



# An Invariance Principle for Stochastic Series II. Non Gaussian Limits

Vlad Bally, Lucia Caramellino

► **To cite this version:**

Vlad Bally, Lucia Caramellino. An Invariance Principle for Stochastic Series II. Non Gaussian Limits. 2016. <hal-01413533>

**HAL Id: hal-01413533**

**<https://hal-upec-upem.archives-ouvertes.fr/hal-01413533>**

Submitted on 9 Dec 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# An Invariance Principle for Stochastic Series II. Non Gaussian Limits

VLAD BALLY\*  
LUCIA CAMELLINO†

July 14, 2016

## Abstract

We study the convergence in total variation distance for series of the form

$$S_N(c, Z) = \sum_{l=1}^N \sum_{i_1 < \dots < i_l} c(i_1, \dots, i_l) Z_{i_1} \dots Z_{i_l}$$

where  $Z_k, k \in \mathbb{N}$  are independent centered random variables with  $\mathbb{E}(Z_k^2) = 1$ . This enters in the framework of the  $U$ -statistics theory which plays a major role in modern statistic. In the case when  $Z_k, k \in \mathbb{N}$  are standard normal,  $S_N(c, Z)$  is an element of the sum of the first  $N$  Wiener chaoses and, starting with the seminal paper of D. Nualart and G. Peccati, the convergence of such functionals to the Gaussian law has been extensively studied. So the interesting point consists in studying invariance principles, that is, to replace Gaussian random variables with random variables with a general law. This has been done in several papers using the Fortet–Mourier distance, the Kolmogorov distance or the total variance distance. In particular, estimates of the total variance distance in terms of the fourth order cumulants has been given in the part I of the present paper. But, as the celebrated Fourth Moment Theorem of Nualart and Peccati shows, such estimates are pertinent to deal with Gaussian limits. In the present paper we study the convergence to general limits which may be non Gaussian, and then the estimates of the error has to be done in terms of the low influence factor only.

**Keywords:** invariance principles, nonlinear Central Limit Theorem, Malliavin calculus.

**2010 MSC:** 60F05, 60H07.

## Contents

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Introduction and main results</b>        | <b>2</b> |
| <b>2</b> | <b>Proofs of the main results</b>           | <b>6</b> |
| 2.1      | Notation and preliminary results . . . . .  | 6        |
| 2.2      | Estimate of the covariance matrix . . . . . | 9        |
| 2.3      | Proof of the main results . . . . .         | 11       |

---

\*Université Paris-Est, LAMA (UMR CNRS, UPEMLV, UPEC), INRIA, F-77454 Marne-la-Vallée, France.  
Email: [bally@univ-mlv.fr](mailto:bally@univ-mlv.fr).

†Dipartimento di Matematica, Università di Roma - Tor Vergata, Via della Ricerca Scientifica 1, I-00133 Roma, Italy. Email: [caramell@mat.uniroma2.it](mailto:caramell@mat.uniroma2.it)

|          |  |           |
|----------|--|-----------|
| <b>3</b> | <b>Examples</b>                                | <b>13</b> |
| 3.1      | Approximation with a chi-squared law . . . . . | 13        |
| 3.2      | An example of quadratic CLT . . . . .          | 14        |
| 3.3      | A variance-type estimator . . . . .            | 15        |
| <b>A</b> | <b>An iterated Hoeffding’s inequality</b>      | <b>19</b> |
| <b>B</b> | <b>Computations around an integral</b>         | <b>22</b> |
|          | <b>References</b>                              | <b>28</b> |

# 1 Introduction and main results

Let us introduce the objects involved in our paper. We consider a sequence of independent random variables  $Z_k, k \in \mathbb{N}$  with  $\mathbb{E}(Z_k) = 0$  and  $\mathbb{E}(Z_k^2) = 1$ . We assume that the law of each of them is locally lower bounded by the Lebesgue measure, that is  $\mathbb{P}(Z_k \in dy) \geq \varepsilon dy$  for  $y \in B(z_k, 2r)$ . More precisely, there exists  $r, \varepsilon > 0$  and  $z_k \in \mathbb{R}$  such that, for every measurable function  $f : \mathbb{R} \rightarrow \mathbb{R}_+$

$$\mathbb{E}(f(Z_k)) \geq \varepsilon \int f(z) 1_{B(0,2r)}(z - z_k) dz. \tag{1.1}$$

All along the paper we will fix some  $r, \varepsilon \in (0, 1)$  and an increasing sequence  $M_p \in (1, \infty), p \in \mathbb{N}$ . These are arbitrary but fixed (without any supplementary mention). We use the notation  $\mathcal{L}((M_p)_{p \in \mathbb{N}}, r, \varepsilon)$  to indicate the sequences of independent random variables  $Z = (Z_k)_{k \in \mathbb{N}}$  with  $\mathbb{E}(Z_k) = 0$  and  $\mathbb{E}(Z_k^2) = 1$  which verifies (1.1) with  $r, \varepsilon$  and such that  $\|Z_k\|_p \leq M_p$  for every  $k, p \in \mathbb{N}$ . Notice that the random variables  $Z_k$  are not identically distributed. However, the fact that we may choose  $(M_p)_{p \in \mathbb{N}}, r, \varepsilon$  to be the same for all of them represents an uniformity property.

We consider a family of coefficients  $c = \{c(\alpha) : \alpha \in \mathbb{N}^m, m \in \mathbb{N}\}$  and for a multi-index  $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$  we denote  $|\alpha| = m$  the length of  $\alpha$ . We also denote  $Z^\alpha = Z_{\alpha_1} \cdots Z_{\alpha_m}$ . We denote by  $\mathcal{C}$  the class of the coefficients  $c$  which are symmetric and null on the diagonals. And we look to stochastic series of the following type:

$$S_N(c, Z) = \sum_{1 \leq |\alpha| \leq N} c(\alpha) Z^\alpha \tag{1.2}$$

This enters in the framework of  $U$ -statistics introduced by Hoeffding [12] and Fisher [11], which play a major role in modern statistics (see for example Lee [15]). Moreover we denote

$$\delta_1(c) = \max_k |c(k)|, \quad \delta_m(c) = \max_k \left( \sum_{|\alpha|=m-1} c^2(\alpha, k) \right)^{1/2} \quad m \geq 2, \quad \bar{\delta}_N(c) = \sum_{m=1}^N \delta_m(c). \tag{1.3}$$

$\bar{\delta}_N(c)$  is the so called “influence factor”:  $\sum_{|\alpha|=m-1} c^2(\alpha, k)$  may be considered as the measure of the action of the particle  $k$  on all the other particles, at level  $m$ . And if  $\bar{\delta}_N(c)$  is small we say that we have “low influence”.

We will also use the following semi-norms

$$|c|_m = \left( \sum_{|\alpha|=m} c^2(\alpha) \right)^{1/2} \quad \text{and} \quad \|c\|_N^2 = \sum_{m=1}^N |c|_m^2 \tag{1.4}$$

We are now able to give our first result:

**Theorem 1.1** *We consider a sequence  $Z^n = (Z_k^n)_{k \in \mathbb{N}} \in \mathcal{L}((M_p)_{p \in \mathbb{N}}, r, \varepsilon)$ . Let  $N \in \mathbb{N}$  be fixed and let  $c_n \in \mathcal{C}$  be a sequence of coefficients such that*

$$\limsup_n \sum_{|\alpha| \leq N} c_n^2(\alpha) < \infty. \quad (1.5)$$

*We assume that they verify the “low influence condition”:*

$$\lim_{n \rightarrow \infty} \bar{\delta}_N(c_n) = 0. \quad (1.6)$$

*We also assume that the following non degeneracy condition holds:*

$$\liminf_{n \rightarrow \infty} \sum_{|\alpha|=N} c_n^2(\alpha) > 0. \quad (1.7)$$

*Let  $\mu$  be a probability measure. Then  $\lim_{n \rightarrow \infty} S_N(c_n, Z) = X$  in law implies (and so is equivalent to) convergence in total variation distance.*

**Remark 1.2** *This is a generalization of the celebrated Prohorov’s Theorem (see [29]) concerning convergence in total variation in the CLT (which corresponds to  $N = 1$ ). And as it is clear from Prohorov’s theorem, the condition (1.1) appears as natural when dealing with convergence in total variation distance (in contrast with convergence in law or in Kolmogorov distance when such a condition is not necessary). A more particular variant of this result has already been obtained recently by Nourdin and Poly in [23].*

**Remark 1.3** *Notice that the non degeneracy condition (1.7) is much stronger than the one in [3] where  $\sum_{|\alpha|=N} c_n^2(\alpha)$  is replaced by  $\sum_{|\alpha| \leq N} c_n^2(\alpha)$ . So here we ask that the higher line of  $S_N$  is non degenerated while in [3] all the coefficients  $c(\alpha)$  in the sum contribute to the non degeneracy condition. But there we also need that the cumulants tend to zero (not only the influence factor) and if this is true then  $\mu$  is a Gaussian probability measure.*

We will now give some (non asymptotic) estimates for the errors involved in the limit in total variation distance. We denote  $\mathbb{N}_*$  the set of the positive integers and given  $N \in \mathbb{N}_*$  we will use the following constants:

$$c_N(r, \varepsilon) = \left( \frac{\varepsilon \sqrt{r}}{\sqrt{2}} \right)^{2N} \frac{1}{N} \quad (1.8)$$

and we use the generic notation  $C_N(r, \varepsilon)$  for every constant of the form

$$C_N(r, \varepsilon) = C(N!)^{q_1} e^{q_2 M_p} r^{-q_3} \varepsilon^{-q_2} \quad (1.9)$$

where  $C, p, q_i \in \mathbb{N}_*, i = 1, \dots, 4$  are universal constants (independent of the parameters  $M_p, \varepsilon, r$  and on  $N$ ) and which may change from a line to another.

We first estimate the error which is done by replacing a sequence  $Z = (Z_k)_{k \in \mathbb{N}}$  with another sequence  $\bar{Z} = (\bar{Z}_k)_{k \in \mathbb{N}}$ : this is the invariance principle. We recall first Theorem 3.1 from [3] which concerns smooth test functions (notice that here the hypothesis (1.1) is not necessary):

**Theorem 1.4** *Let  $Z = (Z_k)_{k \in \mathbb{N}}$  and  $\bar{Z} = (\bar{Z}_k)_{k \in \mathbb{N}}$  be two sequences of centered independent random variables such that  $\mathbb{E}(Z_k^2) = \mathbb{E}(\bar{Z}_k^2) = 1$ . We also assume that  $\mathbb{E}(|Z_k|^3) \leq M_3$  and  $\mathbb{E}(|\bar{Z}_k|^3) \leq M_3$ . Then for every  $f \in C_b^3(\mathbb{R})$  and every  $c \in \mathcal{C}$*

$$|\mathbb{E}(f(S_N(c, Z))) - \mathbb{E}(f(S_N(c, \bar{Z})))| \leq (N+1)!^2 M_3^{4N} \|f'''\|_\infty \|c\|_N \bar{\delta}_N(c). \quad (1.10)$$

The aim of the present paper is to obtain a similar estimate but to replace  $\|f'''\|_\infty$  by  $\|f\|_\infty$ , that is to work in total variation distance. This has already been done in [3] (see Theorem 6.1 therein) but there the estimate involves the fourth cumulant (and not only  $\bar{\delta}_N(c)$ ). So, if we aim to use such estimates in order to study the convergence of a sequence  $S_N(c_n, Z)$ ,  $n \in \mathbb{N}$ , then the limit has to be a Gaussian random variable (this is a consequence of the Fourth Moment Theorem of Nualart and Peccati [19]). In the present paper we prove the following estimate in terms of  $\bar{\delta}_N(c)$  (which allows to study the convergence to general laws):

**Theorem 1.5** *Let  $Z = (Z_k)_{k \in \mathbb{N}}$  and  $\bar{Z} = (\bar{Z}_k)_{k \in \mathbb{N}}$  be two sequences of random variables which belong to  $\mathcal{L}((M_p)_{p \in \mathbb{N}}, r, \varepsilon)$  and let  $c \in \mathcal{C}$ . Then, for every  $N$  and for every bounded and measurable function  $f : \mathbb{R} \rightarrow \mathbb{R}$*

$$\begin{aligned} |\mathbb{E}(f(S_N(c, Z))) - \mathbb{E}(f(S_N(c, \bar{Z})))| \leq & C_{N+1}(r, \varepsilon) \|f\|_\infty (1 + \|c\|_N) \\ & \times \left( \frac{\bar{\delta}_N^{\frac{1}{4+3p_*N}}(c)}{|c|_N^{\frac{6p_*}{4+3p_*N}}} + \exp\left(-\frac{c_N(r, \varepsilon) |c|_N^2}{\bar{\delta}_N^2(c)}\right) \right), \end{aligned} \quad (1.11)$$

with  $\|c\|_N$  and  $|c|_N$  defined in (1.4) and  $p_*$  is the universal constant which appears in (2.8).

