

# ASYMPTOTIC BEHAVIOR OF LARGE GAUSSIAN CORRELATED WISHART MATRICES

IVAN NOURDIN AND GUANGQU ZHENG

ABSTRACT. We consider high-dimensional Wishart matrices  $d^{-1}\mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$ , associated with a rectangular random matrix  $\mathcal{X}_{n,d}$  of size  $n \times d$  whose entries are jointly Gaussian and correlated. Even if we will consider the case of overall correlation among the entries of  $\mathcal{X}_{n,d}$ , our main focus is on the case where the rows of  $\mathcal{X}_{n,d}$  are independent copies of a  $n$ -dimensional stationary centered Gaussian vector of correlation function  $s$ . When  $s$  belongs to  $\ell^{4/3}(\mathbb{Z})$ , we show that a proper normalization of  $d^{-1}\mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$  is close in *Wasserstein distance* to the corresponding Gaussian ensemble *as long as  $d$  is much larger than  $n^3$* , thus recovering the main finding of [3, 9] and extending it to a larger class of matrices. We also investigate the case where  $s$  is the correlation function associated with the fractional Brownian noise of parameter  $H$ . This example is very rich, as it gives rise to a great variety of phenomena with very different natures, depending on how  $H$  is located with respect to  $1/2$ ,  $5/8$  and  $3/4$ . Notably, when  $H > 3/4$ , our study highlights a new probabilistic object, which we have decided to call the *Rosenblatt-Wishart matrix*. Our approach crucially relies on the fact that the entries of the Wishart matrices we are dealing with are double Wiener-Itô integrals, allowing us to make use of multivariate bounds arising from the Malliavin-Stein method and related ideas. To conclude the paper, we analyze the situation where the row-independence assumption is relaxed and we also look at the setting of random  $p$ -tensors ( $p \geq 3$ ), a natural extension of Wishart matrices.

## 1. INTRODUCTION AND MAIN RESULTS

Let  $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$  be a  $n \times d$  random matrix whose entries are *independent* copies of a real centered random variable with unit variance (say). The (real) *Wishart matrix*  $d^{-1}\mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$ , introduced by the statistician J. Wishart [21] in the late 20s, has been very useful in the multivariate statistics and it arises naturally as the sample covariance matrix. When  $n$  is *fixed* and  $d$  goes to infinity, the matrix  $d^{-1}\mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$  converges almost surely to the identity matrix  $\mathcal{I}_n$ , according to the strong law of large numbers. Moreover, the multivariate central limit theorem implies that the fluctuations of Wishart matrix around  $\mathcal{I}_n$  are Gaussian, provided  $X_{11}$  has the finite fourth moment.

For a long time, the case where  $n$  is fixed was enough for applications. But in the current world filled with large data sets, there has been a change of paradigm: that both  $d$  and  $n$  are large simultaneously has now become the rule rather than the exception; see *e.g.* Johnstone's ICM survey [10]. In such a context, one can no longer merely rely on the law of large numbers

---

*Date:* April 18, 2018.

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

*Key words and phrases.* Stein's method; Malliavin calculus; High-dimensional regime; Rosenblatt-Wishart matrix.

I. Nourdin was partially supported by the Grant F1R-MTH-PUL-15CONF at Luxembourg University.

and the classical central limit theorem to analyze the asymptotic behavior of the Wishart matrix.

In order to describe the asymptotic behavior of  $d^{-1}\mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$  when *both*  $d$  and  $n$  go to infinity, a classical strategy in random matrix theory consists in analyzing the weak convergence of its *empirical spectral distribution*  $\mu_{n,d}$ , defined as  $\mu_{n,d} = n^{-1} \sum_{i=1}^n \delta_{\lambda_i(n,d)}$ , where  $\delta_\lambda$  stands for the Dirac mass at  $\lambda$  and  $\lambda_1(n,d) \leq \dots \leq \lambda_n(n,d)$  are the eigenvalues of  $d^{-1}\mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$ . It is known since Marchenko and Pastur [11] that, if  $n = n(d)$  *diverges* in a way such that  $n/d \rightarrow c \in (0, \infty)$ , then  $\mu_{n,d}$  converges weakly to  $\mu_c = \max\{0, 1 - c^{-1}\}\delta_0 + (2\pi c)^{-1} \sqrt{(a_+ - x)(x - a_-)} \mathbf{1}_{[a_-, a_+]}(x) dx$  with  $a_\pm = (1 \pm \sqrt{c})^2$ . At the fluctuation level, it is also known that linear statistics of eigenvalues of Wishart matrix satisfies central limit theorems under certain conditions; see for instance [5] and references therein.

