

March 2021

## Noncentral Limit Theorem for Large Wishart Matrices with Hermite Entries

Charles-Philippe Diez

CNRS, Université de Lille, Laboratoire Paul Painlevé UMR 8524, F-59655 Villeneuve d'Ascq, France,  
chphdiez@gmail.com

Ciprian A. Tudor

CNRS, Université de Lille, Laboratoire Paul Painlevé UMR 8524, F-59655 Villeneuve d'Ascq, France,  
ciprian.tudor@univ-lille.fr

Follow this and additional works at: <https://digitalcommons.lsu.edu/josa>



Part of the [Analysis Commons](#), and the [Other Mathematics Commons](#)

---

### Recommended Citation

Diez, Charles-Philippe and Tudor, Ciprian A. (2021) "Noncentral Limit Theorem for Large Wishart Matrices with Hermite Entries," *Journal of Stochastic Analysis*: Vol. 2 : No. 1 , Article 2.

DOI: 10.31390/josa.2.1.02

Available at: <https://digitalcommons.lsu.edu/josa/vol2/iss1/2>

## NONCENTRAL LIMIT THEOREM FOR LARGE WISHART MATRICES WITH HERMITE ENTRIES

CHARLES-PHILIPPE DIEZ AND CIPRIAN A. TUDOR\*

ABSTRACT. We analyze the limit behavior of the Wishart matrix  $\mathcal{W}_{n,d} = \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T$  constructed from an  $n \times d$  random matrix  $\mathcal{X}_{n,d}$  whose entries are given by the increments of the Hermite process. These entries are correlated on the same row, independent from one row to another and their probability distribution is different on different rows. We prove that the Wishart matrix converges in law, as  $d \rightarrow \infty$ , to a diagonal random matrix whose diagonal elements are random variables in the second Wiener chaos. We also estimate the Wasserstein distance associated to this convergence.

### 1. introduction

The Wishart matrices constitutes a class of random matrices with various applications. Given a  $n \times d$  random matrix  $\mathcal{X}_{n,d}$ , its associated Wishart matrix  $\mathcal{W}_{n,d}$  can be defined as  $\mathcal{W}_{n,d} = \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T$ , where “ $T$ ” denotes the transpose matrix. More complete descriptions of the Wishart matrices and of their applications in practice can be found e.g. in the surveys [1] or [7].

Of particular interest is the behavior of the Wishart matrices for large set data, i.e. when  $d \rightarrow \infty$  (“one dimensional regime”) or when both  $n, d \rightarrow \infty$  (“high-dimensional regime”). Different approaches have been used (eigenvalues analysis, empirical distribution, Wasserstein distance etc) in order to understand the behavior of such matrices when the dimensions  $n, d$  go the infinity.

The limit behavior of the Wishart matrix  $\mathcal{W}_{n,d}$  obviously depends on the probability distribution of the entries of the starting matrix  $\mathcal{X}_{n,d}$ . In their majority, the references related to this subject assume the independence (and many times the same distribution) for the entries of  $\mathcal{X}_{n,d}$  (see e.g. [3], [4], [6], [11]). Relaxing the independence hypothesis constitutes a topic of interest with important potential for applications (see e.g. the discussion in [4]). Such a step has been done in the work [8], where the authors consider a starting matrix with correlated Gaussian elements on the rows, the correlation being given by the fractional Brownian noise. A further step has been made in [2], where the same correlation on rows has been considered, but the law of the entries of the matrix  $\mathcal{X}_{n,d}$  are not Gaussian anymore, being random variables in the second Wiener chaos (actually, they are increments

---

Received 2020-9-3; Accepted 2020-10-24; Communicated by the editors.

2010 *Mathematics Subject Classification*. Primary 60H07; Secondary 60B20, 60F05, 60G22.

*Key words and phrases*. Wishart matrix, multiple stochastic integrals, Hermite process, Rosenblatt process, fractional Brownian motion, high-dimensional regime.

\* Corresponding author.

of the so-called Rosenblatt process). Both references [8] and [2] use the techniques of Malliavin calculus or the stochastic analysis on Wiener chaos in order to prove the main results.

Our work continues the line of research in [8] and [2]. The novelty is that we begin with a random matrix whose entries are elements in a Wiener chaos of arbitrary order and we also allow correlation for the entries on the same row of the matrix. The entries on different lines are independent but not necessarily with the same distribution. More precisely, the entries on the  $i$ th line are assumed to be given by the increments of the Hermite process of order  $q_i$  with  $q_i \geq 2$  and with  $q_i$  possibly different of  $q_j$  if  $i \neq j$ . Recall that the Hermite process is a non-Gaussian self-similar process. It also has stationary increments and it lives in the  $q$ th Wiener chaos. For  $q = 1$  it coincides with the fractional Brownian motion while for  $q = 2$  it is known as the Rosenblatt process. In this sense, we extend the approach in [2] (where  $q_i = 2$  for every  $i = 1, \dots, n$ ).

We show that the limit distribution as  $d \rightarrow \infty$  of the (properly normalized) Wishart matrix constructed from a matrix with Hermite entries in the  $q$ th Wiener chaos converges (componentwise) in distribution to a  $n \times n$  diagonal random matrix, whose diagonal elements are Rosenblatt random variables (i.e. the value at time 1 of a Rosenblatt process) living in a second Wiener chaos or they vanish. We also measure the Wasserstein distance between the renormalized Wishart matrix and its limiting random matrix. Our proofs rely on the properties of the random variables in Wiener chaos and in particular on the behavior of the quadratic and cross variation of the Hermite process.

We organized our work as follows. In Section 2 we recall the basic facts related to the Hermite processes and to the Wasserstein distance. In Section 3 we introduce the random matrix  $\mathcal{X}_{n,d}$  and we analyze the limit behavior in distribution of its associated Wishart matrix. In Section 4, we estimate the Wasserstein distance which corresponds to this convergence in law. The last section (Section 5) is the Appendix which contains the definition and the properties of the multiple stochastic integrals.

## 2. Preliminaries

In this preliminary part, we recall the definition and the main properties of the Hermite processes and of the Wasserstein distance between two random matrices and between two random vectors.

