fact, we have that $\nu(n,k) = 1$ for all $1 \le k \le q$ and $\ell_n = q$.

Proof of Theorem 3.2. In this case $\dim_{\mathbb{Q}} \operatorname{span}\left\{\alpha_{\infty,1}^2, \cdots, \alpha_{\infty,q}^2\right\} = q$. Then, condition $\sum_{k=1}^{q} \nu(n,k)\alpha_{\infty,k}^2 = 1$ necessarily implies $\nu(n,k) = 1$ for all $k = 1, \ldots, q$ and thus $\ell_n = q$ (recall Remark 3.2). □

Proof of Theorem 3.1. For any $r = 2, \cdots, q+1$, one can write

$$\kappa_r(F_n) = \kappa_r(W_1) \sum_{k=1}^{\ell_n} \alpha_{\max}^r(n,k) + \kappa_r(W_1) \sum_{k=\ell_n+1}^{\infty} \alpha_{\max}^r(n,k).$$

Therefore, according to (3.30), we obtain

$$\left| \kappa_r(F_n) - \kappa_r(W_1) \sum_{k=1}^{\ell_n} \alpha_{\max}^r(n,k) \right| \leq C\, \Delta(F_n). \tag{3.32}$$

Moreover, from the Cauchy-Schwarz inequality and (3.31), we have

$$\left| \sum_{k=1}^{\ell_n} \alpha_{\max}^r(n,k) - \sum_{k=1}^{q} \nu(n,k)\alpha_{\infty,k}^r \right| \leq C \sqrt{\Delta(F_n)} \tag{3.33}$$

for any $r = 2, \cdots, q+1$. Combining estimates (3.32) and (3.33) together, we arrive to

$$\left| \kappa_r(F_n) - \kappa_r(W_1) \sum_{k=1}^{q} \nu(n,k)\alpha_{\infty,k}^r \right| \leq C \sqrt{\Delta(F_n)}. \tag{3.34}$$

Now, we introduce the following $q \times q$ so-called Vandermonde matrix which is invertible because we assumed that the coefficients $\alpha_{\infty,k}$ are pairwise distinct

$$\mathbb{M} = \begin{bmatrix} \alpha_{\infty,1}^2 & \alpha_{\infty,2}^2 & \cdots & \alpha_{\infty,q}^2 \\ \cdots & \cdots & & \\ \alpha_{\infty,1}^{q+1} & \alpha_{\infty,2}^{q+1} & \cdots & \alpha_{\infty,q}^{q+1} \end{bmatrix}.$$

Set $\mathbf{V}_n := (\nu(n,1), \cdots, \nu(n,q))^t$ (t stands for transposition) and consider the vector

$$\Xi = \left(\frac{\kappa_2(F_\infty)}{\kappa_2(W_1)}, \ldots, \frac{\kappa_{q+1}(F_\infty)}{\kappa_{q+1}(W_1)} \right)^t.$$

Note that by our assumption $\kappa_r(W_1) \neq 0$ for all $2 \leq r \leq q+1$. Now, inequality (3.34) entails that

$$\|\mathbb{M}\mathbf{V}_n - \Xi\|_\infty \leq C \left(\sqrt{\Delta(F_n)} + \sum_{r=2}^{q+1} |\kappa_r(F_\infty) - \kappa_r(F_n)| \right), \tag{3.35}$$

which, by inversion, implies that

$$\|\mathbf{V}_n - \mathbb{M}^{-1}\Xi\|_\infty \leq C \left(\sqrt{\Delta(F_n)} + \sum_{r=2}^{q+1} |\kappa_r(F_\infty) - \kappa_r(F_n)| \right). \tag{3.36}$$

If the right hand side of the estimate (3.36) is strictly less than 1, say $\frac{1}{3}$, then because $\mathbf{V}_n$ is a vector of integer numbers, we necessarily have where $\lfloor \cdot \rfloor$ denotes the standard integer part. In order to conclude the proof, it remains to show that $\lfloor \mathbb{M}^{-1}\Xi \rfloor$ is the vector $(1, \cdots, 1)^t$ and $\ell_n = q$. Note that $\lfloor \mathbb{M}^{-1}\Xi \rfloor$ does not depend on the approximating sequence F_n but only on the target random variable F_∞. So we might place ourselves in the obvious situation where the approximating sequence F_n is nothing else than the target itself. In this case $\mathbf{V}_n = (1, \cdots, 1)^t$ and necessarily

$$\lfloor \mathbb{M}^{-1}\Xi \rfloor = (1, \cdots, 1)^t.$$

□

Remark 3.6. Here, we would like to strongly highlight that in order to show that all the multiplicity numbers $\nu(n, k) = 1$ for all $k = 1, \cdots, q$, one needs that the all distances $|\kappa_r(F_n) - \kappa_r(F_\infty)|$ are very small (in the above sense) for all $r = 2, \cdots, q + 1$. However, after doing that taking into account the estimate (3.34), one can immediately observe that the true convergence rate is given by $\sqrt{\Delta(F_n)}$.

Acknowledgements

BA's research is supported by a Welcome Grant from the Université de Liège. YS gratefully acknowledges support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy).

References

[1] Azmoodeh, A., Campese, S., and Poly, G. (2014) Fourth Moment Theorems for Markov Diffusion Generators. *J. Funct. Anal.* 266, No. 4, 2341-2359.