However, for some applications the previous way to describe the asymptotic behavior of a large Wishart matrix might not be appropriate (for several possible reasons: because  $n$  and  $d$  are not of the same order, or because we do not have access to eigenvalues, etc.) In this work, we take another approach, recently introduced in [3, 9] and that we shall describe now. To ease the presentation, we start with the following definition.

**Definition 1.1.** *For each  $n \geq 1$ , suppose that we have two families  $\{\mathcal{W}_{n,d} : d \geq 1\}$ ,  $\{\mathcal{Z}_{n,d} : d \geq 1\}$  of  $n \times n$  random matrices. Consider a function  $\phi : \mathbb{N}^* \times \mathbb{N}^* \rightarrow [0, \infty]$ . We say that  $\mathcal{W}_{n,d}$  is  $\phi$ -close to  $\mathcal{Z}_{n,d}$  if the Fortet-Mourier distance  $d_{\text{FM}}(\mathcal{W}_{n,d}, \mathcal{Z}_{n,d})$  between them tends to zero, when  $d, n \rightarrow \infty$  and  $\phi(n, d) \rightarrow 0$ .*

In Definition 1.1, we use the *Fortet-Mourier distance* between two random variables with values in  $\mathcal{M}_n(\mathbb{R})$ , the space of  $n \times n$  real matrices. Let us recall its definition: if  $\mathcal{X}$  and  $\mathcal{Y}$  are two such random matrices, then

$$d_{\text{FM}}(\mathcal{X}, \mathcal{Y}) := \sup \{ \mathbb{E}[g(\mathcal{X})] - \mathbb{E}[g(\mathcal{Y})] : \|g\|_\infty + \|g\|_{\text{Lip}} \leq 1 \}, \quad (1.1)$$

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

where  $\|\cdot\|_{\text{HS}}$  is the Hilbert-Schmidt norm on  $\mathcal{M}_n(\mathbb{R})$ . Since  $(\mathcal{M}_n(\mathbb{R}), \|\cdot\|_{\text{HS}})$  is a Polish space, it is well known that  $d_{\text{FM}}$  characterizes the weak convergence of probability measures on  $\mathcal{M}_n(\mathbb{R})$ ; see *e.g.* [7, Section 11.3]. We will also use the *Wasserstein distance*, which is a stronger distance and defined in a similar way:

$$d_{\text{Wass}}(\mathcal{X}, \mathcal{Y}) = \sup \{ \mathbb{E}[g(\mathcal{X})] - \mathbb{E}[g(\mathcal{Y})] : \|g\|_{\text{Lip}} \leq 1 \}. \quad (1.2)$$

It is trivial that  $d_{\text{FM}}(\mathcal{X}, \mathcal{Y}) \leq d_{\text{Wass}}(\mathcal{X}, \mathcal{Y})$ , so that any bound on the Wasserstein distance implies the same bound for the Fortet-Mourier distance.

With the above definition and notation in mind, let us go back to the study of high-dimensional fluctuation of Wishart matrices by considering a normalized version of  $d^{-1}\mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$ , namely

$$\mathcal{W}_{n,d} = \sqrt{d} \left( \frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - \mathcal{I}_n \right). \quad (1.3)$$