**2.1. The Hermite process.** We will denote by  $(Z_t^{(q,H)})_{t \geq 0}$  the Hermite process of order  $q \geq 1$  and with self-similarity index  $H \in (\frac{1}{2}, 1)$ . It is a self-similar process, it has stationary increments and long memory and its trajectories are Hölder continuous of order  $\delta$  for every  $\delta \in (0, H)$ . For every  $t > 0$ , the random variable  $Z_t^{(q,H)}$  is an element of the  $q$ th Wiener chaos and consequently it can be represented as a multiple stochastic integral of order  $q$  with respect to the Brownian motion. There exist several such representations of the Hermite processes, see e.g. [10] or [12]. Here we made the choice to work with its representation on a finite interval, that is, for every  $t \geq 0$ ,

$$Z_t^{(q,H)} = d(q, H) \int_{[0,t]^q} \left( \int_{y_1 \vee \dots \vee y_q}^t \partial_1 K^{H'}(u, y_1) \dots \partial_1 K^{H'}(u, y_q) du \right) dB(y_1) \dots dB(y_q) \quad (2.1)$$

where  $(B(y))_{y \geq 0}$  is a Wiener process. We used the notation

$$H' = 1 + \frac{H-1}{q} \text{ and } d(H, q) = \frac{(H(2H-1))^{1/2}}{(q!(H'(2H'-1))^q)^{1/2}} \quad (2.2)$$

and for  $H \in (\frac{1}{2}, 1)$  and  $t > s$ ,

$$K^H(t, s) = c(H) s^{\frac{1}{2}-H} \int_s^t (u-s)^{H-\frac{3}{2}} u^{H-\frac{1}{2}} du \quad (2.3)$$

with  $c(H) = \left( \frac{H(2H-1)}{\beta(2-2H, H-\frac{1}{2})} \right)^{\frac{1}{2}}$  and with  $\beta$  standing for the beta function (see e.g. [9]). The notation  $\partial_1 K^H(t, s)$  means the derivative of the function  $K^H$  with respect to its first variable and from (2.3), we have

$$\partial_1 K^H(t, s) = c(H) \left( H - \frac{1}{2} \right) \left( \frac{s}{t} \right)^{\frac{1}{2}-H} (t-s)^{H-\frac{3}{2}}. \quad (2.4)$$

We can also write  $Z^{(q,H)}$  as

$$Z_t^{(q,H)} = I_q(L_t)$$

where  $I_q$  denotes the multiple Wiener-Itô integral with respect to the (two-sided) Brownian motion  $B$  and the kernel  $L$  of the Hermite process is given by

$$L_t(y_1, y_2, \dots, y_q) = d(q, H) 1_{[0,t] \times \dots \times [0,t]}(y_1, y_2, \dots, y_q) \int_{y_1 \vee \dots \vee y_q}^t \partial_1 K^{H'}(u, y_1) \dots \partial_1 K^{H'}(u, y_q) du \quad (2.5)$$

for every  $y_1, y_2, \dots, y_q \in \mathbb{R}$  and  $t \geq 0$ .

The covariance of the Hermite process is the same for all  $q \geq 1$  and it is given by

$$\mathbf{E} Z_t^{(q,H)} Z_s^{(q,H)} = \frac{1}{2} (t^{2H} + s^{2H} - |t-s|^{2H}) \text{ for all } s, t \geq 0. \quad (2.6)$$

When  $q = 1$ , the Hermite process  $Z^{(1,H)}$  is nothing else than the fractional Brownian motion with Hurst parameter  $H \in (\frac{1}{2}, 1)$  and in this case the covariance (2.6) determines its law. This is not the case for  $q \geq 2$ , the Hermite process being not Gaussian in this situation. For  $q = 2$ , the process  $Z^{(2,H)}$  is known in the literature as the Rosenblatt process.

A random variable is called a Hermite random variable of order  $q$  if its law coincides with the law of  $Z_1^{(q,H)}$ . If  $q = 2$ , we will call such a random variable a Rosenblatt random variable.

Let  $(Z_j^{(q,H)} - Z_{j-1}^{(q,H)}, j \geq 1)$  be the noise generated by the Hermite process. From the covariance formula (2.6) we deduce that this sequence is stationary and

its correlation is given by, for  $j, k \geq 1$ ,

$$\mathbf{E} \left( Z_j^{(q,H)} - Z_{j-1}^{(q,H)} \right) \left( Z_k^{(q,H)} - Z_{k-1}^{(q,H)} \right) = \rho_H(j - k)$$

where for every  $v \in \mathbb{Z}$ ,

$$\rho_H(v) = \frac{1}{2} (|v + 1|^{2H} + |v - 1|^{2H} - 2|v|^{2H}). \quad (2.7)$$

In particular, the function  $\rho_H$  satisfies  $\rho_H(0) = 1$  and  $\rho_H(-v) = \rho_H(v)$  for every  $v \in \mathbb{Z}$ .

**2.2. Wasserstein distance between random matrices and vectors.** Let  $\mathcal{X}, \mathcal{Y}$  be two random matrices with values in  $\mathcal{M}_n(\mathbb{R})$ ,  $n \geq 1$ . We will denote by  $d_W$  the Wasserstein distance between the probability distributions of  $\mathcal{X}$  and  $\mathcal{Y}$ . That is,

$$d_W(\mathcal{X}, \mathcal{Y}) = \sup_{\|g\|_{\text{Lip}} \leq 1} |\mathbf{E}(g(\mathcal{X})) - \mathbf{E}(g(\mathcal{Y}))|,$$

where the Lipschitz norm  $\|\cdot\|_{\text{Lip}}$  of  $g: \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$  is defined by

$$\|g\|_{\text{Lip}} = \sup_{A \neq B, A, B \in \mathcal{M}_n(\mathbb{R})} \frac{|g(A) - g(B)|}{\|A - B\|_{\text{HS}}},$$

with  $\|\cdot\|_{\text{HS}}$  denoting the Hilbert-Schmidt norm on  $\mathcal{M}_n(\mathbb{R})$ .