Similar but less precise results have been obtained before. Assume for a moment that we replace  $S_N(c, Z)$  by  $\Phi_N(c, Z) := \sum_{|\alpha|=N} c(\alpha) Z^\alpha$ . A first result, concerning convergence in law, has been obtained in the pioneering papers of de Jong [9, 10]. Afterwards, in [17] the authors prove convergence in Kolmogorov distance, that is

$$\sup_x |\mathbb{E}(1_{(-\infty, x]}(\Phi_N(c, Z))) - \mathbb{E}(1_{(-\infty, x]}(\Phi_N(c, \bar{Z})))| \rightarrow 0 \quad \text{as } \delta_N(c) \rightarrow 0.$$

These results hold for general random variables  $Z_k$ , condition (1.1) being not needed. And recently, Nourdin and Poly in [23] assume (1.1), and they prove that

$$\sup_{\|f\|_\infty \leq 1} |\mathbb{E}(f(\Phi_N(c, Z))) - \mathbb{E}(f(\Phi_N(c, \bar{Z})))| \rightarrow 0 \quad \text{as } \delta_N(c) \rightarrow 0.$$

A first progress in our paper is that we consider a general sum  $S_N(c, Z)$  and not only  $\Phi_N(c, Z)$ . And more important, we obtain an estimate of the error - and this is not asymptotic, but holds for every fixed  $c \in \mathcal{C}$ .

The drawback of the estimate (1.11) is that it rapidly degrades as  $N$  becomes large. This point is a consequence of the techniques we use here: we use a stochastic variation calculus (analogues to the Malliavin calculus) and the delicate point is to estimate the Malliavin covariance matrix associated to our series; in order to do this we use Carbery-Wright inequality which concerns general polynomials and which make appear  $1/N$  as a power of  $\bar{\delta}_N(c)$ . One may compare this estimate with the similar one which is given in Theorem 6.2 in [3]. There the upper bound is given in terms of the fourth cumulant  $\kappa_{4,N}(c)$  of  $\Phi_N(c, G)$ , where  $G = (G_k)_{k \in \mathbb{N}}$  with  $G_k$

are independent standard normal random variables. And that upper bound is of the form  $C_N(r, \varepsilon) \kappa_{4,N}^{1/2}(c)$  when  $c(\alpha) = 0$  for  $|\alpha| = 1$ , otherwise the power is no more 1/2 but 1/4. In any case, the power of  $\kappa_{4,N}(c)$  does not depend on  $N$ . However we stress that the two estimates may not be directly compared because  $\bar{\delta}_N(c) \leq \kappa_{4,N}(c)$ , and it is possible that  $\bar{\delta}_N(c)$  is much smaller than  $\kappa_{4,N}(c)$  (see e.g. the example developed in in Section 3.2).

The estimate of the Malliavin covariance matrix is done in [3] using some martingale techniques which take into account the specific structure of the stochastic series at hand and so are more powerful than estimates concerning general polynomials (as in the Carbery-Wright inequality). But they make appear the fourth cumulant  $\kappa_{4,N}(c)$  which does not converge to zero, except in the case when we focus on a Gaussian limit (as it is pointed out by the fourth moment theorem of Nualart and Peccati [27]). So, if we aim to general limits, we have to come back to the Carbery-Wright lemma (which does not involve cumulants).

We give now some estimates of the error in the convergence in total variation of a sequence  $S_N(c_n, Z)$ ,  $n \in \mathbb{N}$  to a probability measure  $\mu$ . We will work with the metrics

$$d_k(F, G) = \sup\{|\mathbb{E}(f(F)) - \mathbb{E}(f(G))| : \|f\|_{k,\infty} \leq 1\} \quad (1.12)$$

where

$$\|f\|_{k,\infty} := \sum_{p=0}^k \|f^{(p)}\|_{\infty}. \quad (1.13)$$

In particular  $d_0 = d_{TV}$  is the total variation distance and  $d_1 = d_{FM}$  is the Fortet Mourier distance (which metrizes the convergence in law).

**Theorem 1.6** *Let  $X$  be a random variable which is the limit in law of  $(S_M(c_n, \bar{Z}))_n$ , for some  $M \in \mathbb{N}_*$ , where  $(\bar{Z}_k)_{k \in \mathbb{N}} \in \mathcal{L}((M_p)_{p \in \mathbb{N}}, r, \varepsilon)$  and, for every  $n \in \mathbb{N}$ ,  $c_n \in \mathcal{C}$  satisfies (1.5), (1.6) and (1.7) with  $N$  replaced by  $M$ . Set*

$$\bar{C}_{M,X} = \limsup_n \|c_n\|_M \quad \text{and} \quad \underline{C}_{M,X} = \liminf_n |c_n|_M. \quad (1.14)$$

*Then for every  $c \in \mathcal{C}$  and  $(Z_k)_{k \in \mathbb{N}} \in \mathcal{L}((M_p)_{p \in \mathbb{N}}, r, \varepsilon)$  one has*

$$d_0(S_N(c, Z), X) \leq C_{N \vee M}(r, \varepsilon)(1 + \|c\|_N + \bar{C}_{M,X}) \left( \frac{d_1^{2+p_*N \vee M}(S_N(c, Z), X)}{(|c|_N^{2/N} \wedge \underline{C}_{M,X}^{2/M})^{\frac{p_*N \vee M}{2+p_*N \vee M}}} + \exp\left(-\frac{c_N(r, \varepsilon)|c|_N^2}{\delta_N^2(c)}\right) \right), \quad (1.15)$$

$c_N(r, \varepsilon)$  and  $C_N(r, \varepsilon)$  being as in (1.8) and (1.9) respectively and  $p_*$  is the universal constant from (2.8).

We can rewrite Theorem 1.6 by using the concept of “ $M$ -attainability”.

**Definition 1.7** *Given  $M \in \mathbb{N}$  and  $(M_p)_{p \in \mathbb{N}}, \varepsilon, r > 0$  we say that  $X$  is  $M$ -attainable of class  $(M_p)_{p \in \mathbb{N}}, \varepsilon, r$  if there exists a sequence of coefficients  $c_n \in \mathcal{C}$  which satisfy (1.5), (1.6) and (1.7) with  $N$  replaced by  $M$  and a sequence  $Z^n = (Z_k^n)_{k \in \mathbb{N}} \in \mathcal{L}((M_p)_{p \in \mathbb{N}}, r, \varepsilon)$  such that  $\lim_n S_M(c_n, Z^n) = X$  in law. If  $X$  is  $M$ -attainable, we set  $\bar{C}_{M,X}$  and  $\underline{C}_{M,X}$  as in (1.14). We denote by  $\mathcal{A}_M((M_p)_{p \in \mathbb{N}}, \varepsilon, r)$  this class.*

If  $M = 1$  the CLT for non identically distributed random variables shows that the only 1-attainable random variable is the standard normal one. And if  $M = 2$  a characterization of the 2-attainable laws is given in [24] (see also [30]). Of course they include random variables equal in law to elements in the second chaos. And more generally, elements in a fixed  $M$  chaos are  $M$ -attainable. So, as an immediate consequence of Theorem 1.6 we obtain the following

**Corollary 1.8** *Let  $Z \in \mathcal{L}((M_p)_{p \in \mathbb{N}}, r, \varepsilon)$ ,  $c \in \mathcal{C}$  and  $X \in \mathcal{A}_M((M_p)_{p \geq 1}, r, \varepsilon)$ . Let  $p_*$  be the universal constant from (2.8). Then  $d_0(S_M(c, Z), X)$  satisfies inequality (1.15).*

The proofs of the above results are given in Section 2.3.

Finally we give several examples of applications.

First, in Theorem 3.2, we estimate the distance between  $\Phi_N(c, Z)$  and a  $\chi_2$  law with  $m$  degrees of freedom. This significantly straighten a result of Nourdin and Peccati from [20] concerning approximation of the law of a multiple stochastic integral by a  $\chi_2$  law with  $m$  degrees of freedom: the result in [20] concerns Wiener multiple integrals and the estimate is in  $d_1$  distance, while here we have a general sequence of random variables  $Z_k$  and the estimate is in terms of  $d_0$ .

In a second application we prove that

$$S_n(Z) = \frac{n!}{\sqrt{2n \ln n}} \sum_{1 < i < j \leq n} \frac{1}{\sqrt{j-i}} Z_i Z_j$$

converges to the standard normal distribution and the total variation distance to the limit is upper bounded by  $(n^{-1} \ln^2 n)^{1/43(1+2p_*)}$ . We notice that if the interaction potential  $|j-i|^{-1/2}$  is replaced by  $|j-i|^{-p}$  with  $p \in (0, 1/2)$ , then (with a suitable renormalization) the above sum converges to a double stochastic integral. So, if  $p = 1/2$  we have a contraction phenomenon whereas such a phenomenon does not exist if  $p < 1/2$ .

Finally, in the third example we consider

$$X_i = \frac{1}{\sqrt{n}} \sum_{\substack{j=1 \\ j \neq i}}^n \frac{1}{\sqrt{|j-i|}} Z_j \quad \text{and} \quad V_n(Z) = \sum_{i=1}^n X_i^2 - EX_i^2$$

and we prove that  $V_n(Z)$  converges to a double Wiener integral and the total variation distance is upper bounded by  $(n^{-1} \ln^2 n)^{1/4(3+2p_*)}$ .

Finally in Appendix A we prove an iterated version of Hoeffding's inequality (which may be of own interest) and in Appendix B we give some estimates for integrals needed in the last two examples.

## 2 Proofs of the main results

### 2.1 Notation and preliminary results

All along we consider some sequence  $(M_p)_{p \in \mathbb{N}}$  and some  $\varepsilon, r > 0$  to be given, and we employ the notation already settled in the introduction. We consider a sequence of random variables  $Z = (Z_k)_{k \in \mathbb{N}} \in \mathcal{L}((M_p)_{p \geq 1}, r, \varepsilon)$ , so each  $Z_k$  satisfies (1.1). Then we construct a function  $\psi_r$  in the following way:

$$\theta_r(z) = 1 - \frac{1}{1 - (\frac{z}{r} - 1)^2} \quad \psi_r(z) = 1_{\{|z| \leq r\}} + 1_{\{r < |z| \leq 2r\}} e^{\theta_r(|z|)}. \quad (2.1)$$

We denote

$$m(r) = \int_{\mathbb{R}} \psi_r(|z|^2) dz \leq 2\sqrt{2r}, \quad v(r) = \frac{1}{m(r)} \int_{\mathbb{R}} z^2 \psi_r(|z|^2) dz \geq \frac{r}{3\sqrt{2}}. \quad (2.2)$$

Then  $m(r)^{-1}\psi_r(|z|^2)$  is a probability density and the corresponding random variable has mean zero and variance  $v(r)$ .

Since  $\psi_r \leq 1_{B(0,2r)}$ , the inequality (1.1) holds with  $1_{B(0,2r)}$  replaced by  $\psi_r$ . This allows to use a splitting method in order to give the following representation of the law of  $Z_k$ . We consider some independent random variables  $\chi_k, U_k, V_k, k \in \mathbb{N}$ , with

$$\begin{aligned} \mathbb{P}(\chi_k = 1) &= \varepsilon m(r), & \mathbb{P}(\chi_k = 0) &= 1 - \varepsilon m(r), \\ \mathbb{P}(U_k \in dz) &= \frac{1}{m(r)} \psi_r(|z - z_k|^2) dz \\ \mathbb{P}(V_k \in dz) &= \frac{1}{1 - \varepsilon m(r)} (\mathbb{P}(Z_k \in dz) - \varepsilon \psi_r(|z - z_k|^2) dz). \end{aligned}$$

(we stress that in in [3] the role of  $U$  and  $V$  are inverted).

Then  $\chi_k U_k + (1 - \chi_k) V_k$  has the same law as  $Z_k$  so from now on we assume that

$$Z_k = \chi_k U_k + (1 - \chi_k) V_k.$$

We will work with stochastic series based on  $Z_k$ , which we introduce now. We denote  $\Gamma_m = \mathbb{N}_*^m$ . Any  $\alpha \in \Gamma_m$  is named a multi-index and we define  $|\alpha| = m$  its length. We set  $\Gamma = \cup_m \Gamma_m$ . For  $J \in \mathbb{N}_*$  we denote  $\Gamma_m(J) = \{\alpha \in \Gamma_m : 1 \leq \alpha_i \leq J\}$  and  $\Gamma(J) = \cup_{m=1}^{\infty} \Gamma_m(J)$ . Moreover, for  $z_i \in \mathbb{R}, i \in \mathbb{N}$  and  $\alpha = (\alpha_1, \dots, \alpha_m) \in \Gamma_m$ , we denote  $z^\alpha = \prod_{i=1}^m z_{\alpha_i}$ . We denote by  $\mathcal{C}$  the class of the coefficients  $c : \Gamma \rightarrow \mathbb{R}$  which are symmetric and null on the diagonals. Then we consider a family of coefficients  $c \in \mathcal{C}$  and we work with the stochastic series

$$S_N(c, Z) = \sum_{1 \leq |\alpha| \leq N} c(\alpha) Z^\alpha = \sum_{m=1}^N \sum_{\alpha \in \Gamma_m} c(\alpha) Z^\alpha. \quad (2.3)$$

In [3] we developed a stochastic variational calculus based on  $U_k, k \in \mathbb{N}$  (the explicit expression of the density of the law of  $U_k$  is central in that calculus) but here we do not need to recall all this – we will just recall some consequences which are used in the present paper. We denote

$$c_j(\alpha) = (1 + |\alpha|)c(\alpha, j) \quad (2.4)$$

with the convention that, if  $\alpha$  is void, then  $|\alpha| = 0$  and  $c_j(\alpha) = c(j)$ . Then we define

$$\lambda_N = \lambda_{S_N(c, Z)} := \sum_{j=1}^{\infty} \chi_j |\partial_{Z_j} S_N(c, Z)|^2 = \sum_{j=1}^{\infty} \chi_j |c(j) + S_{N-1}(c_j, Z)|^2. \quad (2.5)$$

This is the ‘‘Malliavin covariance matrix’’ (in our one-dimensional case, this is a scalar) associated to  $S_N(c, Z)$  and plays a central role in our estimates. Moreover we recall the seminorms  $|c|_m$  and  $\|c\|_N$  in (1.4) and we define

$$N_q(c, M) = \left( \sum_{m=q}^N M^{m-q} \times \frac{m!}{(m-q)!} \times m! \sum_{|\alpha|=m} c^2(\alpha) \right)^{1/2} \leq N! e^{\frac{1}{2}M} \|c\|_N. \quad (2.6)$$



where  $\|c\|_N$  is defined in (1.4). In [3] (see (4.17) therein we have defined the Sobolev norms  $\|S_N(c, Z)\|_{q,p}$  and in [3] Proposition 5.5, formula (5.14), we have proved that

$$\|S_N(c, Z)\|_{q,p} \leq \frac{C}{r^{q-1}} \left( \sum_{n=1}^{q-1} |c|_n + N_q(c, M_p^2) \right).$$

So, using (2.6) we have

$$\|S_N(c, Z)\|_{q,p} \leq \frac{C}{r^{q-1}} N! e^{M_p/2} \|c\|_N. \quad (2.7)$$

In fact the only way in which  $\|S_N(c, Z)\|_{q,p}$  comes on in the present paper is just by means of the above inequality, so the reader does not need to go further in the knowledge of this quantity.