Also, consider the GOE matrix

$$\mathcal{Z}_n = (Z_{ij})_{1 \leq i, j \leq n}, \quad (1.4)$$

where  $Z_{ii} \sim N(0, 2)$ ,  $Z_{ij} \sim N(0, 1)$  for  $i < j$ ,  $Z_{ij} = Z_{ji}$  for  $i > j$ , and  $\{Z_{ij}, i \leq j\}$  are independent. Interestingly, the following phenomenon discovered independently in [3, 9] arises: when the entry distribution of  $\mathcal{X}_{n,d}$  is *standard Gaussian*, then  $\mathcal{W}_{n,d}$  and  $\mathcal{Z}_n$  are  $\phi$ -close (with respect to the total variation distance<sup>1</sup>) for  $\phi(n, d) = n^3/d$ . Otherwise stated, one cannot distinguish between the laws of  $\mathcal{W}_{n,d}$  and  $\mathcal{Z}_n$  when  $n$  and  $d$  go to infinity with  $d$  much larger than  $n^3$ . It is also proved in [3] that the condition  $n^3 = o(d)$  is sharp in the following sense: if we assume this time that  $d = o(n^3)$ , then the total variation distance between the laws of  $\mathcal{W}_{n,d}$  and  $\mathcal{Z}_n$  goes to 1; see [3] for precise statements and see also [19] for a proof that this phase transition from  $d = o(n^3)$  to  $n^3 = o(d)$  is smooth. The work [3] was soon generalized in [4] to the setting that the entries of  $\mathcal{X}_{n,d}$  are *i.i.d. log-concave*<sup>2</sup> random variables and some very similar results on the high-dimensional regime were obtained therein.

In the present paper, we take another path of generalization, by relaxing the full independence assumption that is made in [3, 4] on the entries of  $\mathcal{X}_{n,d}$ . In fact, the most basic phenomena in multivariate analysis is that of correlation – the tendency of quantities to vary together. And the appearance of correlation usually increases drastically the complexity of the problem at hand, see *e.g.* [4, Section 6]. As a first step, we have decided to mainly focus on the case where the entries of  $\mathcal{X}_{n,d}$  are *Gaussian* and exhibit *row independence*, that is, we will allow the columns of  $\mathcal{X}_{n,d}$  to be correlated, but not its rows; see Section 4 for further results on the case of overall correlation.

Here is our precise framework. Let  $\mathfrak{H}$  be a real separable Hilbert space equipped with the inner product  $\langle \cdot, \cdot \rangle_{\mathfrak{H}}$  and the Hilbert norm  $\|\cdot\|_{\mathfrak{H}}$ , and let  $\{e_{ij} : i, j \geq 1\} \subset \mathfrak{H}$  be a family such that

$$\langle e_{ij}, e_{i'j'} \rangle_{\mathfrak{H}} = \mathbf{1}_{\{i=i'\}} s(j - j'). \quad (1.5)$$

In (1.5),  $s : \mathbb{Z} \rightarrow \mathbb{R}$  stands for some correlation function satisfying suitable assumptions to be given later on. One of them is that  $s(0) = 1$ , implying in particular that  $\|e_{ij}\|_{\mathfrak{H}} = 1$  for all  $i, j \geq 1$ .

Consider the corresponding Gaussian sequence  $X_{ij} = X(e_{ij}) \sim N(0, 1)$ , where  $X = \{X(h), h \in \mathfrak{H}\}$  is a centered Gaussian process indexed by  $\mathfrak{H}$  such that  $\mathbb{E}[X(g)X(h)] = \langle g, h \rangle_{\mathfrak{H}}$  for all  $g, h \in \mathfrak{H}$ , that is,  $X$  is an isonormal process over  $\mathfrak{H}$ . As before, let  $\mathcal{X}_{n,d}^{(s)}$  be the  $n \times d$  random matrix given by

$$\mathcal{X}_{n,d}^{(s)} = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1d} \\ X_{21} & X_{22} & \dots & X_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nd} \end{pmatrix}. \quad (1.6)$$

Given the form of (1.5), we note that the rows  $\{(X_{i1}, \dots, X_{id}), i = 1, \dots, n\}$  of  $\mathcal{X}_{n,d}^{(s)}$  are independent and identically distributed.

We are now in a position to state our first main result, in which we basically extend the main result of [3, 9] to rectangular random matrices of the type (1.6), that is, to a situation where the entries of  $\mathcal{X}_{n,d}$  are *Gaussian* and *partially correlated*.

<sup>1</sup>Due to the explicit density functions of these two random matrix ensembles, the authors of [3] were indeed able to compute the total variation distance as the  $L_1$  distance between densities; see also the earlier work [9], in which the same conclusion was derived independently, using quite involved spectral analysis.

<sup>2</sup>that is,  $X_{11}$  has a density function  $\phi$  such that  $\log \phi$  is a concave function.