In our work, the distances between random matrices will be evaluated via the distances between their associated random vectors. We recall that if  $X, Y$  are two  $n$ -dimensional random vectors, then the Wasserstein distance between them is defined to be

$$d_W(X, Y) = \sup_{\|g\|_{\text{Lip}} \leq 1} |\mathbf{E}(g(X)) - \mathbf{E}(g(Y))|, \quad (2.8)$$

where the Lipschitz norm  $\|\cdot\|_{\text{Lip}}$  of  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  is defined by

$$\|g\|_{\text{Lip}} = \sup_{x \neq y \in \mathbb{R}^n} \frac{|g(x) - g(y)|}{\|x - y\|_{\mathbb{R}^n}},$$

with  $\|\cdot\|_{\mathbb{R}^n}$  denoting the Euclidean norm on  $\mathbb{R}^n$ .

If  $\mathcal{X} = (X_{ij})_{1 \leq i, j \leq n}$  is an  $n \times n$  symmetric random matrix, we associate to it its ‘‘half-vector’’ defined to be the  $n(n + 1)/2$ -dimensional random vector

$$\mathcal{X}^{\text{half}} = (X_{11}, X_{12}, \dots, X_{1n}, X_{22}, X_{23}, \dots, X_{2n}, \dots, X_{nn}). \quad (2.9)$$

The following result has been proven in [8]. It gives the link between the distance of two symmetric random matrices and their associated half vectors: if  $\mathcal{X}, \mathcal{Y}$  are two symmetric random matrices with values in  $\mathcal{M}_n(\mathbb{R})$  then

$$d_W(\mathcal{X}, \mathcal{Y}) \leq \sqrt{2} d_W(\mathcal{X}^{\text{half}}, \mathcal{Y}^{\text{half}}),$$

where  $\mathcal{X}^{\text{half}}, \mathcal{Y}^{\text{half}}$  are the associated half-vectors defined in (2.9).

### 3. The Wishart Matrix with Hermite Entries

Let  $B = (B^{(1)}, \dots, B^{(n)})$  be a  $n$ -dimensional Brownian motion. We will denote by  $I_q^{(i)}$  the multiple stochastic integral of order  $q \geq 1$  with respect to the Wiener process  $B^{(i)}$  for every  $i = 1, \dots, n$ , see the Appendix.

We fix in the sequel  $n$  integer numbers  $q_1, \dots, q_n \geq 2$ . We let, for  $i = 1, \dots, n$  and for  $t \geq 0$ ,

$$Z^{(q_i, H, i)} = I_{q_i}^{(i)}(L_t)$$

with  $L_t$  given by (2.5). That means that  $Z^{(q_i, H, i)}$ ,  $1 \leq i \leq n$  are  $n$  independent Hermite processes, all of them with self-similarity index  $H \in (\frac{1}{2}, 1)$  and of order  $q_i \geq 2$  respectively.

We will start with a  $n \times d$  random matrix  $\mathcal{X} = (X_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d}$  whose entries are given by

$$X_{i,j} = Z_j^{(q_i, H, i)} - Z_{j-1}^{(q_i, H, i)} = I_{q_i}^{(i)}(L_j - L_{j-1}) \text{ for } 1 \leq i \leq n, 1 \leq j \leq d. \quad (3.1)$$

The elements of the matrix  $\mathcal{X}_{n,d}$  satisfy the following properties:

- all the entries on the line  $i$  have the same probability distribution, which coincides with the law of a Hermite random variable  $Z_1^{(q_i, H)}$ . In particular, for every  $1 \leq i \leq n, 1 \leq j \leq d$

$$\mathbf{E}X_{i,j}^2 = q_i! \|L_j - L_{j-1}\|_{L^2(\mathbb{R}^2)}^2 = 1.$$

- the elements on the same row are correlated and this correlation is given by the correlation of the Hermite noise. That is, for  $1 \leq i \leq n$  and  $1 \leq j, k \leq d$ ,

$$\mathbf{E}X_{i,j}X_{i,k} = \mathbf{E} \left( Z_j^{(q_i, H, i)} - Z_{j-1}^{(q_i, H, i)} \right) \left( Z_k^{(q_i, H, i)} - Z_{k-1}^{(q_i, H, i)} \right) = \rho_H(j - k)$$

with  $\rho_H$  given by (2.7).

- the elements on different rows are independent random variables. The distribution of the random variable on different lines is not the same, it depends on the order  $q_i$  of the corresponding Hermite process. This is more general than the assumption in [2], where all the elements of the matrix  $\mathcal{X}_{n,d}$  were identically distributed.

We define the (centered) Wishart matrix  $\mathcal{W}_{n,d}$  associated with  $\mathcal{X}_{n,d}$  by

$$\mathcal{W}_{n,d}^o = \frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - \mathcal{I}_n \quad (3.2)$$

where  $\mathcal{I}_n$  stands for the  $n \times n$  identity matrix. The matrix  $\mathcal{W}_{n,d}^o = (W_{i,j})_{1 \leq i, j \leq n}$  is a symmetric random matrix with values in  $\mathcal{M}_n(\mathbb{R})$  and its entries are given by

$$W_{i,i} = \frac{1}{d} \sum_{k=1}^d (X_{i,k}^2 - 1) = \frac{1}{d} \sum_{k=1}^d \left[ \left( Z_k^{(q_i, H, i)} - Z_{k-1}^{(q_i, H, i)} \right)^2 - 1 \right] \quad (3.3)$$

and for  $i \neq j$

$$W_{i,j} = \frac{1}{d} \sum_{k=1}^d X_{i,k} X_{j,k} = \frac{1}{d} \sum_{k=1}^d \left( Z_k^{(q_i, H, i)} - Z_{k-1}^{(q_i, H, i)} \right) \left( Z_k^{(q_j, H, j)} - Z_{k-1}^{(q_j, H, j)} \right). \quad (3.4)$$