We use now a regularization lemma from [3]. Let  $\psi_1$  be the function defined in (2.1),  $m(1)$  the normalization constant from (2.2) (with  $r = 1$ ) and, for  $\delta > 0$ , let

$$\gamma_\delta(z) = \frac{1}{m(1)\sqrt{\delta}} \psi_1(\delta^{-1} |z|^2).$$

For  $f : \mathbb{R} \rightarrow \mathbb{R}$  we set  $f * \gamma_\delta$  the convolution between  $f$  and  $\gamma_\delta$ , whenever it is well defined. Using the regularization Lemma 4.6 from [3] and (2.7) we obtain

**Lemma 2.1** *There exist some universal constants  $C, p_* \geq 1$  such that for every  $\eta > 0, \delta > 0$  and for every bounded and measurable  $f : \mathbb{R} \rightarrow \mathbb{R}$  one has*

$$|\mathbb{E}(f(S_N(c, Z))) - \mathbb{E}(f * \gamma_\delta(S_N(c, Z)))| \leq C_N(r, \varepsilon) \|c\|_N \|f\|_\infty \left( \mathbb{P}(\lambda_N < \eta) + \frac{\sqrt{\delta}}{\eta^{p_*}} \right) \quad (2.8)$$

with  $C_N(r, \varepsilon) = Cr^{-2} N! e^{\frac{1}{2} M_{p_*}}$ .

We will use the following easy consequence, which is a slightly more precise version of Theorem 2.7 from [2].

**Lemma 2.2** *Let  $Z, \bar{Z} \in \mathcal{L}((M_p)_{p \geq 1}, r, \varepsilon)$  and  $c, \bar{c} \in \mathcal{C}$ . Let  $p_*$  be the universal constant from (2.8). For every  $k \in \mathbb{N}$ ,  $N, M \in \mathbb{N}_*$  there exists a universal constant  $C_{N \vee M}(r, \varepsilon)$  (depending on  $k$ ) such that for every  $\eta > 0$  one has*

$$\begin{aligned} d_0(S_N(c, Z), S_M(\bar{c}, \bar{Z})) &\leq C_{N \vee M}(r, \varepsilon) (1 + \|c\|_N + \|\bar{c}\|_M) \times \\ &\quad \times \left( \frac{1}{\eta^{\frac{k p_*}{k+1}}} d_k^{\frac{1}{k+1}}(S_N(c, Z), S_M(\bar{c}, \bar{Z})) + \mathbb{P}(\lambda_N < \eta) + \mathbb{P}(\bar{\lambda}_M < \eta) \right) \end{aligned} \quad (2.9)$$

where  $\lambda_N = \lambda_{S_N(c, Z)}$  and  $\bar{\lambda}_M = \lambda_{S_M(\bar{c}, \bar{Z})}$  are defined in (2.5),  $d_k$  is defined in (1.12) and  $C_N(r, \varepsilon)$  is a constant of the form (1.9).

**Proof.** To simplify the notation we put  $S_N = S_N(c, Z)$  and  $\bar{S}_M = S_M(\bar{c}, \bar{Z})$  and  $C$  a constant of the form  $C_{N \vee M}(r, \varepsilon) (1 + \|c\|_N + \|\bar{c}\|_M)$  (which changes from a line to another). Let  $\delta > 0$  and let  $f \in C(\mathbb{R})$  with  $\|f\|_\infty \leq 1$ . Since  $\|f * \gamma_\delta\|_{k, \infty} \leq C \delta^{-k/2}$  we have

$$|\mathbb{E}(f * \gamma_\delta(S_N)) - \mathbb{E}(f * \gamma_\delta(\bar{S}_M))| \leq C \delta^{-k/2} d_k(S_N, \bar{S}_M).$$

Then, using (2.8),

$$|\mathbb{E}(f(S_N)) - \mathbb{E}(f(\bar{S}_M))| \leq C\delta^{-k/2}d_k(S_N, \bar{S}_M) + C\left(\mathbb{P}(\lambda_N < \eta) + \mathbb{P}(\bar{\lambda}_M < \eta) + \frac{\delta^{1/2}}{\eta^{p^*}}\right).$$

We optimize over  $\delta$ : we take

$$\delta^{(k+1)/2} = d_k(S, \bar{S})\eta^{p^*}.$$

We insert this in the previous inequality and we obtain (2.9).  $\square$

## 2.2 Estimate of the covariance matrix

Our aim is to estimate  $\mathbb{P}(\lambda_N \leq \eta)$  with  $\lambda_N$  defined in (2.5) and this will be done using the Carbery-Wright inequality (we follow here an idea from [23]). In order to do this we need the following lemma.

**Lemma 2.3** *We denote by  $\mathbb{E}_{V,\chi}$  the conditional expectation with respect to  $\sigma(V_i, \chi_i, i \in \mathbb{N})$ . Let  $v(r)$  be as in (2.2). Then*

$$\mathbb{E}_{V,\chi}(\lambda_N) \geq \frac{v^{N-1}(r)}{N} \sum_{|\alpha|=N} c^2(\alpha)\chi^\alpha. \quad (2.10)$$

**Proof.** We denote

$$\bar{U}_i = U_i - \mathbb{E}(U_i) \quad \text{and} \quad \bar{V}_i = (1 - \chi_i)V_i + \chi_i\mathbb{E}(U_i)$$

so that

$$Z_i = \chi_i U_i + (1 - \chi_i)V_i = \chi_i \bar{U}_i + \bar{V}_i.$$

Then we define

$$\bar{Z}^\alpha = \sum_{\substack{(\beta,\gamma)=\alpha, \\ \gamma \neq \emptyset}} \chi^\beta \bar{U}^\beta \times \bar{V}^\gamma$$

and we write

$$Z^\alpha = \bar{Z}^\alpha + \chi^\alpha \bar{U}^\alpha.$$

Notice that for every multi-indexes  $\alpha \in \Gamma_m$  with  $m \leq N - 1$  and  $\theta \in \Gamma_N$  we have

$$\mathbb{E}_{V,\chi}(\bar{Z}^\alpha \bar{U}^\theta) = \sum_{\substack{(\beta,\gamma)=\alpha, \\ \gamma \neq \emptyset}} \chi^\beta \mathbb{E}_{V,\chi}(\bar{U}^\beta \bar{U}^\theta) \times \bar{V}^\gamma = 0. \quad (2.11)$$

This is because  $|\beta| < m \leq N - 1$ , so there is at least one  $\theta_i \notin \beta$  and  $\mathbb{E}_{V,\chi}(\bar{U}^{\theta_i}) = 0$ . We take now  $\kappa \in \mathbb{R}$  and we consider the r.v.  $X = \kappa + S_N(c, Z)$ . We write  $\kappa + S_N(c, Z) = S' + S''$  with  $S' = \sum_{|\alpha|=N} c(\alpha)\chi^\alpha \bar{U}^\alpha$  and  $S'' = \kappa + S_N(c, Z) - S'$ . By (2.11),  $S'$  and  $S''$  are orthogonal in  $L^2(\mathbb{P}_{V,\chi})$  so that

$$\mathbb{E}_{V,\chi}((\kappa + S_N(c, Z))^2) \geq \mathbb{E}_{V,\chi}(S'^2) = \sum_{|\alpha|=N} v^N(r)c^2(\alpha)\chi^\alpha, \quad (2.12)$$

the last equality being a consequence of  $\mathbb{E}(\bar{U}_i^2) = v(r)$ .

Consider now  $\lambda_N$ : by (2.5),  $\lambda_N = \sum_{j=1}^{\infty} \chi_j |c(j) + S_{N-1}(c_j, Z)|^2$ . We use (2.12) with  $\kappa = c(j)$  and  $c$  replaced by  $c_j$  (see (2.4)) and we obtain

$$\begin{aligned} \mathbb{E}_{V,\chi}(\lambda_N) &= \sum_{j=1}^{\infty} \chi_j \mathbb{E}_{V,\chi}(|c(j) + S_{N-1}(c_j, Z)|^2) \geq v^{N-1}(r) \sum_{j=1}^{\infty} \chi_j \sum_{|\alpha|=N-1} c_j^2(\alpha) \chi^\alpha \\ &= v^{N-1}(r) \sum_{j=1}^{\infty} \chi_j \sum_{|\alpha|=N-1} c^2(\alpha, j) \chi^\alpha = \frac{1}{N} v^{N-1}(r) \sum_{|\beta|=N} c^2(\beta) \chi^\beta. \end{aligned}$$

□

We are now able to give our estimate:

**Lemma 2.4** *Let  $c \in \mathcal{C}$ . For every  $\eta > 0$ ,*

$$\mathbb{P}(\lambda_N \leq \eta) \leq C_N(r, \varepsilon) \left( \left( \frac{\eta}{|c|_N^2} \right)^{1/N} + \exp \left( - \frac{c_N(r, \varepsilon) |c|_N^2}{\delta_N^2(c)} \right) \right) \quad (2.13)$$

with  $C_N(r, \varepsilon)$  denotes a constant of the type (1.9) and  $c_N(r, \varepsilon)$  is given in (1.8).

**Proof.** We chose  $J$  sufficiently large in order to have

$$\sum_{\alpha \in \Gamma_N(J)} c^2(\alpha) \geq \frac{1}{2} \sum_{\alpha \in \Gamma_N} c^2(\alpha) = \frac{1}{2} |c|_N^2. \quad (2.14)$$

We will use the Carbery–Wright inequality that we recall here (see Theorem 8 in [8]). Let  $\mu$  be a probability law on  $\mathbb{R}^J$  which is absolutely continuous with respect to the Lebesgue measure and has a log-concave density. There exists a universal constant  $K$  such that for every polynomial  $Q(x)$  of order  $N$  and for every  $\eta > 0$  one has

$$\mu(x : |Q(x)| \leq \eta) \leq KN(\eta/V_\mu(Q))^{1/N} \quad (2.15)$$

with  $V_\mu(Q) = (\int Q^2(x) d\mu(x))^{1/2}$ .