**Theorem 1.2** (Gaussian approximation). *Let  $\mathcal{X}_{n,d}^{(s)}$  be given by (1.6), and consider*

$$\mathcal{W}_{n,d}^{(s)} = (W_{ij})_{1 \leq i, j \leq n} = \sqrt{d} \left( \frac{1}{d} \mathcal{X}_{n,d}^{(s)} (\mathcal{X}_{n,d}^{(s)})^T - \mathcal{I}_n \right). \quad (1.7)$$

Let  $\mathcal{G}_{n,d}^{(s)} = (G_{ij})_{1 \leq i, j \leq n}$  be a  $n \times n$  symmetric random matrix such that the associated random vector

$$(G_{11}, \dots, G_{1n}, G_{21}, \dots, G_{2n}, \dots, G_{n1}, \dots, G_{nn})^T$$

of  $\mathbb{R}^{n^2}$  is Gaussian and has the same covariance matrix as

$$(W_{11}, \dots, W_{1n}, W_{21}, \dots, W_{2n}, \dots, W_{n1}, \dots, W_{nn})^T.$$

Then, for any  $n, d \geq 1$ , one has

$$\begin{aligned} d_{\text{Wass}}(\mathcal{W}_{n,d}^{(s)}, \mathcal{G}_{n,d}^{(s)}) &\leq \sqrt{\frac{192n^3}{\sum_{|k| \leq d} (1 - \frac{|k|}{d}) s(k)^2} \times \frac{1}{d} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^3} \\ &\leq \sqrt{\frac{192n^3}{d} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^3} \quad \left( \text{since } \sum_{|k| \leq d} (1 - \frac{|k|}{d}) s(k)^2 \geq s(0)^2 = 1 \right). \end{aligned} \quad (1.8)$$

Moreover, if  $s \in \ell^2(\mathbb{Z})$  and if we let  $\mathcal{Z}_n^{(s)} = (Z_{ij})_{1 \leq i, j \leq n}$  denote a  $n \times n$  symmetric random matrix such that  $Z_{ii} \sim N(0, 2\|s\|_{\ell^2(\mathbb{Z})}^2)$ ,  $Z_{ij} \sim N(0, \|s\|_{\ell^2(\mathbb{Z})}^2)$  for  $i < j$ ,  $Z_{ij} = Z_{ji}$  for  $i > j$ , and  $\{Z_{ij}, 1 \leq i < j \leq n\}$  independent, we have

$$d_{\text{Wass}}(\mathcal{G}_{n,d}^{(s)}, \mathcal{Z}_n^{(s)}) \leq \frac{2\sqrt{n(n+1)}}{\|s\|_{\ell^2(\mathbb{Z})}} \left( \sum_{|k| > d} s(k)^2 + \frac{1}{d} \sum_{|k| \leq d} |k| s(k)^2 \right). \quad (1.9)$$

**Remark 1.3.** (i) For instance, if  $s(r) = \mathbf{1}_{\{r=0\}}$ , that is, if the entries  $\{X_{ij} : i, j \geq 1\}$  of  $\mathcal{X}_{n,d}$  are i.i.d standard Gaussian, then we recover from Theorem 1.2 the result of [3, 9];  $\mathcal{W}_{n,d}$  given by (1.3) is close to  $\mathcal{Z}_n$  given by (1.4) when  $n^3/d \rightarrow 0$ .

(ii) If we assume that  $s \in \ell^{4/3}(\mathbb{Z})$ , then (1.8) leads to  $d_{\text{Wass}}(\mathcal{W}_{n,d}^{(s)}, \mathcal{G}_{n,d}^{(s)}) = O(\sqrt{n^3/d})$ . In this case,  $\mathcal{W}_{n,d}^{(s)}$  continues to be close to  $\mathcal{G}_{n,d}^{(s)}$  as soon as  $n^3/d \rightarrow 0$ , exactly like in the full independent case considered in [3, 4, 9]; see also (i).