The limit behavior as  $d \rightarrow \infty$  of the elements of the Wishart matrix  $\mathcal{W}_{n,d}$  is related to the behavior of the quadratic variations of the Hermite process. Actually, due to the self-similarity property of the Hermite process, for every  $1 \leq i, j \leq n$ , the random variables  $W_{i,i}$  and  $W_{i,j}$  with  $i \neq j$  have respectively the same distribution as

$$A_{i,i} = \frac{1}{d} \sum_{k=1}^d \left[ \frac{\left( Z_{\frac{k}{d}}^{(q_i, H, i)} - Z_{\frac{k-1}{d}}^{(q_i, H, i)} \right)^2}{d^{-2H}} - 1 \right] \quad (3.5)$$

and

$$A_{i,j} = d^{2H-1} \sum_{k=1}^d \left( Z_{\frac{k}{d}}^{(q_i, H, i)} - Z_{\frac{k-1}{d}}^{(q_i, H, i)} \right) \left( Z_{\frac{k}{d}}^{(q_j, H, j)} - Z_{\frac{k-1}{d}}^{(q_j, H, j)} \right). \quad (3.6)$$

Let us recall the following result from [5] which gives the behavior of the quadratic variations of the Hermite process and which plays a key role in the sequel.

**Theorem 3.1.** *Fix  $H \in (\frac{1}{2}, 1)$  and an integer  $q \geq 2$ . Let  $(Z_t^{(q, H)})_{t \geq 0}$  be a Hermite process given by (2.1). Define, for  $d \geq 1$ ,*

$$V_d = c_{q,H} d^{\frac{2-2H}{q}-1} \sum_{k=0}^{d-1} \left[ \frac{\left( Z_{\frac{k+1}{d}}^{(q, H)} - Z_{\frac{k}{d}}^{(q, H)} \right)^2}{d^{-2H}} - 1 \right]. \quad (3.7)$$

Then  $V_d$  converges in  $L^2(\Omega)$ , as  $d \rightarrow \infty$ , to a Rosenblatt random variable  $Z_1^{(2, 2H'-1)}$  with  $H'$  given by (2.2). The explicit expression of the constant  $c_{q,H}$  can be found in [5].

Let

$$q_0 := \min\{q_1, \dots, q_n\}. \quad (3.8)$$

We define the renormalized Wishart matrix  $\widetilde{\mathcal{W}}_{n,d} = (\widetilde{W}_{i,j})_{1 \leq i, j \leq n}$  by

$$\widetilde{W}_{i,j} = c_{q_0, H} d^{\frac{2-2H}{q_0}} W_{i,j} \text{ for } 1 \leq i, j \leq n \quad (3.9)$$

with  $c_{q_0, H}$  appearing in (3.7).

Theorem 3.1 immediately gives the limit behavior of the diagonal terms of the matrix  $\widetilde{\mathcal{W}}_{n,d}$ . We denote by " $\rightarrow^{(L)}$ " the convergence in law.

**Corollary 3.2.** *Let  $\widetilde{\mathcal{W}}_{n,d} = (\widetilde{W}_{i,j})_{1 \leq i, j \leq n}$  with  $\widetilde{W}_{i,j}$  given by (3.9). Then for every  $1 \leq i \leq n$  such that  $q_i = q_0$ , we have*

$$\widetilde{W}_{i,i} \rightarrow_{d \rightarrow \infty}^{(L)} Z^{(2, 2H'-1)}$$

with  $H'$  as in (2.2). For every  $1 \leq i \leq n$  with  $q_i > q_0$ ,

$$\widetilde{W}_{i,i} \rightarrow_{d \rightarrow \infty} 0 \text{ in } L^2(\Omega).$$

*Proof.* By Theorem 3.1 and the scaling property of the Hermite process, each diagonal term  $\widetilde{W}_{i,i}$  (given by (3.9)) of the matrix  $\widetilde{\mathcal{W}}_{n,d}$  with  $q_i = q_0$  converges to a Rosenblatt random variable with self-similarity index  $2H' - 1$  with  $H'$  from (2.2). The other diagonal terms of  $\widetilde{\mathcal{W}}_{n,d}$  (those  $\widetilde{W}_{i,i}$  with  $q_i > q_0$ ) converges to zero in  $L^2(\Omega)$ . Indeed, if  $q_i > q_0$ ,

$$\begin{aligned} \mathbf{E}|\widetilde{W}_{i,i}|^2 &= c_{q_0,H}^2 \mathbf{E} \left| d^{\frac{2-2H}{q_0}} A_{i,i} \right|^2 \\ &= c_{q_0,H}^2 d^{2(2H-2)(\frac{1}{q_0} - \frac{1}{q_i})} \mathbf{E} \left| d^{\frac{2-2H}{q_i}} A_{i,i} \right|^2 \rightarrow_{d \rightarrow \infty} 0. \end{aligned}$$

□

Let us regard the asymptotic behavior of the non diagonal terms. This will follow from the following result.

**Proposition 3.3.** *Let  $Z^{(q_1,H)}, Z^{(q_2,H)}$  be two independent Hermite processes of orders  $q_1, q_2 \geq 2$  respectively. Assume  $q_1 \leq q_2$  and define*

$$F_d(q_1, q_2) = d^{\frac{2-2H}{q_1}-1} \sum_{k=0}^{d-1} \left( Z_{k+1}^{(q_1,H)} - Z_k^{(q_1,H)} \right) \left( Z_{k+1}^{(q_2,H)} - Z_k^{(q_2,H)} \right). \quad (3.10)$$

Then

$$\mathbf{E}F_d(q_1, q_2)^2 \leq c \begin{cases} d^{\frac{4-4H}{q_1}-1} & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4}\right) \\ \log(d)d^{\frac{1}{q_1}-1} & \text{if } H = \frac{3}{4} \\ d^{(4H-4)(1-\frac{1}{q_1})} & \text{if } H \in \left(\frac{3}{4}, 1\right) \end{cases}, \quad (3.11)$$

In particular  $F_d(q_1, q_2)$  converges to zero in  $L^2(\Omega)$  as  $d \rightarrow \infty$ .