We will use this result in the following framework. We recall that  $\mathbb{P}_{V,\chi}$  is the conditional probability with respect to  $\sigma(V_i, \chi_i, i \in \mathbb{N})$  and we look to

$$Q(U_1, \dots, U_J) := \sum_{j=1}^{\infty} \chi_j |\partial_{Z_j} S_N(c1_{\Gamma(J)}, Z)|^2 =: \lambda_{N,J}$$

as to a polynomial of order  $N$  of  $U_1, \dots, U_J$ . It is easy to see that the density of the law  $\mu$  of  $(U_1, \dots, U_J)$  (under  $\mathbb{P}_{V,\chi}$ ) is log-concave. So we are able to use (2.15). Using (2.10)

$$\begin{aligned} V_\mu(Q) &= \left( \int Q^2(x) d\mu(x) \right)^{1/2} \geq \int |Q(x)| d\mu(x) = \mathbb{E}_{V,\chi} \left( \sum_{j=1}^{\infty} \chi_j |\partial_{Z_j} S(c1_{\Gamma(J)}, Z)|^2 \right) \\ &\geq \frac{v^{N-1}(r)}{N} \sum_{|\beta|=N} c^2(\beta) 1_{\Gamma(J)}(\beta) \chi^\beta. \end{aligned}$$

We take now  $\theta > 0$  (to be chosen in a moment) and we use (2.15) in order to obtain

$$\begin{aligned}
\mathbb{P}(\lambda_{N,J} \leq \eta) &= \mathbb{P}(Q(U_1, \dots, U_J) \leq \eta) \\
&\leq \mathbb{P}(V_\mu(Q) \leq \theta) + \mathbb{E}(\mathbb{P}_{V,\chi}(Q(U_1, \dots, U_J) \leq \eta) 1_{\{V_\mu(Q) \geq \theta\}}) \\
&\leq \mathbb{P}\left(\sum_{|\beta|=N} c^2(\beta) 1_{\Gamma(J)}(\beta) \chi^\beta \leq \frac{\theta N}{v^{N-1}(r)}\right) + KN(\eta/\theta)^{1/N}. \tag{2.16}
\end{aligned}$$

The first term in the above right hand side is estimated in Appendix A: we apply Lemma A.1 with  $x = \theta N/v^{N-1}(r)$  and with the coefficients  $\bar{c}_J(\alpha) = c(\alpha) 1_{\Gamma_N(J)}(\alpha)$ , so that  $S_N(\bar{c}_J^2, \chi) = \sum_{|\beta|=N} c^2(\beta) 1_{\Gamma(J)}(\beta) \chi^\beta$ . By (2.14) we have

$$\frac{|c|_N^2}{2} \leq \|\bar{c}_J\|_N^2 = |\bar{c}_J|_N^2 \leq |c|_N^2.$$

We recall that in Lemma A.1 we use  $p = \varepsilon m(r)$  and that we need (see (A.1)) that

$$\theta = \frac{v^{N-1}(r)}{N} x \leq \frac{v^{N-1}(r)}{2N} \left(\frac{p}{4}\right)^{2N} |c|_N^2. \tag{2.17}$$

We take  $\theta$  equal to the quantity in the right hand side of the above inequality so that

$$x = \frac{\theta N}{v^{N-1}(r)} = \frac{1}{2} \left(\frac{p}{4}\right)^{2N} |c|_N^2.$$

Then (A.2) gives

$$\mathbb{P}\left(S_N(\bar{c}_J^2, \chi) \leq \frac{\theta N}{v^{N-1}(r)}\right) \leq \frac{2e^3}{9} N \exp\left(-\frac{1}{4} \left(\frac{p}{4}\right)^{4N} |c|_N^4 \frac{1}{N \bar{\delta}_N^2(\bar{c}_J) \|\bar{c}_J\|_N^2}\right).$$

Since  $\bar{\delta}_N^2(\bar{c}_J) \leq \delta_N^2(c)$  and  $\|\bar{c}_J\|_N^2 \leq |c|_N^2$  we upper bound the above term with

$$\frac{2e^3}{9} N \exp\left(-\frac{1}{4N} \left(\frac{p}{4}\right)^{4N} \frac{|c|_N^2}{\delta_N^2(c)}\right).$$

Inserting this in (2.16) we obtain

$$\begin{aligned}
\mathbb{P}(\lambda_N \leq \eta) &\leq \mathbb{P}(\lambda_{N,J} \leq \eta) \\
&\leq \frac{2e^3}{9} N \exp\left(-\frac{1}{4N} \left(\frac{\varepsilon m(r)}{4}\right)^{4N} \frac{|c|_N^2}{\delta_N^2(c)}\right) + \frac{KN}{v(r) \varepsilon^2 m^2(r) |c|_N^{2/N}} \eta^{1/N}.
\end{aligned}$$

and the proof is completed.  $\square$

### 2.3 Proof of the main results

Our basic lemma is the following:

**Lemma 2.5** Let  $Z, \bar{Z} \in \mathcal{L}((M_p)_{p \geq 1}, r, \varepsilon)$  and  $c, \bar{c} \in \mathcal{C}$ . We denote  $S_N = S_N(c, Z)$  and  $\bar{S}_M = S_M(\bar{c}, \bar{Z})$ . Let  $p_*$  be the universal constant from (2.8). For every  $k \in \mathbb{N}$  there exist a constant  $C_{N \vee M}(r, \varepsilon)$  as in (1.9) such that

$$\begin{aligned} d_0(S_N, \bar{S}_M) \leq & C_{N \vee M}(r, \varepsilon)(1 + \|c\|_N + \|\bar{c}\|_M) \left( \frac{d_k^{\frac{1}{k+1+kp_*N \vee M}}(S_N, \bar{S}_M)}{(|c|_N^{2/N} \wedge |\bar{c}|_M^{2/M})^{\frac{kp_*N \vee M}{k+1+kp_*N \vee M}}} \right. \\ & \left. + \exp\left(-\frac{c_N(r, \varepsilon)|c|_N^2}{\delta_N^2(c)}\right) + \exp\left(-\frac{c_M(r, \varepsilon)|\bar{c}|_M^2}{\delta_M^2(\bar{c})}\right) \right), \end{aligned} \quad (2.18)$$

$c_N(\varepsilon, r)$  being given in (1.8).

**Proof.** We use Lemma 2.2 and in the estimate (2.9), we replace  $\mathbb{P}(\det \sigma_{S_N(c, Z)} < \eta)$  by the expression from (2.13). So, we obtain

$$\begin{aligned} d_0(S_N, \bar{S}_M) \leq & C_{N \vee M}(r, \varepsilon)(1 + \|c\|_N + \|\bar{c}\|_M) \left( \frac{1}{\eta^{\frac{kp_*}{k+1}}} d_k^{\frac{1}{k+1}}(S_N, \bar{S}_M) + \right. \\ & \left. + \frac{1}{|c|_N^{2/N} \wedge |\bar{c}|_M^{2/M}} (\eta^{1/N} + \eta^{1/M}) + \exp\left(-\frac{c_N(r, \varepsilon)|c|_N^2}{\delta_N^2(c)}\right) + \exp\left(-\frac{c_M(r, \varepsilon)|\bar{c}|_M^2}{\delta_M^2(\bar{c})}\right) \right). \end{aligned}$$

This holds true for every  $\eta > 0$ . We optimize over  $\eta$  and we obtain (2.18).  $\square$

**Proof of Theorem 1.1.** Let  $S_N(c_n, Z^n)$ ,  $n \in \mathbb{N}$ , be the sequence considered in the statement of the theorem. Since this sequence converges in law to  $\mu$ , it follows that it is a Cauchy sequence in  $d_1$ . And since  $\bar{\delta}_N(c_n) \rightarrow 0$ , and  $\liminf_{n \rightarrow \infty} |c|_N^2 > 0$  the inequality (2.18) says that the sequence is Cauchy in  $d_0$ . It follows that it converges to  $\mu$  in  $d_0$ .  $\square$

**Proof of Theorem 1.5.** By Theorem 1.4,  $d_3(S_N(c, Z), S_N(c, \bar{Z})) \leq C_N(r, \varepsilon)\bar{\delta}_N(c)$ , so, using (2.18) with  $k = 3$  and  $N = M$  we obtain (1.11).  $\square$

**Proof of Theorem 1.6.** The hypotheses ensure that  $\lim_n \bar{\delta}_M(c_n) = 0$ ,  $\limsup_n \|c_n\|_M = \bar{C}_{M, X}$ ,  $\liminf_n \|c_n\|_M = \underline{C}_{M, X} > 0$  and  $\lim_n d_1(S_M(c_n, \bar{Z}), X) = 0$ . Notice that by Theorem 1.1 we know that  $\lim_n d_0(S_M(c_n, \bar{Z}), X) = 0$ . We write

$$\begin{aligned} d_0(S_N(c, Z), X) & \leq d_0(S_N(c, Z), S_M(c_n, \bar{Z})) + d_0(S_M(c_n, \bar{Z}), X) \\ & \leq C_{N \vee M}(r, \varepsilon)(1 + \|c\|_N + \|c_n\|_M) \left( \frac{d_1^{\frac{1}{2+2p_*N \vee M}}(S_N(c, Z), S_M(c_n, \bar{Z}))}{(|c|_N^{2/N} \wedge |c_n|_M^{2/M})^{\frac{p_*N \vee M}{2+2p_*N \vee M}}} \right. \\ & \left. + \exp\left(-\frac{c_N(r, \varepsilon)|c|_N^2}{\delta_N^2(c)}\right) + \exp\left(-\frac{c_M(r, \varepsilon)|c_n|_M^2}{\delta_M^2(c_n)}\right) \right) + d_0(S_M(c_n, \bar{Z}), X) \end{aligned}$$

the second inequality being (2.18). Since  $d_1(S_M(c_n, \bar{Z}), X) \rightarrow 0$  then  $d_1(S(c, Z), S_M(c_n, \bar{Z})) \rightarrow d_1(S(c, Z), X)$ . We also have  $\exp\left(-\frac{c_M(r, \varepsilon)|c_n|_M^2}{\delta_M^2(c_n)}\right) + d_0(S_M(c_n, \bar{Z}), X) \rightarrow 0$  so (1.15) is proved.  $\square$

**Proof of Corollary 1.8.** Since  $X \in \mathcal{A}_M((M_p)_{p \geq 1}, r, \varepsilon)$  we may find a sequence  $Z^{(n)} = (Z_k^{(n)})_{k \in \mathbb{N}} \in \mathcal{L}((M_p)_{p \geq 1}, \varepsilon, r)$  and a sequence  $(c_n)_n \subset \mathcal{C}$  that verifies the requests of Theorem 1.6. So, the statement holds by repeating the proof of Theorem 1.6.  $\square$

### 3 Examples

#### 3.1 Approximation with a chi-squared law

In [20], Nourdin and Peccati give sufficient conditions in order to estimate the Fortet-Mourier distance ( $d_1$  in our notation) between a multiple Wiener integral and a random variable with a centred Gamma distribution. It is not clear if the Gamma distribution with fractional coefficient is attainable in the sense of Definition 1.7, so we are not able to use our results in the general case. But for an integer parameter  $\nu = 2m$ , the Gamma distribution coincides with the  $\chi^2$  distribution with  $m$  degrees of freedom, and this law is clearly attainable (just represent it as  $\sum_{k=0}^{m-1} 2 \int_k^{k+1} W_s dW_s + m$  and then use approximation with Riemann sums). So we restrict ourself to this case. One looks to

$$\Phi_N(c, Z) = \sum_{\alpha \in \Gamma_N} c(\alpha) Z^\alpha.$$

If  $Z_k, k \in \mathbb{N}$ , are standard Gaussian random variables, then  $\Phi_N(c, Z)$  is a multiple stochastic integral and in this case Nourdin and Peccati in [20] have proved the following result. In order to present it we have to introduce some notation. For  $0 \leq r \leq N$  and  $\alpha, \beta \in \Gamma_{N-r}$  one denotes  $c \otimes_r c(\alpha, \beta) = \sum_{\gamma \in \Gamma_r} c(\alpha, \gamma) c(\beta, \gamma)$  with the convention that for  $r = 0$  we put  $c \otimes_0 c(\alpha, \beta) = c(\alpha) c(\beta)$  and for  $r = N$ ,  $c \otimes_N c = \sum_{\gamma \in \Gamma_N} c(\gamma) c(\gamma)$ . Notice that even if  $c$  is symmetric,  $c \otimes_r c$  is not symmetric, so we introduce  $c \tilde{\otimes}_r c$  to be the symmetrization of  $c \otimes_r c$ . Finally, if  $N$  is an even number, we introduce

$$\begin{aligned} \kappa_{m,N}(c) &= (m - N! |c|_N^2)^2 + 4N! |\theta_N \times c \tilde{\otimes}_{N/2} c - c|_N^2 \\ &\quad + N^2 \sum_{\substack{r \in \{1, \dots, N-1\} \\ r \neq N/2}} (2N - 2r)! (r - 1)!^2 \binom{N-1}{r-1}^4 |c \otimes_r c|_N^2 \end{aligned}$$

with  $\theta_N = \frac{1}{4} (N/2)! \binom{N}{N/2}$ . Combining Theorem 3.11 and Proposition 3.13 from [20] one obtains the following:

**Theorem 3.1** *Let  $N$  be an even integer and let  $F(m) = \sum_{k=1}^m G_k^2 - m$  with  $G_k$  independent standard Gaussian random variables. Assume also that  $Z_k, k \in \mathbb{N}$  are independent standard Gaussian random variables. Then*

$$d_1(\Phi_N(c, Z), F(m)) \leq K_1(m) \kappa_{m,N}^{1/2}(c)$$

with  $K_1(m) = \max\{\sqrt{\pi/m}, 1/2m + 1/2m^2\}$ .