As we just pointed out in the previous remark,  $\mathcal{W}_{n,d}^{(s)}$  and  $\mathcal{G}_{n,d}^{(s)}$  are asymptotically close as long as  $s \in \ell^{4/3}(\mathbb{Z})$  and  $n^3/d \rightarrow 0$ . What happens when  $s \in \ell^2(\mathbb{Z}) \setminus \ell^{4/3}(\mathbb{Z})$  or when  $s \notin \ell^2(\mathbb{Z})$ ? And how close are  $\mathcal{G}_{n,d}^{(s)}$  and  $\mathcal{Z}_n^{(s)}$ ? To exhibit an interesting situation where different behaviors may arise, starting from now on we will focus on the case where  $s$  is the correlation function of the fractional Brownian noise of Hurst index  $H \in (0, 1)$ , that is,

$$s(k) = s_H(k) = \frac{1}{2} (|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}), \quad k \in \mathbb{Z}. \quad (1.10)$$

When  $H \neq \frac{1}{2}$ , one has  $|s_H(k)| \sim c|k|^{2H-2}$  as  $|k| \rightarrow \infty$  (with  $c > 0$  a constant whose value is immaterial and may change from one instance to another). It is a straightforward exercise to check that  $s_H \in \ell^2(\mathbb{Z})$  if and only if  $H \in (0, \frac{3}{4})$ . Moreover, as  $d \rightarrow \infty$ ,

$$\begin{aligned} \frac{1}{d} \left( \sum_{|k| \leq d} |s_H(k)|^{4/3} \right)^3 &\sim c \begin{cases} 1/d & \text{if } 0 < H < 5/8 \\ (\log d)^3/d & \text{if } H = 5/8 \\ d^{8H-6} & \text{if } 5/8 < H < 1 \end{cases}; \\ \sum_{|k| > d} s_H(k)^2 &\sim c d^{4H-3} \quad \text{for } H \in (0, 1/2) \cup (1/2, 3/4); \\ \sum_{|k| \leq d} \left(1 - \frac{|k|}{d}\right) s_H(k)^2 &\sim c \log d \quad \text{for } H = 3/4; \\ \frac{1}{d} \sum_{|k| \leq d} |k| s_H(k)^2 &\sim c \begin{cases} 1/d & \text{if } 0 < H < 1/2 \\ d^{4H-3} & \text{if } 1/2 < H < 1 \end{cases}. \end{aligned}$$

As a result, for  $s_H$  given by (1.10), we deduce from (1.8) that  $\mathcal{W}_{n,d}^{(s_H)}$  is  $\phi$ -close to  $\mathcal{G}_{n,d}^{(s)}$  when  $H \in (0, \frac{3}{4}]$  and

$$\phi(n, d) = \frac{n^3}{d} \mathbf{1}_{\{0 < H < 5/8\}} + n^3 \frac{\log^3 d}{d} \mathbf{1}_{\{H=5/8\}} + n^3 d^{8H-6} \mathbf{1}_{\{5/8 < H < 3/4\}} + \frac{n^3}{\log d} \mathbf{1}_{\{H=3/4\}}. \quad (1.11)$$

On the other hand, when  $H \in (0, \frac{3}{4})$  (otherwise  $\mathcal{Z}_n^{(s)}$  is not defined), we deduce from (1.9) that  $\mathcal{G}_{n,d}^{(s)}$  is  $\psi$ -close to  $\mathcal{Z}_n^{(s)}$  for

$$\psi(n, d) = \frac{n}{d} \mathbf{1}_{\{0 < H < 1/2\}} + n d^{4H-3} \mathbf{1}_{\{1/2 < H < 3/4\}}. \quad (1.12)$$

We summarize the above discussion as follows.

**Corollary 1.4** (High-dimensional regime for fractional noise entries and Gaussian approximation). *Assume that  $s = s_H$  is given by (1.10) with  $H \in (0, \frac{3}{4}]$ . Then*