*Proof.* We have for every  $d \geq 1$ ,

$$\begin{aligned} &\mathbf{E}F_d(q_1, q_2)^2 \\ &= d^{\frac{4-4H}{q_1}-2} \sum_{k,\ell=0}^{d-1} \mathbf{E} \left[ \left( Z_{k+1}^{(q_1,H)} - Z_k^{(q_1,H)} \right) \left( Z_{k+1}^{(q_2,H)} - Z_k^{(q_2,H)} \right) \right. \\ &\quad \left. \times \left( Z_{k+1}^{(q_1,H)} - Z_k^{(q_1,H)} \right) \left( Z_{k+1}^{(q_2,H)} - Z_k^{(q_2,H)} \right) \right] \\ &= d^{\frac{4-4H}{q_1}-2} \sum_{k,\ell=0}^{d-1} \mathbf{E} \left[ \left( Z_{k+1}^{(q_1,H)} - Z_k^{(q_1,H)} \right) \left( Z_{\ell+1}^{(q_1,H)} - Z_\ell^{(q_1,H)} \right) \right] \\ &\quad \times \mathbf{E} \left[ \left( Z_{k+1}^{(q_2,H)} - Z_k^{(q_2,H)} \right) \left( Z_{\ell+1}^{(q_2,H)} - Z_\ell^{(q_2,H)} \right) \right] \\ &= d^{\frac{4-4H}{q_1}-2} \sum_{k,\ell=0}^{d-1} \rho_H(k-l)^2 \\ &= d^{\frac{4-4H}{q_1}-1} \sum_{v \in \mathbb{Z}} \rho_H(|v|)^2 \left( 1 - \frac{|v|}{n} \right) 1_{(|v| < n)} \end{aligned}$$



with  $\rho_H$  from (2.7). Now we use the fact that  $\rho_H(|k|)$  behaves as  $H(2H-1)|k|^{2H-2}$  as  $|k| \rightarrow \infty$ .

The fact that  $F_d(q_1, q_2)$  converges to zero in  $L^2(\Omega)$  as  $d \rightarrow \infty$  can be deduced from (3.11) since for  $q_1 \geq 2$  we have for  $H \in (\frac{1}{2}, 1)$  and  $q \geq 2$ ,  $\frac{4-4H}{q_1} - 1 < \frac{2}{q_1} - 1 \leq 0$ ,  $(4H-4)(1 - \frac{1}{q_1}) < 0$  and  $\frac{1}{q_1} - 1 < 0$ .  $\square$

As a consequence of Proposition 3.3, we have

**Corollary 3.4.** *Let  $\widetilde{\mathcal{W}}_{n,d} = (\widetilde{W}_{i,j})_{1 \leq i,j \leq n}$  be the renormalized Wishart matrix with  $\widetilde{W}_{i,j}$  given by (3.9). Then for every  $1 \leq i, j \leq n$  with  $i \neq j$  and  $q_i \leq q_j$ , there exists a positive constant  $c$  such as*

$$\mathbf{E}|\widetilde{W}_{i,j}|^2 \leq c \begin{cases} d^{\frac{4-4H}{q_i}-1} & \text{if } H \in (\frac{1}{2}, \frac{3}{4}) \\ \log(d)d^{\frac{1}{q_i}-1} & \text{if } H = \frac{3}{4} \\ d^{(4H-4)(1-\frac{1}{q_i})} & \text{if } H \in (\frac{3}{4}, 1). \end{cases},$$

In particular  $\widetilde{W}_{i,j}$  converges to zero in  $L^2(\Omega)$  as  $d \rightarrow \infty$ .

*Proof.* Let  $i \neq j$  and assume  $q_i \leq q_j$ . Then

$$\begin{aligned} \mathbf{E}|\widetilde{W}_{i,j}|^2 &= c_{q,H}^2 d^{2(2H-2)(\frac{1}{q_0}-\frac{1}{q_i})} \\ &\quad \mathbf{E} \left| d^{\frac{2-2H}{q_i}-1} \sum_{k=0}^{d-1} \left( Z_{k+1}^{(q_i,H,i)} - Z_k^{(q_i,H,i)} \right) \left( Z_{k+1}^{(q_j,H,j)} - Z_k^{(q_j,H,j)} \right) \right|^2 \\ &\leq c_{q,H}^2 \mathbf{E} \left| d^{\frac{2-2H}{q_i}-1} \sum_{k=0}^{d-1} \left( Z_{k+1}^{(q_i,H,i)} - Z_k^{(q_i,H,i)} \right) \left( Z_{k+1}^{(q_j,H,j)} - Z_k^{(q_j,H,j)} \right) \right|^2 \end{aligned}$$

and it suffices to apply Proposition 3.3.  $\square$

From Theorem 3.1 and Corollaries 3.2 and 3.4 we deduce that the renormalized Wishart matrix  $\widetilde{\mathcal{W}}_{n,d}$  with the entries (3.9) converges componentwise to a diagonal  $n \times n$  matrix  $\mathcal{R}_n^H = (R_{i,j}^H)_{1 \leq i,j \leq n}$  with independent diagonal entries given by

$$R_{i,i}^H = Z_1^{(q_i, 2H'-1, i)} 1_{q_i=q_0}. \quad (3.12)$$

If there exists only one order  $q_i$  such that  $q_i = q_0$  (assume that this is  $q_1$ ) then the diagonal elements of the matrix  $\mathcal{R}_n^H$  are

$$R_{1,1}^H = Z^{(q_1, 2H'-1, 1)} \text{ and } R_{i,i}^H = 0 \text{ for } i = 2, \dots, n.$$

If  $q_1 = q_2 = \dots = q_n = q$ , then for every  $i = 1, \dots, n$ ,

$$R_{i,i}^H = Z^{(q, 2H'-1, i)}.$$

#### 4. The Wasserstein Distance Between the Wishart Matrix and its Limiting Matrix

We need to evaluate how fast the sequence  $V_d$  converges to its Rosenblatt limit in  $L^2(\Omega)$ .