As an immediate consequence of Corollary 1.8 we obtain the following result:

**Theorem 3.2** *Let  $N$  be an even integer,  $m \in \mathbb{N}_*$  and let  $F(m) = \sum_{k=1}^m G_k^2 - m$  with  $G_k$  independent standard Gaussian random variables. Assume also that  $Z \in \mathcal{L}((M_p)_{p \geq 1}, r, \varepsilon)$ , and  $c \in \mathcal{C}$ . Then*

$$d_0(S_N(c, Z), X) \leq C_{N \vee 2}(r, \varepsilon) (1 + \|c\|_N) \left( \frac{\kappa_{m,N}^{\frac{1}{4+2p_* N \vee M}}(c)}{|c|_N^{\frac{2p_* N \vee 2}{N(2+p_* N \vee 2)}}} + \exp\left(-\frac{c_N(r, \varepsilon) |c|_N^2}{\delta_N^2(c)}\right) \right).$$

### 3.2 An example of quadratic CLT

An easy way to construct examples of invariance principles is to take a double stochastic integral, to discretize it, and then to replace the Brownian increments (renormalized) with some general random variables. So, for example, starting with  $\int_0^1 \int_0^t f(t, s) dW_s dW_t$  we construct the approximation

$$\sum_{0 \leq i < j \leq 1} f\left(\frac{i}{n}, \frac{j}{n}\right) \frac{\Delta_i}{\sqrt{n}} \frac{\Delta_j}{\sqrt{n}}$$

with  $\Delta_i, i \in \mathbb{N}$  independent standard Gaussian random variables. Then we replace  $\Delta_i$  by some general  $Z_i$  and we obtain our invariance principle. Notice however that using this strategy double sums give double integrals - so we remain in the same chaos. This is true if  $f$  is a square integrable function. In contrast, if we work with some  $f$  which is not square integrable then we may pass from a double sum to a Gaussian limit (so to an element of the first chaos): a contraction phenomenon is at work. In this section we give an example which illustrates this fact. We will study the convergence to normality of the following stochastic series. We denote

$$S_n(Z) = \frac{1}{\sqrt{2n \ln n}} \sum_{i, j \geq 1}^n \frac{1}{\sqrt{|i-j|}} \mathbf{1}_{\{i \neq j\}} Z_i Z_j.$$

Notice that

$$S_n(Z) = S_2(c_n, Z) \quad \text{with} \quad c_n(i, j) = \frac{1}{\sqrt{2n \ln n}} \frac{1}{\sqrt{|i-j|}} \mathbf{1}_{\{i \neq j\}} \mathbf{1}_{(i, j) \in \Gamma_2(n)} \quad (3.1)$$

**Theorem 3.3 A.** *Let  $Z = (Z_i)_{i \in \mathbb{N}} \in \mathcal{L}((M_p)_p, r, \varepsilon)$  and let  $G = (G_i)_{i \in \mathbb{N}}$  be a sequence of standard normal random variables. Then*

$$d_0(S_n(Z), S_n(G)) \leq \frac{C_2(r, \varepsilon)}{n^{1/(4+6p_*)}} \quad (3.2)$$

where  $C_2(r, \varepsilon) = C(M_p r^{-1} \varepsilon^{-1})^q$  with some universal constants  $C, p, q$ , and  $p_*$  is the universal constant from (2.8)

**B.** *Let  $W$  be a standard normal random variable. Then*

$$d_0(S_n(Z), W) \leq \frac{C_2(r, \varepsilon)}{\ln n}. \quad (3.3)$$

**Remark 3.4** *Using the strategy mentioned in the beginning of this section we may easily prove that, for  $p < \frac{1}{2}$ ,*

$$\frac{1}{n^{1-p}} \sum_{i < j \leq n} \frac{1}{(j-i)^p} Z_i Z_j = \sum_{i < j \leq n} \frac{1}{\left(\frac{j}{n} - \frac{i}{n}\right)^p} \frac{Z_i}{\sqrt{n}} \frac{Z_j}{\sqrt{n}} \xrightarrow{\mathcal{L}} \int_0^1 \int_0^t \frac{1}{(t-s)^p} dW_s dW_t.$$

Notice that in this case we start with the function  $f(t, s) = |t-s|^{-p}$  which is square integrable for  $p < \frac{1}{2}$ . So, with a soft singularity ( $p < 1/2$ ) we remain in the second chaos. But with a strong singularity ( $p = 1/2$ ), a contraction phenomenon is at work and we pass in the first chaos.

**Proof. A.** We apply Theorem 1.5. Here,  $N = 2$ ,  $\|c_n\|_2 = |c_n|_2$  and  $\bar{\delta}_2(c_n) = \delta_2(c_n)$ . So, by using (1.11) we have

$$d_0(S_n(Z), S_n(G)) \leq C_3(r, \varepsilon)(1 + \|c_n\|_2) \left( \frac{\bar{\delta}_2^{\frac{1}{4+6p^*}}(c_n)}{|c_n|_2^{\frac{6p^*}{4+6p^*}}} + \exp\left(-\frac{c_2(r, \varepsilon)|c_n|_2^2}{\bar{\delta}_2^2(c_n)}\right) \right),$$

and (3.2) immediately follows by using the estimates in (B.6) e (B.7).

**B.** Let us prove (3.3). We notice that

$$S_n(G) \stackrel{\mathcal{L}}{=} \int_0^\infty \int_0^t f_n(t, s) dW_s dW_t = I_2(f_n),$$

where  $(W_t)_t$  denotes a Brownian motion and

$$f_n(s, t) = \sum_{i, j=1}^n c_n(i, j) 1_{(i, i+1]}(s) 1_{(j, j+1]}(s).$$

Then we can use the results in [20] and we have that  $d(S_n(G), W) \leq C\sqrt{\kappa(f_n)}$  where  $\kappa(f_n)$  is the fourth cumulant of  $I_2(f_n)$ . And since  $\kappa(f_n) \leq C\|f_n \otimes_1 f_n\|_{\mathcal{L}^2}^2 = C\sum_{i, j} (c_n \otimes_1 c_n)^2(i, j)$ , (3.3) is a consequence of (B.13).  $\square$

### 3.3 A variance-type estimator

We denote

$$X_i = \frac{1}{\sqrt{n}} \sum_{\substack{j=1, \\ j \neq i}}^n \frac{1}{\sqrt{|i-j|}} Z_j$$

and we study the asymptotic behavior of

$$V_n(Z) = \sum_{i=1}^n (X_i^2 - \mathbb{E}(X_i^2)).$$

The limit will be given by the double stochastic integral

$$I_2(\phi) = \int_0^1 \int_0^1 \phi(t, s) dW_t dW_s$$

where the function  $\phi$  is defined in (B.2):

$$\phi(t, s) = \int_0^1 \frac{du}{\sqrt{|(t-u)(s-u)|}} = \pi + 2 \ln \frac{\sqrt{1-t} + \sqrt{1-s}}{|\sqrt{t} - \sqrt{s}|}$$

**Proposition 3.5** *Let  $Z = (Z_i)_{i \in \mathbb{N}} \in \mathcal{L}((M_p)_p, r, \varepsilon)$  Then*

$$d_0(V_n(Z), I_2(\phi)) \leq C_2(r, \varepsilon) \left(\frac{\ln^2 n}{n}\right)^{1/4(1+2p^*)}. \quad (3.4)$$



**Proof.** In this proof we refer several times to some computations and estimates which are developed in Appendix B.

**Step 1.** We denote

$$\begin{aligned} a(i, j) &= 1_{\{i \neq j\}} |i - j|^{-1/2}, \\ \bar{c}_n(i, j) &= \frac{1}{n} (a \otimes_1 a)(i, j) := \frac{1}{n} \sum_{k=1}^n a(i, k) a(j, k). \end{aligned}$$

We recall that in (B.9), (B.10) and (B.11) one proves that

$$\delta_2^2(\bar{c}_n) \leq \frac{C \ln^2 n}{n}, \quad 0 < c_* \leq |\bar{c}_n|^2 \leq C \quad \text{and} \quad \sum_{k=1}^n \bar{c}_n^2(k, k) \leq \frac{C \ln^2 n}{n}. \quad (3.5)$$

We decompose

$$V_n(Z) = V_n'(Z) + V_n''(Z)$$

with

$$\begin{aligned} V_n'(Z) &= \frac{2}{n} \sum_{j < j'} a \otimes_1 a(j, j') Z_j Z_{j'} = \Phi_2(\bar{c}_n, Z) \quad \text{and} \\ V_n''(Z) &= \frac{1}{n} \sum_{j=1}^n \left( \sum_{i=1}^n a^2(i, j) \right) (Z_j^2 - 1). \end{aligned}$$

Since  $V_n''(Z)$  contains terms of the form  $Z_j^2$  we may not use directly the results from the previous sections, and we are obliged to develop a slight variant of them.

**Step 2.** By (1.10)

$$d_3(V_n'(Z), V_n'(G)) \leq C \bar{\delta}_2(\bar{c}_n) = C \delta_2(\bar{c}_n) \leq \frac{C \ln^2 n}{n}.$$

And by the isometry property

$$\begin{aligned} \mathbb{E}(|V_n''(Z)|^2) &= \frac{1}{n^2} \sum_{j=1}^n \left( \sum_{i=1}^n a^2(i, j) \right)^2 \mathbb{E}((Z_j^2 - 1)^2) \\ &\leq \frac{1}{n^2} \max_j (\mathbb{E}(Z_j^4) - 1) \sum_{j=1}^n \left( \sum_{i=1}^n a^2(i, j) \right)^2 \\ &= \frac{1}{n^2} \max_j (\mathbb{E}(Z_j^4) - 1) \sum_{j=1}^n \bar{c}_n^2(j, j) \\ &\leq \frac{C \ln^2 n}{n}. \end{aligned}$$

So

$$d_3(V_n(Z), V_n(G)) \leq \frac{C \ln^2 n}{\sqrt{n}}. \quad (3.6)$$

**Step 3.** We will use the stochastic calculus of variations for  $V_n(Z)$  so we have to estimate the Sobolev norms and the covariance matrix. First

$$\| \| V_n(Z) \| \|_{q,p} \leq \| \| V_n'(Z) \| \|_{q,p} + \| \| V_n''(Z) \| \|_{q,p} \leq C_2(r, \varepsilon) |\bar{c}_n|_2^2. \quad (3.7)$$

This is because the estimate of  $\|V_n'(Z)\|_{q,p}$  is already given in (2.7) and the estimate of  $\|V_n''(Z)\|_{q,p}$  is analogous (it suffices to follow the computations in Proposition 5.3 and 5.4 in [3]), so we skip it.

We estimate now the covariance matrix (scalar in our case) defined in (2.5):

$$\lambda_{V_n(Z)} = \sum_{k=1}^n \chi_k |\partial_{Z_k} V_n(Z)|^2.$$

We have

$$\begin{aligned} \partial_{Z_k} V_n(Z) &= 2 \sum_{i=1}^n X_i \partial_{Z_k} X_i = \frac{2}{\sqrt{n}} \sum_{i=1}^n X_i a(i, k) \chi_k \\ &= \frac{2}{n} \chi_k \sum_{i=1}^n a(i, k) \sum_{j=1}^n a(i, j) Z_j \\ &= 2 \chi_k \sum_{j=1}^n Z_j \left( \frac{1}{n} \sum_{i=1}^n a(i, k) a(i, j) \right) = 2 \chi_k \sum_{j=1}^n Z_j \bar{c}_n(j, k) \end{aligned}$$

so that

$$\lambda_{V_n(Z)} = 4 \sum_{k=1}^n \chi_k \left| \sum_{j=1}^n Z_j \bar{c}_n(j, k) \right|^2.$$

This expression is strongly similar to  $\lambda_{S_2(\bar{c}_n, Z)}$  defined in (2.5), but there is one difference: we do not have the property  $\bar{c}_n(j, j) = 0$ . So we have to eliminate the diagonal terms. We define  $c'_n(i, j) = 1_{\{i \neq j\}} \bar{c}_n(i, j)$  and we use the inequality  $(a + b)^2 \geq \frac{1}{2}a^2 - b^2$  in order to obtain

$$\begin{aligned} \lambda_{V_n(Z)} &\geq 2 \sum_{k=1}^n \chi_k \left| \sum_{j=1}^n Z_j c'_n(j, k) \right|^2 - 4 \sum_{k=1}^n \chi_k \bar{c}_n^2(k, k) Z_k^2 \\ &= 2 \lambda_{S_2(c'_n, Z)} - 4 \sum_{k=1}^n \chi_k \bar{c}_n^2(k, k) Z_k^2. \end{aligned}$$

Using (3.5), for  $n$  sufficiently large we have

$$|c'_n|_2^2 \geq \frac{1}{2} |\bar{c}_n|_2^2 - \sum_{k=1}^n \bar{c}_n^2(k, k) \geq \frac{c_*}{2} - \frac{C \ln^2 n}{n} \geq \frac{c_*}{4}.$$

Then, by (2.13) first and by (3.5) then, for every  $\eta > 0$ ,

$$\begin{aligned} \mathbb{P}(\lambda_{S_2(c'_n, Z)} \leq \eta) &\leq C_2(r, \varepsilon) \left( \left( \frac{\eta}{|c'_n|_2^2} \right)^{1/2} + \exp \left( - \frac{c_2(r, \varepsilon) |c'_n|_2^2}{\delta_2^2(c'_n)} \right) \right) \\ &\leq C_2(r, \varepsilon) (\eta^{1/2} + \exp(-c_2(r, \varepsilon)n)). \end{aligned}$$

And again by (3.5)

$$\mathbb{E} \left( \sum_{k=1}^n \chi_k c_n^2(k, k) Z_k^2 \right) \leq \sum_{k=1}^n c_n^2(k, k) \leq \frac{C \ln^2 n}{n}$$

so that

$$\begin{aligned}
\mathbb{P}(\lambda_{V_n(Z)} \leq \eta) &\leq \mathbb{P}\left(2\lambda_{S_2(c'_n, Z)} \leq 2\eta\right) + \mathbb{P}\left(\sum_{k=1}^n \chi_k c_n^2(k, k) Z_k^2 \geq \eta\right) \\
&\leq C_2(r, \varepsilon) \left(\eta^{1/2} + \exp(-c_2(r, \varepsilon)n) + \frac{C \ln^2 n}{\eta n}\right) \\
&\leq C_2(r, \varepsilon) \left(\eta^{1/2} + \frac{C \ln^2 n}{\eta n}\right).
\end{aligned}$$