[2] Azmoodeh, E., Malicet, D., Mijoule, G., Poly, G. (2015) Generalization of the Nualart-Peccati criterion. To appear in *Ann. Probab.*

[3] Azmoodeh, E., Peccati, G., and Poly, G. (2014) Convergence towards linear combinations of chi-squared random variables: a Malliavin-based approach. *SÃl'minaire de ProbabilitÃl's* XLVII (Special volume in memory of Marc Yor), 339-367.

[4] Barbour, A. D. (1990) Stein's method for diffusion approximations. *Probab. Theory Related Fields* 84, 297-322.

[5] Bai, S., Taqqu, M. (2015) Behavior of the generalized Rosenblatt process at extreme critical exponent values. To appear in *Ann. Probab.*

[6] Barbour, A. D., Holst, L., and Janson, S. (1992) *Poisson approximation.* Oxford: Clarendon Press.

[7] Chatterjee, S., and Meckes, E. (2008) Multivariate normal approximation using exchangeable pairs. *ALEA* 4, 257-283.

[8] Chen, L.HY, Goldstein, L. and Shao, Q.M. (2010) *Normal approximation by Stein's method.* Springer Science Business Media.

[9] Döbler, C. (2015) Stein's method of exchangeable pairs for the Beta distribution and generalizations. *Electron. J. Probab.* 20, No. 109, 1-34.

[10] Döbler, C., Gaunt, R. E., and Vollmer, S. J. (2015) An iterative technique for bounding derivatives of solutions of Stein equations. Preprint `http://arxiv.org/abs/1510.02623`.

[11] Eden, R., Viquez, J. (2015) Nourdin-Peccati analysis on Wiener and Wiener-Poisson space for general distributions. *Stoch. Proc. Appl.* 125, No. 1, 182-216.

[12] Eichelsbacher, P., Thäle, C. (2014) Malliavin-Stein method for Variance-Gamma approximation on Wiener space. Preprint `http://arxiv.org/abs/1409.5646`.

[13] Gaunt, R. E. (2015) Products of normal, beta and gamma random variables: Stein characterisations and distributional theory. Preprint `http://arxiv.org/abs/1507.07696`.

[14] Gaunt, R. E. (2015) On Stein's method for products of normal random variables and zero bias couplings. Preprint `http://arxiv.org/abs/1309.4344`.

[15] Gaunt, R. E. (2014) Variance-Gamma approximation via Stein's method. *Electron. J. Probab.* 19, No. 38, 1-33.

[16] Gaunt, R. E., Pickett, A., and Reinert, G. (2015) Chi-square approximation by Stein's method with application to Pearson's statistic. Preprint `http://arxiv.org/abs/1507.01707`.

[17] Ley, C., Reinert, G., and Swan, Y. (2014) Approximate computation of expectations: a canonical Stein operator. Preprint `http://arxiv.org/abs/1408.2998`.

[18] Ley, C., Swan, Y. (2013) Local Pinsker inequalities via Stein's discrete density approach. *IEEE Trans. Inf. Theory* 59, No. 9, 5584-5591.

[19] Marinucci, D., Peccati, G., Rossi, M., Wigman, I. (2015) Non-Universality of Nodal Length Distribution for Arithmetic Random Waves. Preprint http://arxiv.org/abs/1508.00353.

[20] Nourdin, I., Peccati, G. (2012) *Normal Approximations Using Malliavin Calculus: from Stein's Method to Universality.* Cambridge Tracts in Mathematics. Cambridge University.

[21] Nourdin, I., Peccati, G. (2009) Noncentral convergence of multiple integrals. *Ann. Probab.* 37, No. 4, 1412-1426.

[22] Nourdin, I., Peccati, G. (2009) Stein's method on Wiener chaos. *Probab. Theory Related Fields.* 145, No. 1-2, 75-118.

[23] Nourdin, I., Peccati, G., and Swan, Y. (2014). Entropy and the fourth moment phenomenon. *J. Funct. Anal.*, 266, No. 5, 3170-3207.

[24] Nourdin, I., Poly, G. (2012) Convergence in law in the second Wiener/Wigner chaos. *Electron. Commun. Probab.* 17, No. 36, 1-12.

[25] Peköz, E., Röllin, A. and Ross, N. (2013) Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Prob.* 23, 1188-1218.

[26] Pickett, A. (2004) *Rates of Convergence of χ^2 approximations via Stein's Method.* DPhil thesis, University of Oxford.

[27] Pike, J., Ren, H. (2014) Stein's method and the Laplace distribution. Preprint http://arxiv.org/abs/1210.5775.

[28] Stein, C. (1972) A bound for the error in the normal approximation to the distribution of a sum of dependent random variables Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, II, 583–602.

[29] Stein, C. (1986) *Approximate computation of expectations.* IMS, Lecture Notes-Monograph Series 7.

[30] Kusuoka, S., Tudor, C. A. (2012) Stein's method for invariant measures of diffusions via Malliavin calculus. *Stoch. Proc. Appl.* 122, No. 4, 1627-1651.

[31] Villani, C. (2009) *Optimal transport. Old and new.* Grundlehren der Mathematischen Wissenschaften 338. Springer.

(B. Arras and Y. Swan) Mathematics department, Université de Liège, Liège, Belgium

(E. Azmoodeh) Department of Mathematics and Statistics, University of Vaasa, Finland

(G. Poly) Institut de Recherche Mathématiques de Rennes, Université de Rennes 1, Rennes, France

E-mail address, B. Arras barras@ulg.ac.be
E-mail address, E. Azmoodeh ehsan.azmoodeh@uva.fi
E-mail address, G. Poly guillaume.poly@univ-rennes1.fr
E-mail address, Y. Swan yswan@ulg.ac.be