**Proposition 4.1.** *Let  $V_d$  be given by (3.7) and let  $Z_1^{(2,2H'-1)}$  be its limit in  $L^2(\Omega)$  as  $d \rightarrow \infty$ . Then, for  $q \geq 3$  and  $d$  large enough,*

$$\mathbf{E} \left| V_d - Z_1^{(2,2H'-1)} \right|^2 \leq c \begin{cases} d^{\frac{4-4H}{q}-1} & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4}\right) \\ d^{-\frac{1}{q}} \log(d), & \text{if } H = \frac{3}{4} \\ d^{\frac{4H-4}{q}} & \text{if } H \in \left(\frac{3}{4}, 1\right) \end{cases}$$

and for  $q = 2$  and  $d$  large enough,

$$\mathbf{E} \left| V_d - Z_1^{(2,2H'-1)} \right|^2 \leq c \begin{cases} d^{1-2H} & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4}\right) \\ d^{-\frac{1}{2}} \log(d), & \text{if } H = \frac{3}{4} \\ d^{2H-2} & \text{if } H \in \left(\frac{3}{4}, 1\right) \end{cases}$$

*Proof.* The case  $q = 2$  has been treated in [2]. Assume  $q \geq 3$ . The random variable  $V_d$  admits the following chaos expansion

$$V_d = c_{q,H} d^{\frac{2-2H}{q}-1} \sum_{k=0}^{d-1} [d^{2H} I_q(f_{k,d})^2 - 1]$$

with (the kernel  $L$  is given by (2.5)),

$$f_{k,d} = L_{\frac{k+1}{d}} - L_{\frac{k}{d}}, \text{ for } d \geq 1, k = 0, \dots, d-1.$$

By the product formula for multiple stochastic integrals (5.4), we find

$$\begin{aligned} V_d &= c_{q,H} d^{\frac{2-2H}{q}-1} \sum_{k=0}^{d-1} \left[ d^{2H} \sum_{r=0}^q r! (C_q^r)^2 I_{2q-2r}(f_{k,d} \otimes_r f_{k,d}) - 1 \right] \\ &= c_{q,H} d^{\frac{2-2H}{q}-1} \sum_{k=0}^{d-1} d^{2H} \sum_{r=0}^{q-1} r! (C_q^r)^2 I_{2q-2r}(f_{k,d} \otimes_r f_{k,d}) \\ &:= \sum_{r=0}^{q-1} S_{2q-2r} \end{aligned}$$

with, for  $r = 0, \dots, q-1$ ,

$$S_{2q-2r} = c_{q,H} d^{\frac{2-2H}{q}+2H-1} \sum_{k=0}^{d-1} r! (C_q^r)^2 I_{2q-2r}(f_{k,d} \otimes_r f_{k,d}) \quad (4.1)$$

Notice that the random variable  $S_{2q-2r}$  is an element of the  $2q-2r$ th Wiener chaos.

We will write the difference  $V_d - Z_1^{(2,2H'-1)}$  as

$$V_d - Z_1^{(2,2H'-1)} = \sum_{r=0}^{q-2} S_{2q-2r} + (S_2 - Z_1^{(2,2H'-1)})$$

and by the orthogonality of the Wiener chaoses of different orders

$$\mathbf{E} \left| V_d - Z_1^{(2,2H'-1)} \right|^2 = \sum_{r=0}^{q-2} \mathbf{E} |S_{2q-2r}|^2 + \mathbf{E} \left| S_2 - Z_1^{(2,2H'-1)} \right|^2. \quad (4.2)$$

The  $L^2(\Omega)$  norm of the terms  $S_{2q-2r}$  with  $r = 0, \dots, q-2$  has been estimated in [5]. From Proposition 4.2 in [5] (see relation (3.9)) we have that for  $r = 1, \dots, q-2$

$$d^{\frac{4H-4}{q}-2(2H'-2)(q-r)} \mathbf{E}|S_{2q-2r}|^2 \rightarrow_{d \rightarrow \infty} c_{r,q,H}$$

with some explicit strictly positive constants  $c_{r,q,H}$  (see formula (3.10) in [5]). Therefore, for  $d$  sufficiently large, since  $H'$  is given by (2.2),

$$\mathbf{E}|S_{2q-2r}|^2 \leq Cd^{\frac{4-4H}{q}+2(2H'-2)(q-r)} = Cd^{\frac{4H-4}{q}(q-r-1)}. \quad (4.3)$$

The term  $S_{2q}$  in the  $2q$ th Wiener chaos (obtained for  $r = 0$ ) has a different behavior. Actually, if  $H \in (\frac{3}{4}, 1)$ , it obeys the same estimate as above, i.e.

$$\mathbf{E}S_{2q}^2 \leq cd^{(4H-4)(1-\frac{1}{q})} \quad (4.4)$$

while for  $H \in (\frac{1}{2}, \frac{3}{4})$ , one has

$$\mathbf{E}S_{2q}^2 \leq cd^{\frac{4-4H}{q}-1} \quad (4.5)$$

and for  $H = \frac{3}{4}$ ,

$$\mathbf{E}S_{2q}^2 \leq cd^{\frac{1}{q}-1} \log(d). \quad (4.6)$$

It remains to estimate  $\mathbf{E} \left| S_2 - Z_1^{(2,2H'-1)} \right|^2$ . Actually, the estimate for this term follows from the proof of Proposition 4.3 in [2], by replacing in this proof  $H$  by  $2H' - 1 = \frac{2-2H}{q} - 1$ . We will have, for  $d$  large, for all  $H \in (\frac{1}{2}, 1)$ ,

$$\mathbf{E} \left| S_2 - Z_1^{(2,2H'-1)} \right|^2 \leq Cd^{1-2(2H'-1)} = Cd^{\frac{4-4H}{q}-1}. \quad (4.7)$$

To finish the proof, we use the bounds (4.3)-(4.7) and we notice that for every  $r = 1, \dots, q-2$ ,

$$\frac{4-4H}{q} - 1 \geq \frac{4H-4}{q}(q-r-1) \text{ if } H \in \left( \frac{1}{2}, \frac{3}{4} \right)$$

and

$$\frac{4-4H}{q} - 1 \leq \frac{4H-4}{q}(q-r-1) \text{ if } H \in \left( \frac{3}{4}, 1 \right).$$

When  $H = \frac{3}{4}$ , we need to compare the right-hand sides of (4.3), (4.6) and (4.7). Since for every  $r = 1, \dots, q-2$  and  $d$  large enough,  $d^{-\frac{1}{q}} \geq d^{-\frac{1}{q}(q-r-1)} \geq \log(d)d^{\frac{1}{q}-1}$  we get the conclusion.  $\square$

*Remark 4.2.* The different behavior in the cases  $q = 2$  and  $q \geq 3$  can be noticed for the limiting situation when  $H = \frac{3}{4}$ . This comes from the presence of the terms  $S_{2q-2r}$  with  $r = 1, \dots, q-2$  given by (4.1) which appear only when  $q \geq 3$ . For  $H = \frac{3}{4}$  and  $q \geq 3$  this is the dominant term in the right-hand side of (4.2).