**Step 4.** We have all the ingredients in order that the regularization Lemma 2.1 holds for  $V_n(Z)$  and  $V_n(G)$  and we can prove for both of them an estimate as in (2.9). By using it, we obtain, for  $\eta < 1$ ,

$$\begin{aligned}
d_0(V_n(Z), V_n(G)) &\leq C \left(\frac{1}{\eta^{3p_*/4}} d_3^{\frac{1}{4}}(V_n(Z), V_n(G)) + \mathbb{P}(\lambda_{V_n(Z)} < \eta) + \mathbb{P}(\lambda_{V_n(G)} < \eta)\right) \\
&\leq C \left(\frac{1}{\eta^{3p_*/4}} \left(\frac{\ln n}{\sqrt{n}}\right)^{1/4} + \eta^{1/2} + \frac{\ln^2 n}{\eta n}\right) \leq C \left(\frac{1}{\eta^{p_*}} \left(\frac{\ln n}{\sqrt{n}}\right)^{1/4} + \eta^{1/2}\right).
\end{aligned}$$

We optimize over  $\eta < 1$  and we obtain

$$d_0(V_n(Z), V_n(G)) \leq C \left(\frac{\ln n}{\sqrt{n}}\right)^{1/4(2p_*+1)}.$$

**Step 5.** Here we set  $G_k = n(W_{k/n} - W_{(k-1)/n})$ , where  $W_t$  denotes the Brownian motion on which  $I_2(\phi)$  is written. We estimate

$$\|V'_n(G) - I_2(\phi)\|_2^2 = \int_0^1 \int_0^1 |\psi_n(x, y) - \phi(x, y)|^2 dx dy,$$

where

$$\psi_n(x, y) = a \otimes_1 a(i, j) \quad \text{for } x \in I_i, y \in I_j.$$

By (B.3)

$$|\psi_n(x, y) - \phi(x, y)| \leq c \left(\frac{1}{\sqrt{n}\sqrt{|x-y|}} + \frac{1}{n(x+y)}\right)$$

so that

$$\|V'_n(G) - I_2(\phi)\|_2^2 \leq \frac{C}{n}.$$

Since  $\lim_n \|V''_n(G)\|_2 = 0$  we conclude that  $\lim_n V_n(G) = I_2(\phi)$  in  $L^2$ .

Let  $m \geq n$ . Using exactly the same argument as above we obtain, as  $\eta < 1$ ,

$$\begin{aligned}
d_0(V_n(G), V_m(G)) &\leq C \left(\frac{1}{\eta^{p_*}} d_1^{\frac{1}{2}}(V_n(G), V_m(G)) + \mathbb{P}(\lambda_{V_n(G)} < \eta) + \mathbb{P}(\lambda_{V_m(G)} < \eta)\right) \\
&\leq C \left(\frac{1}{\eta^{p_*}} \frac{1}{n^{1/2}} + \eta^{1/2} + \frac{\ln^2 n}{\eta n}\right) \leq C \left(\frac{1}{\eta^{p_*}} \frac{1}{n^{1/2}} + \eta^{1/2}\right).
\end{aligned}$$

We optimize for  $\eta < 1$  in order to obtain

$$d_0(V_n(G), V_m(G)) \leq \frac{C}{n^{1/2(2p_*+1)}}.$$

So  $V_n(G), n \in \mathbb{N}$  is a Cauchy sequence in  $d_0$  and consequently converges to some limit which has to be  $I_2(\phi)$ . And the estimate of the error is the one given above.  $\square$

## A An iterated Hoeffding's inequality

In this section we estimate  $\mathbb{P}(S_N(c^2, \chi) \leq x)$  with

$$S_N(c^2, \chi) = \sum_{m=1}^{m_0} \sum_{\alpha \in \Gamma_m} c^2(\alpha) \chi^\alpha.$$

Essentially this amounts to an iterated application of Hoeffding's inequality. In order to implement this strategy we will use an extension of Hoeffding's inequality to martingales, due to Benktus [6]. We recall that  $\|c\|_N^2 = \sum_{m=1}^N |c|_m^2 = \sum_{1 \leq |\alpha| \leq N} c^2(\alpha)$  and  $\bar{\delta}^2(c)$  is defined in (1.3).

**Lemma A.1** *Let  $p = \mathbb{P}(\chi_j = 1) = \varepsilon m(r)$ . If*

$$x \leq \left(\frac{p}{4}\right)^{2N} \|c\|_N^2 \tag{A.1}$$

Then

$$\mathbb{P}(S_N(c^2, \chi) \leq x) \leq \frac{2e^3}{9} N \exp\left(-\frac{x^2}{N \bar{\delta}^2(c) \|c\|_N^2}\right). \tag{A.2}$$

**Proof.** We proceed by recurrence on  $N$ . If  $N = 1$  we have

$$\begin{aligned} \mathbb{P}(S_N(c^2, \chi) \leq x) &= \mathbb{P}\left(\sum_j c^2(j) \chi_j \leq x\right) \\ &\leq \mathbb{P}\left(p \sum_j c^2(j) \leq 2x\right) + \mathbb{P}\left(\sum_j c^2(j)(p - \chi_j) \geq x\right). \end{aligned}$$

Since

$$\sum_j c^2(j) = \|c\|_1^2 \geq \left(\frac{4}{p}\right)^2 x > \frac{2x}{p}$$

the first term is zero (here comes on the hypothesis (A.1)). And by Hoeffding's inequality

$$\mathbb{P}\left(\sum_j c^2(j)(p - \chi_j) \geq x\right) \leq \exp\left(-\frac{2x^2}{\sum_j c^4(j)}\right).$$

Since

$$\sum_j c^4(j) \leq \max_j c^2(j) \times \sum_j c^2(j) = \bar{\delta}_1^2(c) \|c\|_1^2$$

our inequality is verified.

Suppose now that (A.2) holds for  $N - 1$  and let us prove it for  $N$ . We recall that  $\Gamma_m(j) = \{\alpha = (\alpha_1, \dots, \alpha_m) : \alpha_i \leq j\}$  and we denote  $\Gamma_m^o(j) = \{\alpha \in \Gamma_m(j) : \alpha_1 < \alpha_2 < \dots < \alpha_m\}$ . We also set  $\Gamma_m^o = \{\alpha \in \Gamma_m : \alpha_1 < \alpha_2 < \dots < \alpha_m\}$ . We write

$$\begin{aligned} S_N(c^2, \chi) &= \sum_{m=1}^N m! \sum_{\alpha \in \Gamma_m^o} c^2(\alpha) \chi^\alpha \\ &= \sum_{j=1}^{\infty} c^2(j) \chi_j + \sum_{m=2}^N m! \sum_{j=1}^{\infty} \chi_j \sum_{\alpha \in \Gamma_{m-1}^o(j-1)} c^2(\alpha, j) \chi^\alpha \\ &= \sum_{j=1}^{\infty} \chi_j (c^2(j) + H_j) \\ &= A + pB \end{aligned}$$

with

$$H_j = \sum_{m=2}^N m! \sum_{\alpha \in \Gamma_{m-1}^o(j-1)} c^2(\alpha, j) \chi^\alpha.$$

and

$$A = \sum_{j=1}^{\infty} (\chi_j - p)(c^2(j) + H_j), \quad B = \sum_{j=1}^{\infty} (c^2(j) + H_j).$$

We take  $x$  which satisfies (A.1) and we write

$$\mathbb{P}(S_N(c^2, \chi) \leq x) \leq \mathbb{P}(B \leq 2x/p) + \mathbb{P}(-A \geq x) =: b + a.$$

Let us estimate  $b$ . For  $\alpha = (\alpha_1, \dots, \alpha_m)$  we denote  $\bar{\alpha} = \max_{j=1, \dots, m} \alpha_j$  and

$$\bar{c}^2(\alpha) = m \sum_{j > \bar{\alpha}} c^2(\alpha, j)$$

and we write

$$\begin{aligned} \sum_{j=1}^{\infty} H_j &= \sum_{m=2}^N m! \sum_{\alpha \in \Gamma_{m-1}^o} \sum_{j > \bar{\alpha}} c^2(\alpha, j) \chi^\alpha = \sum_{m=2}^N \sum_{\alpha \in \Gamma_{m-1}} \sum_{j > \bar{\alpha}} m c^2(\alpha, j) \chi^\alpha \\ &= \sum_{m=2}^N \sum_{\alpha \in \Gamma_{m-1}} \bar{c}^2(\alpha) \chi^\alpha = S_{N-1}(\bar{c}^2, \chi). \end{aligned}$$

It follows that

$$B = \sum_{j=1}^{\infty} c^2(j) + S_{N-1}(\bar{c}^2, \chi).$$

**Case 1.** We suppose that

$$\sum_{j=1}^{\infty} c^2(j) \geq \frac{1}{2} \|c\|_N^2. \quad (\text{A.3})$$

By (A.1)

$$\frac{2}{p} x \leq \left(\frac{p}{2}\right)^{2N-1} \|c\|_N^2 < \frac{1}{2} \|c\|_N^2 \leq \sum_{j=1}^{\infty} c^2(j)$$

so that

$$b = \mathbb{P}\left(\sum_{j=1}^{\infty} c^2(j) + S_{N-1}(\bar{c}^2, \chi) \leq \frac{2}{p} x\right) = 0.$$

**Case 2.** We suppose that

$$\sum_{j=1}^{\infty} c^2(j) < \frac{1}{2} \|c\|_N^2. \quad (\text{A.4})$$

Then ignore  $\sum_{j=1}^{\infty} c^2(j)$  and we write

$$b \leq \mathbb{P}(S_{N-1}(\bar{c}^2, \chi) \leq \frac{2}{p} x).$$

We will use the recurrence hypothesis. Before doing this, we verify that

$$\bar{\delta}_{N-1}^2(\bar{c}) \leq \bar{\delta}_N^2(c) \quad \text{and} \quad \frac{1}{4} \|c\|_N^2 \leq \|\bar{c}\|_{N-1}^2 \leq \|c\|_N^2. \quad (\text{A.5})$$

Let  $m \geq 2$ . We have

$$\begin{aligned} \delta_{m-1}^2(\bar{c}) &= \max_j \sum_{\alpha \in \Gamma_{m-1}} \bar{c}^2(\alpha, j) = \max_j m \sum_{\alpha \in \Gamma_{m-1}} \sum_{i > j\sqrt{\alpha}} c^2(\alpha, j, i) \\ &\leq \max_j \sum_{\beta \in \Gamma_m} c^2(\beta, j) = \delta_m^2(c). \end{aligned}$$

Summing over  $m$  we obtain  $\bar{\delta}_{N-1}^2(\bar{c}) \leq \bar{\delta}_N^2(c)$ .

We write now

$$\begin{aligned} \|\bar{c}\|_{N-1}^2 &= \sum_{m=1}^{N-1} m! \sum_{\alpha \in \Gamma_m^o} \bar{c}^2(\alpha) = \sum_{m=1}^{N-1} m! \sum_{\alpha \in \Gamma_m^o} m \sum_{i > \alpha_m} c^2(\alpha, i) \\ &= \sum_{m=1}^{N-1} m! m \sum_{\beta \in \Gamma_{m+1}^o} c^2(\beta) = \sum_{m=1}^{N-1} \frac{m}{m+1} \sum_{\beta \in \Gamma_{m+1}} c^2(\beta) \leq \|c\|_N^2. \end{aligned}$$

And, since  $\frac{m}{m+1} \geq \frac{1}{2}$ , we use (A.4) and we obtain

$$\|\bar{c}\|_{N-1}^2 \geq \frac{1}{2} \sum_{m=1}^{N-1} \sum_{\beta \in \Gamma_{m+1}} c^2(\beta) = \frac{1}{2} (\|c\|_N^2 - \sum_{j=1}^{\infty} c^2(j)) \geq \frac{1}{4} \|c\|_N^2$$

so (A.5) is proved.