**Theorem 4.3.** *Let  $\widetilde{\mathcal{W}}_{n,d}$  be the renormalized Wishart matrix with entries (3.9) and let  $\mathcal{R}_n^H$  be its limiting matrix as  $d \rightarrow \infty$ . Then for every  $n \geq 1$ ,  $q \geq 3$  and for  $d$  large enough,*

$$d_W \left( \widetilde{\mathcal{W}}_{n,d}, \mathcal{R}_n^H \right) \leq C \begin{cases} nd^{\frac{2-2H}{q_0}-1/2} & \text{if } H \in \left( \frac{1}{2}, \frac{3}{4} \right) \\ nd^{-\frac{1}{2q_0}}, & \text{if } H = \frac{3}{4} \\ nd^{\frac{2H-2}{q_0}} & \text{if } H \in \left( \frac{3}{4}, 1 \right) \end{cases} \quad (4.8)$$

while for every  $n \geq 1$ ,  $q = 2$  and for  $d$  large enough,

$$d_W \left( \widetilde{\mathcal{W}}_{n,d}, \mathcal{R}_n^H \right) \leq C \begin{cases} nd^{\frac{1}{2}-H} & \text{if } H \in \left( \frac{1}{2}, \frac{3}{4} \right) \\ n\sqrt{\log(d)}d^{-\frac{1}{4}} & \text{if } H = \frac{3}{4} \\ nd^{H-1} & \text{if } H \in \left( \frac{3}{4}, 1 \right) \end{cases}. \quad (4.9)$$

*Proof.* Again take  $q \geq 3$  since the result for  $q = 2$  is known from Theorem 2 in [2]. The independence of the Hermite processes  $Z^{(q_i, H, i)}$  for  $i = 1, \dots, n$  and the scaling property of the Hermite process imply that the matrix  $\widetilde{\mathcal{W}}_{n,d}$  has the same distribution as the matrix  $\mathcal{A}_{n,d} = c_{q_0, H} d^{\frac{2-2H}{q_0}} (A_{i,j})_{1 \leq i, j \leq n}$  with  $A_{i,j}$  given by (3.5) and (3.6). Hence

$$d_W \left( \widetilde{\mathcal{W}}_{n,d}, \mathcal{R}_n^H \right) = d_W \left( \mathcal{A}_{n,d}, \mathcal{R}_n^H \right) \leq \sqrt{\sum_{i,j=1}^n \mathbf{E} \left( c_{q_0, H} d^{\frac{2-2H}{q_0}} A_{i,j} - R_{i,j}^H \right)^2}$$

where the last bound is due to the properties of the Wasserstein distance.

Now, Proposition 4.1 gives the estimate for  $\mathbf{E} \left( c_{q_0, H} d^{\frac{2-2H}{q_0}-1} A_{i,i} - R_{i,i}^H \right)^2$  while for  $i \neq j$ ,

$$\mathbf{E} \left( c_{q_0, H} d^{\frac{2-2H}{q_0}} A_{i,j} - R_{i,j}^H \right)^2 = c_{q_0, H}^2 \mathbf{E} \left( d^{\frac{2-2H}{q_0}} A_{i,j} \right)^2 \leq c_{q_0, H}^2 \mathbf{E} (F_d(q_i, q_j))^2$$

with  $F_d(q_i, q_j)$  given by (3.10) and whose  $L^2(\Omega)$  norm is estimated in Proposition 3.3.  $\square$

*Remark 4.4.*

- If  $q_1 = \dots = q_n = 2$ , we retrieve the main result in [2]. Actually, notice that only the minimal order  $q_0$  of the Wiener chaos related to the elements of the matrix  $\mathcal{X}_{n,d}$  appear in the Wasserstein bound (4.8). This means that if, there exists at least one  $q_i, i = 1, \dots, n$  such that  $q_i = 2$ , the bound for the Wasserstein distance  $d_W \left( \widetilde{\mathcal{W}}_{n,d}, \mathcal{R}_n^H \right)$  is the same as if  $q_1 = q_2 = \dots = q_n = 2$ .
- Another way to interpret the result in Theorem 4.3 is to say that the random matrices  $\widetilde{\mathcal{W}}_{n,d}$  and  $\mathcal{R}_n^H$  are  $\Phi(n, d)$ -close as  $d, n$  go to infinity, where the function  $\Phi(n, d)$  is given by the right-hand side of (4.8). That means that (see e.g. [8]) the Wasserstein distance  $d_W \left( \widetilde{\mathcal{W}}_{n,d}, \mathcal{R}_n^H \right)$  converges to zero when  $\Phi(n, d) \rightarrow 0$ .