We have to verify that  $\bar{x} = \frac{2}{p}x$  verifies (A.1). Using (A.1) for  $x$  and (A.5) we obtain

$$\bar{x} \leq \frac{4}{p}x \leq \left(\frac{p}{4}\right)^{2N-1} \|c\|_N^2 \leq \left(\frac{p}{4}\right)^{2(N-1)} \frac{1}{4} \|c\|_N^2 \leq \left(\frac{p}{4}\right)^{2(N-1)} \|\bar{c}\|_{N-1}^2.$$

Now we may use (A.2) and (A.5) and we obtain (notice that  $x^2 \leq \bar{x}^2$ )

$$\begin{aligned} \mathbb{P}(S_{N-1}(\bar{c}^2, \chi) \leq \bar{x}) &\leq \frac{2e^3}{9} (N-1) \exp\left(-\frac{\bar{x}^2}{(N-1)\bar{\delta}_{N-1}^2(\bar{c}) \|\bar{c}\|_{N-1}^2}\right) \\ &\leq \frac{2e^3}{9} (N-1) \exp\left(-\frac{x^2}{N\bar{\delta}_N^2(c) \|c\|_N^2}\right). \end{aligned}$$

We conclude that in both Case 1 and Case 2 we have

$$b \leq \frac{2e^3}{9} (N-1) \exp\left(-\frac{x^2}{\varepsilon_N^2(c) \|c\|_N^2}\right). \quad (\text{A.6})$$

We estimate now  $a$ . We denote

$$h_j = c^2(j) + \sum_{m=2}^N m! \sum_{\alpha \in \Gamma_{m-1}^o(j-1)} c^2(\alpha, j).$$

Since  $0 \leq \chi^\alpha \leq 1$  we have

$$0 \leq c^2(j) + H_j \leq h_j.$$

Notice that

$$h_j = c^2(j) + \sum_{m=2}^N m \sum_{\alpha \in \Gamma_{m-1}(j-1)} c^2(\alpha, j) \leq N \bar{\delta}_N^2(c)$$

and

$$\begin{aligned} \sum_{j=1}^{\infty} h_j &= \sum_{j=1}^{\infty} c^2(j) + \sum_{m=2}^N m \sum_{j=1}^{\infty} \sum_{\alpha \in \Gamma_{m-1}(j-1)} c^2(\alpha, j) \\ &= \sum_{j=1}^{\infty} c^2(j) + \sum_{m=2}^N \sum_{\beta \in \Gamma_m} c^2(\beta) = \|c\|_N^2. \end{aligned}$$

In particular

$$\sum_{j=1}^{\infty} h_j^2 \leq N \bar{\delta}_N^2(c) \|c\|_N^2.$$

We use now Corollary 1.4 pg 1654 in Bentkus [Be] which asserts the following: if  $M_k, k \in \mathbb{N}$  is a martingale such that  $|M_k - M_{k-1}| \leq h_k$  almost surely, then, for every  $n \in \mathbb{N}$ ,

$$\mathbb{P}(M_n \geq x) \leq \frac{2e^3}{9} \exp\left(-\frac{x^2}{\sum_{j=1}^n h_j^2}\right).$$

In our case this gives

$$a = \mathbb{P}\left(\sum_{j=1}^{\infty} (p - \chi_j)(c^2(j) + H_j) \geq x\right) \leq \frac{2e^3}{9} \exp\left(-\frac{x^2}{N \bar{\delta}_N^2(c) \|c\|_N^2}\right).$$

This, together with (A.6) yields

$$a + b \leq \frac{2e^3}{9} N \exp\left(-\frac{x^2}{N \bar{\delta}_N^2(c) \|c\|_N^2}\right).$$

□

## B Computations around an integral

In this section we compute the following integral:

$$\phi(x, y) = \int_0^1 \theta_{x,y}(z) dz \quad \text{with} \quad \theta_{x,y}(z) = \frac{1}{\sqrt{|x-z||y-z|}}. \quad (\text{B.1})$$

We also discuss the approximation with Riemann sums. We fix  $n \in \mathbb{N}_*$  and we denote  $I_i = [\frac{i}{n}, \frac{i+1}{n})$  and  $x_i = \frac{i}{n}$ .

**Lemma B.1** For  $0 < x < y < 1$ , it holds

$$\phi(x, y) = \pi + 2 \ln \frac{\sqrt{1-x} + \sqrt{1-y}}{|\sqrt{x} - \sqrt{y}|}. \quad (\text{B.2})$$

Moreover, if  $x \in I_i$  and  $y \in I_j$  with  $i < j$  then

$$\left| \phi(x, y) - \frac{1}{n} \sum_{\substack{k=1 \\ k \neq i, k \neq j}}^n \theta_{x_i, x_j}(x_k) \right| \leq \frac{16\sqrt{2}}{\sqrt{n}} \frac{1}{\sqrt{y-x}} + \frac{8}{n(x+y)}. \quad (\text{B.3})$$

**Proof. Step 1.** We consider the decomposition

$$(z-x)(z-y) = z^2 - z(x+y) + xy = \left(z - \frac{x+y}{2}\right)^2 - \frac{(y-x)^2}{4}$$

and we write

$$\phi(x, y) = \int_{(0,x) \cup (y,1)} \frac{1}{\sqrt{\left(z - \frac{x+y}{2}\right)^2 - \left(\frac{y-x}{2}\right)^2}} dz + \int_{(x,y)} \frac{1}{\sqrt{\left(\frac{y-x}{2}\right)^2 - \left(z - \frac{x+y}{2}\right)^2}} dz$$

By using the change of variable  $t = z - \frac{x+y}{2}$  and the fact that

$$\begin{aligned} \int \frac{dt}{\sqrt{t^2 - a^2}} &= \ln \left| t + \sqrt{t^2 - a^2} \right| + C, \quad t^2 > a^2 \\ \int \frac{dt}{\sqrt{a^2 - t^2}} &= \arcsin \frac{t}{a} + C, \quad t^2 < a^2, \end{aligned}$$

straightforward computations give (B.2).

**Step 2.** We set

$$\begin{aligned} I_1(a) &= \int_a^x \theta_{x,y}(z) dz = \ln \frac{y-x}{2} - \ln \left| a - \frac{x+y}{2} + \sqrt{(a-x)(a-y)} \right|, \quad 0 < a < x, \\ I_2'(a) &= \int_x^a \theta_{x,y}(z) dz = \arcsin \frac{a - \frac{x+y}{2}}{\frac{y-x}{2}} + \frac{\pi}{2}, \quad x < a < y, \\ I_2''(a) &= \int_a^y \theta_{x,y}(z) dz = \frac{\pi}{2} - \arcsin \frac{a - \frac{x+y}{2}}{\frac{y-x}{2}}, \quad x < a < y, \\ I_3(a) &= \int_y^a \theta_{x,y}(z) dz = \ln \left| a - \frac{x+y}{2} + \sqrt{(a-x)(a-y)} \right| - \ln \frac{y-x}{2}, \quad y < a < 1. \end{aligned}$$

The above formulas in the last right hand sides follows by using the decomposition and the change of variable as in Step 1.

We first estimate  $I_i(a)$  for  $a$  close to  $x$  or to  $y$ . First we notice that for  $x - \frac{1}{n} < a < x < y$

$$I_1(a) = \int_a^x \frac{dz}{\sqrt{(x-z)(y-z)}} \leq \frac{1}{\sqrt{y-x}} \int_a^x \frac{dz}{\sqrt{x-z}} = \frac{2\sqrt{x-a}}{\sqrt{y-x}} \leq \frac{2}{\sqrt{n}} \frac{1}{\sqrt{y-x}} \quad (\text{B.4})$$



and for  $x < a < \frac{x+y}{2} \wedge (x + \frac{1}{n})$

$$I_2'(a) = \int_x^a \frac{dz}{\sqrt{(x-z)(y-z)}} \leq \frac{1}{\sqrt{y - \frac{x+y}{2}}} \int_x^a \frac{dz}{\sqrt{x-z}} = \frac{2\sqrt{2(x-a)}}{\sqrt{y-x}} \frac{1}{\sqrt{n}} \frac{2\sqrt{2}}{\sqrt{y-x}}. \quad (\text{B.5})$$

Similar estimates hold for  $I_2''(a)$  and for  $I_3(a)$ .

We are now ready to prove (B.3). We decompose

$$S = \frac{1}{n} \sum_{\substack{k=1 \\ k \neq i, k \neq j}}^n \theta_{x_i, x_j}(x_k) = S' + S'' + S'''$$

with

$$S' = \frac{1}{n} \sum_{k=1}^{i-1} \theta_{x_i, x_j}(x_k), \quad S'' = \frac{1}{n} \sum_{k=i+1}^{j-1} \theta_{x_i, x_j}(x_k), \quad S''' = \frac{1}{n} \sum_{k=j+1}^n \theta_{x_i, x_j}(x_k).$$

And we also decompose

$$I = \int_0^1 \theta_{x,y}(z) dz = I' + I'' + I'''$$

with

$$I' = \int_0^x \theta_{x,y}(z) dz, \quad I'' = \int_x^y \theta_{x,y}(z) dz, \quad I''' = \int_y^1 \theta_{x,y}(z) dz.$$

Let us estimate  $I' - S'$ . We have  $x_i \leq x < x_{i+1}$  and  $x_j \leq y < x_{j+1}$  so that

$$x_i - x_k \leq x - x_k \leq x_i - x_{k-1}, \quad x_j - x_k \leq y - x_k \leq x_j - x_{k-1}$$

so that

$$\frac{1}{\sqrt{|x_i - x_{k-1}| |x_j - x_{k-1}|}} \leq \frac{1}{\sqrt{|x - x_k| |y - x_k|}} \leq \frac{1}{\sqrt{|x_i - x_k| |x_j - x_k|}}.$$

Since  $z \mapsto \theta_{x,y}(z)$  is increasing for  $0 < z < x$  we have

$$\frac{1}{n} \sum_{k=0}^{i-1} \frac{1}{\sqrt{|x - x_k| |y - x_k|}} \leq I' \leq \frac{1}{n} \sum_{k=1}^{i-1} \frac{1}{\sqrt{|x - x_k| |y - x_k|}} + \int_{x_i}^x \theta_{x,y}(z) dz.$$

Combining this with the previous inequality one gets

$$\begin{aligned} \frac{1}{n} \sum_{k=0}^{i-2} \frac{1}{\sqrt{|x_i - x_k| |x_j - x_k|}} &\leq \frac{1}{n} \sum_{k=0}^{i-1} \frac{1}{\sqrt{|x - x_k| |y - x_k|}} \\ &\leq I' \leq \frac{1}{n} \sum_{k=1}^{i-1} \frac{1}{\sqrt{|x_i - x_k| |x_j - x_k|}} + \int_{x_i}^x \theta_{x,y}(z) dz \end{aligned}$$

One also has

$$\frac{1}{n} \frac{1}{\sqrt{|x_i - x_{i-1}| |x_j - x_{i-1}|}} \leq \frac{1}{\sqrt{n}} \frac{1}{\sqrt{y-x}}$$

so that finally we obtain

$$\frac{1}{n} \sum_{k=1}^{i-1} \frac{1}{\sqrt{|x_i - x_k| |x_j - x_k|}} - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{y-x}} \leq I' \leq \frac{1}{n} \sum_{k=1}^{i-1} \frac{1}{\sqrt{|x_i - x_k| |x_j - x_k|}} + \int_{x_i}^x \theta_{x,y}(z) dz$$

which, together with (B.4), yields

$$|I' - S'| \leq \frac{1}{\sqrt{n}} \frac{1}{\sqrt{y-x}} + \int_{x_i}^x \theta_{x,y}(z) dz \leq \frac{1}{\sqrt{n}} \frac{3}{\sqrt{y-x}}$$

In a similar way one checks that

$$|I''' - S'''| \leq \frac{1}{\sqrt{n}} \frac{3}{\sqrt{y-x}}.$$

In order to estimate  $|I'' - S''|$  we note that  $z \mapsto \theta_{x,y}(z)$  is increasing for  $x < z < \frac{x+y}{2}$  and decreasing for  $\frac{x+y}{2} < z < y$ . So using similar arguments we obtain, with  $x_l \leq \frac{x+y}{2} < x_{l+1}$ ,

$$\begin{aligned} |I'' - S''| &\leq \frac{4}{\sqrt{n}} \frac{1}{\sqrt{y-x}} + \int_x^{x_{i+1}} \theta_{x,y}(z) dz + \int_{x_j}^y \theta_{x,y}(z) dz + \int_{x_l}^{x_{l+1}} \theta_{x,y}(z) dz \\ &\leq \frac{10\sqrt{2}}{\sqrt{n}} \frac{1}{\sqrt{y-x}} + \int_{x_l}^{x_{l+1}} \theta_{x,y}(z) dz. \end{aligned}$$

It is easy to check that, if  $\frac{1}{n} \leq \frac{y-x}{4}$

$$\int_{\frac{x+y}{2} - \frac{1}{n}}^{\frac{x+y}{2} + \frac{1}{n}} \theta_{x,y}(z) dz \leq \frac{1}{n} \times \frac{8}{y+x}$$

And if  $\frac{1}{n} > \frac{y-x}{4}$  then  $I''$  does not appear, so the above integral does not exist. So

$$|I'' - S''| \leq \frac{10\sqrt{2}}{\sqrt{n}} \frac{1}{\sqrt{y-x}} + \frac{8}{n(x+y)}.$$

We put all these inequalities together and we obtain

$$|I - S| \leq \frac{16\sqrt{2}}{\sqrt{n}} \frac{1}{\sqrt{y-x}} + \frac{8}{n(x+y)}.$$

□

We will use Lemma B.1 in order to compute the following quantities which appear in our calculus. We denote

$$\begin{aligned} a(i, j) &= 1_{\{i \neq j\}} \frac{1}{\sqrt{|i-j|}}, \quad c_n(i, j) = \frac{1}{\sqrt{2n \ln n}} a(i, j) \\ \bar{c}_n(i, j) &= \frac{1}{n} \sum_{k=1}^n a(i, k) a(j, k) = (2 \ln n) \times (c_n \otimes_1 c_n)(i, j). \end{aligned}$$

We also recall that

$$\delta_2^2(c_n) = \max_i \sum_{j=1}^n c_n^2(i, j), \quad |c_n|_2^2 = \sum_{i,j=1}^n c_n^2(i, j).$$

**Lemma B.2 A.** *We have*

$$\delta_2^2(c_n) \leq \frac{2}{n} \quad (\text{B.6})$$

$$1 - \frac{1}{\ln n} \leq |c_n|_2^2 \leq 1 + \frac{1}{\ln n} \quad (\text{B.7})$$