### 5. Appendix: Multiple Stochastic Integrals

Here, we shall only recall some elementary facts; our main reference is [9]. Consider  $\mathcal{H}$  a real separable infinite-dimensional Hilbert space with its associated inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ , and  $(B(\varphi), \varphi \in \mathcal{H})$  an isonormal Gaussian process on a probability space  $(\Omega, \mathfrak{F}, \mathbb{P})$ , which is a centered Gaussian family of random variables such that  $\mathbf{E}(B(\varphi)B(\psi)) = \langle \varphi, \psi \rangle_{\mathcal{H}}$ , for every  $\varphi, \psi \in \mathcal{H}$ . Denote by  $I_q$  the  $q$ th multiple stochastic integral with respect to  $B$ . This  $I_q$  is actually an isometry between the Hilbert space  $\mathcal{H}^{\odot q}$  (symmetric tensor product) equipped with the scaled norm  $\frac{1}{\sqrt{q!}} \|\cdot\|_{\mathcal{H}^{\otimes q}}$  and the Wiener chaos of order  $q$ , which is defined as the closed linear span of the random variables  $H_q(B(\varphi))$  where  $\varphi \in \mathcal{H}$ ,  $\|\varphi\|_{\mathcal{H}} = 1$  and  $H_q$  is the Hermite polynomial of degree  $q \geq 1$  defined by:

$$H_q(x) = (-1)^q \exp\left(\frac{x^2}{2}\right) \frac{d^q}{dx^q} \left( \exp\left(-\frac{x^2}{2}\right) \right), \quad x \in \mathbb{R}. \quad (5.1)$$

The isometry of multiple integrals can be written as: for  $p, q \geq 1$ ,  $f \in \mathcal{H}^{\otimes p}$  and  $g \in \mathcal{H}^{\otimes q}$ ,

$$\mathbf{E}\left(I_p(f)I_q(g)\right) = \begin{cases} q! \langle \tilde{f}, \tilde{g} \rangle_{\mathcal{H}^{\otimes q}} & \text{if } p = q \\ 0 & \text{otherwise} \end{cases} \quad (5.2)$$

where  $\tilde{f}$  denotes the canonical symmetrization of  $f$  and it is defined by:

$$\tilde{f}(x_1, \dots, x_q) = \frac{1}{q!} \sum_{\sigma \in \mathcal{S}_q} f(x_{\sigma(1)}, \dots, x_{\sigma(q)}),$$

in which the sum runs over all permutations  $\sigma$  of  $\{1, \dots, q\}$ . It also holds that:

$$I_q(f) = I_q(\tilde{f}).$$

In the particular case when  $\mathcal{H} = L^2(T, \mathcal{B}(T), \lambda)$  ( $\lambda$  being the Lebesgue measure), the  $r$ th contraction  $f \otimes_r g$  is the element of  $\mathcal{H}^{\otimes(p+q-2r)}$ , which is defined by:

$$\begin{aligned} & (f \otimes_r g)(s_1, \dots, s_{p-r}, t_1, \dots, t_{q-r}) \\ &= \int_{T^r} du_1 \dots du_r f(s_1, \dots, s_{p-r}, u_1, \dots, u_r) g(t_1, \dots, t_{q-r}, u_1, \dots, u_r), \end{aligned} \quad (5.3)$$

for every  $f \in L^2(T^p)$ ,  $g \in L^2(T^q)$  and  $r = 1, \dots, p \wedge q$ . By  $f \tilde{\otimes}_l g$  we denote the symmetrization of the contraction  $f \otimes_l g$ .

The product for two multiple integrals can be expanded into a sum of multiple integrals (see [9]): if  $f \in L^2(T^m)$  and  $g \in L^2(T^n)$  are symmetric functions, then it holds that

$$I_n(f)I_m(g) = \sum_{l=0}^{m \wedge n} l! C_m^l C_n^l I_{m+n-2l}(f \tilde{\otimes}_l g). \quad (5.4)$$

**Acknowledgment.** The authors were supported in part by the Labex CEMPI (ANR-11-LABX-0007-01)

## References

1. Bishop, A. N., Del Moral, P., and Niclas, A.: *A Introduction to Wishart Matrix Moments*. Foundations and Trends of Machine Learning 11:2. NOW publishers, 2018.
2. Bourguin, S., Diez, Ch.-Ph., and Tudor, C. A.: Limiting behavior of large correlated Wishart matrices with chaotic entries. Preprint (2020) to appear in *Bernoulli*.
3. Bubeck, S., Ding, J., Eldan, R., and Rácz, M. Z.: Testing high-dimensional geometry in random graphs, *Random Structures and Algorithms* **49** (2016) no. 3, 503–532.
4. Bubeck, S. and Ganguly, S.: Entropic CLT and phase transition in high dimensional Wishart matrices, *Int. Math. Research Not.* **2** (2016), 588–606.
5. Chronopoulou, A., Tudor, C. A., and Viens, F. G.: Self-similarity parameter estimation and reproduction property for non-Gaussian Hermite processes. *Commun. Stoch. Anal.* **5** (2011) no. 1, 161–185.
6. Jiang, T. and Li, D.: Approximation of rectangular beta-Laguerre ensembles and large deviations, *J. Theoret. Probab.* **28** (2015), no. 3, 804–847.
7. Johnstone, I. M.: High Dimensional Statistical Inference and Random Matrices. Proceedings of the International Congress of Mathematicians, Madrid, Spain, 2006.
8. Nourdin, I. and Zheng, G.: Asymptotic behavior of large Gaussian correlated Wishart matrices, preprint (2019).
9. Nualart, D.: *Malliavin Calculus and Related Topics. Second Edition*. Springer, New York, 2006.
10. Pipiras, V. and Taqqu, M. S.: *Long-range dependence and self-similarity*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2017.
11. Rácz, M. Z. and Richey, J.: A smooth transition from Wishart to GOE, *J. Theoret. Probab.* **32** (2019), no. 2, 898–906.
12. Tudor, C. A.: *Analysis of variations for self-similar processes. A stochastic calculus approach*. Probability and its Applications (New York). Springer, Cham, 2013.

CHARLES-PHILIPPE DIEZ: CNRS, UNIVERSITÉ DE LILLE, LABORATOIRE PAUL PAINLEVÉ UMR 8524, F-59655 VILLENEUVE D’ASCQ, FRANCE.

*E-mail address:* [chphdiez@gmail.com](mailto:chphdiez@gmail.com)

CIPRIAN A. TUDOR: CNRS, UNIVERSITÉ DE LILLE, LABORATOIRE PAUL PAINLEVÉ UMR 8524, F-59655 VILLENEUVE D’ASCQ, FRANCE.

*E-mail address:* [ciprian.tudor@univ-lille.fr](mailto:ciprian.tudor@univ-lille.fr)