**B.** *Let*

$$c_* = \frac{1}{16} \int_{\{|x-y| \geq \frac{1}{4}\}} \phi^2(x, y) dx dy > 0. \quad (\text{B.8})$$

*Then, for  $n \geq (\frac{32\sqrt{2}}{\pi})^2$  one has*

$$\delta_2^2(\bar{c}_n) \leq \frac{C \ln^2 n}{n}, \quad (\text{B.9})$$

$$c_* \leq |\bar{c}_n|_2^2 \leq C, \quad (\text{B.10})$$

$$\sum_{k=1}^n \bar{c}_n^2(k, k) \leq \frac{C \ln^2 n}{n}. \quad (\text{B.11})$$

*where  $C$  is a universal constant.*

**Proof.** We will first check that

$$\ln i + \ln(n - i) \leq \sum_{j=1}^n a^2(i, j) \leq 2 + \ln i + \ln(n - i). \quad (\text{B.12})$$

Let us denote  $x_i = \frac{i}{n}$  so that

$$a(i, j) = \frac{1}{\sqrt{|i-j|}} = \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{|x_i - x_j|}}$$

and then

$$\begin{aligned} \ln(n - i) &= \int_{\frac{1}{n} + x_i}^1 \frac{dy}{y - x_i} \leq \sum_{j=i+1}^{n-1} \frac{1}{x_j - x_i} \times \frac{1}{n} = \sum_{j=i+1}^n a^2(i, j) \\ &\leq 1 + \int_{\frac{1}{n} + x_i}^1 \frac{dy}{y - x_i} = 1 + \ln(n - i). \end{aligned}$$

and

$$\begin{aligned} \ln i &= \int_0^{x_i - 1/n} \frac{dy}{y - x_i} \leq \sum_{j=0}^{i-1} \frac{1}{x_j - x_i} \times \frac{1}{n} = \sum_{j=0}^{i-1} a^2(i, j) \\ &\leq 1 + \int_0^{x_i - 1/n} \frac{dy}{y - x_i} = 1 + \ln i. \end{aligned}$$

Summing these two inequalities we obtain (B.12).

Since  $\int_0^1 \ln x dx = -1$  we have

$$n(\ln n - 1) \leq \sum_{i=1}^n \ln i = n(\ln n + \frac{1}{n} \sum_{i=1}^n \ln \frac{i}{n}) \leq n \ln n$$

so that summing over  $i$  in (B.12) we obtain

$$2n(\ln n - 1) \leq \sum_{i=1}^n \sum_{j=1}^n a^2(i, j) \leq 2n + 2n \ln n$$

which gives (B.7). And by (B.12)

$$\delta_2^2(c_n) = \max_i \sum_{j=1}^n c_n^2(i, j) \leq \frac{2(1 + \ln n)}{2n \ln n} \leq \frac{2}{n}.$$

so (B.6) is also proved.

We will nw check that

$$\sum_{i, j=1}^n (c_n \otimes_1 c_n)^2(i, j) \leq \frac{C}{\ln^2 n}. \quad (\text{B.13})$$

We construct the function

$$\psi_n(x, y) = (a \otimes_1 a)(i, j) \quad \text{for } x \in I_i, y \in I_j$$

so that

$$(c_n \otimes_1 c_n)^2(i, j) = \frac{1}{4n^2 \ln^2 n} (a \otimes_1 a)^2(i, j) = \frac{1}{4 \ln^2 n} \int_{I_i \times I_j} \psi_n^2(x, y) dx dy.$$

Recall the function  $\phi$  defined (B.1). Using (B.3)

$$\begin{aligned} \sum_{|i-j| \geq 2} \int_{I_i \times I_j} \psi_n^2(x, y) dx dy &\leq 2 \sum_{|i-j| \geq 2} \int_{I_i \times I_j} \phi^2(x, y) dx dy \\ &\quad + 2 \sum_{|i-j| \geq 2} \int_{I_i \times I_j} |\psi_n(x, y) - \phi(x, y)|^2 dx dy \\ &\leq 2 \int \phi^2(x, y) dx dy + C \int_{\{|x-y| \geq \frac{1}{n}\}} \frac{1}{n|y-x|} + \frac{1}{n^2(x+y)^2} dx dy \leq C. \end{aligned}$$

So

$$\sum_{|i-j| \geq 2} (c_n \otimes_1 c_n)^2(i, j) \leq \frac{C}{\ln^2 n}. \quad (\text{B.14})$$

And, for  $j \in \{i-1, i, i+1\}$

$$\begin{aligned} \sum_{i=1}^n (c_n \otimes_1 c_n)^2(i, j) &= \sum_{i=1}^n \left( \sum_{k=1}^n c_n(k, i) c_n(k, j) \right)^2 \\ &\leq \sum_{i=1}^n \left( \sum_{k=1}^n c_n^2(k, i) \right) \left( \sum_{k=1}^n c_n^2(k, j) \right) \\ &\leq \delta^2(c_n) |c_n|_2^2 \leq \frac{4}{n}. \end{aligned}$$

So (B.13) is proved.

Let us now prove that

$$\frac{c_*}{4 \ln^2 n} \leq \sum_{i,j=1}^n (c_n \otimes_1 c_n)^2(i, j) \quad (\text{B.15})$$

Using (B.3) and (B.2)

$$\begin{aligned} \psi_n(x, y) &\geq \phi(x, y) - |\phi(x, y) - \psi_n(x, y)| \\ &\geq \frac{1}{2}\phi(x, y) + \frac{\pi}{2} - \frac{16\sqrt{2}}{\sqrt{n}\sqrt{|x-y|}} - \frac{8}{n(x+y)} \end{aligned}$$

Notice that, if  $|x-y| \geq \frac{1}{4}$  then  $x+y \geq \frac{1}{4}$ . Then, if  $\sqrt{n} \geq \frac{256\sqrt{2}}{\pi}$  we have

$$\frac{16\sqrt{2}}{\sqrt{n}\sqrt{|x-y|}} \leq \frac{64\sqrt{2}}{\sqrt{n}} \leq \frac{\pi}{4}, \quad \text{and} \quad \frac{8}{n(x+y)} \leq \frac{32}{n} \leq \frac{\pi}{4}$$

so that  $\psi_n(x, y) \geq \frac{1}{2}\phi(x, y)$ . It follows that

$$\begin{aligned} \sum_{i,j=1}^n (c_n \otimes_1 c_n)^2(i, j) &= \frac{1}{4 \ln^2 n} \int \psi_n^2(x, y) dx dy \\ &\geq \frac{1}{16 \ln^2 n} \int_{\{|x-y| \geq \frac{1}{4}\}} \phi^2(x, y) dx dy = \frac{c_*}{\ln^2 n}. \end{aligned}$$

So (B.15) is proved. And (B.10) follows from (B.14) and (B.15).

Let us prove (B.9). We fix  $i$  and we write

$$\begin{aligned} \sum_{j>i} (c_n \otimes_1 c_n)^2(i, j) &\leq 2 \sum_{j>i} \int_{I_i \times I_j} \phi^2(x, y) dx dy + 2 \sum_{j>i} \int_{I_i \times I_j} |\psi_n(x, y) - \phi(x, y)|^2 dx dy \\ &=: A + B. \end{aligned}$$

We have

$$A \leq 2 \int_{I_i \times I_{i+1}} \phi^2(x, y) dx dy + 2 \sum_{j>i+1} \int_{I_i \times I_j} \phi^2(x, y) dx dy =: A' + A''.$$

Using (B.2),

$$\begin{aligned} A' &\leq \frac{C}{n^2} + C \int_{I_i \times I_{i+1}} \ln^2 \frac{1}{|x-y|} dx dy \leq \frac{C \ln^2 n}{n^2}, \\ A'' &\leq C \int_{x_i}^{x_{i+1}} \int_{x_{i+1} + \frac{1}{n}}^1 \phi^2(x, y) dx dy \leq \frac{C \ln^2 n}{n}. \end{aligned}$$

Using the estimate (B.3) we get similar estimates for  $B$ . And this gives (B.9).

We prove now (B.11). Using (B.6) and (B.7)

$$\sum_{k=1}^n \tilde{c}_n^2(k, k) = \sum_{k=1}^n \left( \frac{1}{n} \sum_{i=1}^n a^2(i, k) \right)^2 \leq \max_i \frac{1}{n} \sum_{i=1}^n a^2(i, k) \times \frac{1}{n} \sum_{k,i=1}^n a^2(i, k) \leq C \frac{\ln^2 n}{n}.$$

□

## References

- [1] V. Bally, L. Caramellino. Asymptotic development for the CLT in total variation distance. *Bernoulli* **22**, 2442–2485, 2016.
- [2] V. Bally, L. Caramellino. On the distances between probability density functions. *Electronic Journal of Probability* **19**, no. 110, 1–33, 2014.
- [3] V. Bally, L. Caramellino. An invariance principle for stochastic series I. Gaussian limits. [ArXiv:1510.03616](https://arxiv.org/abs/1510.03616), 2015.
- [4] V. Bally, E. Clément. Integration by parts formula and applications to equations with jumps. *Probab. Theory Related Fields*, **151**, 613–657, 2011.
- [5] D. Bakry, I. Gentil, M. Ledoux *Analysis and Geometry of Markov Diffusion Semigroups*. Springer, 2014.
- [6] V. Bentkus. On Hoeffding’s inequalities. *Ann. Probab.* **32**, 1650–1673, 2004.
- [7] K. Bichtler, J.-B. Gravereaux, J. Jacod. *Malliavin calculus for processes with jumps*. Gordon and Breach Science Publishers, 1987.
- [8] A. Carbery, J. Wright. Distributional and  $L^q$  norm inequalities for polynomials over convex bodies in  $\mathbb{R}^n$ . *Math. Research Lett.* **8**, 233–248, 2001.
- [9] P. de Jong. A central limit theorem for generalized quadratic forms. *Probab. Th. Rel. Fields* **75**, 261–277, 1987.
- [10] P. de Jong. A central limit theorem for generalized multilinear forms. *Journal of Multivariate Analysis* **34**, 275–289, 1990.
- [11] R.A. Fisher. Moments and product moments of sampling distributions. *Proceedings of the London Mathematical Society* **2**, 199–238, 1929.
- [12] W. Hoeffding. A class of statistics with asymptotically normal distributions. *Ann. Statistics* **19**, 293–325, 1948.
- [13] N. Ikeda, S. Watanabe. *Stochastic Differential Equations and Diffusion processes*. North-Holland Mathematical Library 24, 1989.
- [14] R. Latala. Estimates of moments and tails of Gaussian chaoses. *Ann. Probab.* **34**, 2315–2331, 2006.
- [15] A.J. Lee. *U-Statistics: Theory and Practice*. Marcel Dekker, New York, 1990.
- [16] D. Malicet, G. Poly. Properties of convergence in Dirichlet structures. *J. Funct. Anal.* **264**, 2077–2096, 2013.
- [17] E. Mossel, R. O’Donnell, K. Oleszkiewicz. Noise stability of functions with low influences: Variance and optimality. *Ann. Math.* **171**, 295–341, 2010.
- [18] S. Noredine, I. Nourdin. On the Gaussian approximation of vector-valued multiple integrals. *J. Multiv. Anal.* **102**, 1008-1017, 2011.

- [19] I. Nourdin, G. Peccati. *Normal Approximations Using Malliavin Calculus: from Stein's Method to Universality*. Cambridge Tracts in Mathematics, 192, 2012.
- [20] I. Nourdin, G. Peccati. Stein's method on Wiener chaos. *Probab. Theory Related Fields* **145**, 75–118, 2009.
- [21] I. Nourdin, G. Peccati, G. Reinert. Invariance principles for homogeneous sums: universality of Wiener chaos. *Ann. Probab.* **38**, 1947–1985, 2010.
- [22] I. Nourdin, G. Peccati, A. Réveillac. Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Inst. H. Poincaré Probab. Statist.* **46**, no. 1, 45–58, 2010.
- [23] I. Nourdin, G.Poly. Convergence in total variation on Wiener chaos. *Stochastic Process. Appl.* **123**, 651–674, 2013.
- [24] I. Nourdin, G.Poly. Convergence in law in the second Wiener/Wigner chaos. Convergence in law in the second Wiener/Wigner chaos. *Elect. Comm. in Probab.* **17**, no. 36, 2012.
- [25] D. Nualart. *The Malliavin calculus and related topics. Second Edition*. Springer-Verlag, 2006.
- [26] D. Nualart, S. Ortiz-Latorre. Central limit theorem for multiple stochastic integrals and Malliavin calculus. *Stoch. Processes Appl.* **118**, 614–628, 2008.
- [27] D. Nualart, G. Peccati. Central limit theorems for sequences of multiple stochastic integrals. *Annals of Probability* **33**, 177–193, 2005.
- [28] G. Peccati, C.A. Tudor. Gaussian limits for vector-valued multiple stochastic integrals. *Séminaire de Probabilités XXXVIII*, 247–262, 2004.
- [29] Yu.V. Prohorov. A local theorem for densities. *Doklady Akad. Nauk SSSR (N.S.)* **83**, 797–800, 1952.
- [30] B.A Sevastianov. The class of limit laws for distributions of quadratic forms in normal variables. *Theor. Probability Appl.* **6**, 368—372, 1961.