- (i) when  $H \in (0, 1/2)$ :  $\mathcal{W}_{n,d}^{(s_H)}$  is  $\phi$ -close to  $\mathcal{G}_{n,d}^{(s_H)}$  for  $\phi(n, d) = n^3/d$ , whereas  $\mathcal{G}_{n,d}^{(s_H)}$  is  $\psi$ -close to  $\mathcal{Z}_n^{(s_H)}$  for  $\psi(n, d) = n/d$ ;
- (ii) when  $H = 1/2$ :  $\mathcal{W}_{n,d}^{(s_H)}$  is  $\phi$ -close to  $\mathcal{G}_{n,d}^{(s_H)}$  for  $\phi(n, d) = n^3/d$ , whereas  $\mathcal{G}_{n,d}^{(s_H)}$  and  $\mathcal{Z}_n^{(s_H)}$  have the same law;
- (iii) when  $H \in (1/2, 5/8)$ :  $\mathcal{W}_{n,d}^{(s_H)}$  is  $\phi$ -close to  $\mathcal{G}_{n,d}^{(s_H)}$  for  $\phi(n, d) = n^3/d$ , whereas  $\mathcal{G}_{n,d}^{(s_H)}$  is  $\psi$ -close to  $\mathcal{Z}_n^{(s_H)}$  for  $\psi(n, d) = n d^{4H-3}$ ;
- (iv) for  $H = 5/8$ :  $\mathcal{W}_{n,d}^{(s_H)}$  is  $\phi$ -close to  $\mathcal{G}_{n,d}^{(s_H)}$  for  $\phi(n, d) = (n^3 \log^3 d)/d$ , whereas  $\mathcal{G}_{n,d}^{(s_H)}$  is  $\psi$ -close to  $\mathcal{Z}_n^{(s_H)}$  for  $\psi(n, d) = n/\sqrt{d}$ ;
- (v) for  $H \in (5/8, 3/4)$ :  $\mathcal{W}_{n,d}^{(s_H)}$  is  $\phi$ -close to  $\mathcal{G}_{n,d}^{(s_H)}$  for  $\phi(n, d) = n^3 d^{8H-6}$ , whereas  $\mathcal{G}_{n,d}^{(s_H)}$  is  $\psi$ -close to  $\mathcal{Z}_n^{(s_H)}$  for  $\psi(n, d) = n d^{4H-3}$ ;
- (vi) for  $H = 3/4$ :  $\mathcal{W}_{n,d}^{(s_H)}$  is  $\phi$ -close to  $\mathcal{G}_{n,d}^{(s_H)}$  for  $\phi(n, d) = n^3 (\log d)^{-1}$ , whereas  $\mathcal{Z}_n^{(s_H)}$  is not defined.

**Remark 1.5.** (1) Note that the particular case where  $H = 1/2$  reduces to the case of full independence and gives us the same high-dimensional regime as in [3, 9]. One of the main ingredients in the proof of Theorem 1.2 is the Stein's method developed in the work [14], in which the Malliavin calculus was coupled with the multivariate Stein's method in order to deal with the Wasserstein distance. This explains why, in the present paper, our bounds are for the Wasserstein distance and not the total variation distance  $d_{\text{TV}}$  as in [3, 4, 9]; see also Section 5.

(2) When  $H = 5/8$ , the high-dimensional regime looks similar to that derived in [4], that is, assuming the entries  $X_{ij}$  are i.i.d log-concave and  $(n^3 \log^2 d)/d \rightarrow 0$ , one has  $d_{\text{TV}}(\mathcal{W}_{n,d}, \mathcal{G}_n) \rightarrow 0$ , where  $\mathcal{G}_n$  belongs to the GOE. The log-terms in both regimes seem to be the price paid for deviating away from being Gaussian or independent, see also the critical case  $H = 3/4$ .

In the next result, we finally explain what happens in the case  $H \in (3/4, 1)$ . In this case, an interesting and new phenomenon appears: properly scaled, the Wishart matrix associated with  $\mathcal{X}_{n,d}$  given by (1.6) and  $s = s_H$  given by (1.10) converges to the so-called *Rosenblatt-Wishart matrix* for a suitable range of  $n$  and  $d$ .

**Theorem 1.6** (Rosenblatt approximation). *Consider  $s_H$  given by (1.10) with  $H \in (3/4, 1)$ , set*

$$\widehat{\mathcal{W}}_{n,d}^{(s_H)} = d^{2-2H} \left( \frac{1}{d} \mathcal{X}_{n,d}^{(s_H)} (\mathcal{X}_{n,d}^{(s_H)})^T - \mathcal{I}_n \right) \quad (1.13)$$

and let  $\mathcal{R}_n^{(H)}$  be the  $n \times n$  Rosenblatt-Wishart matrix with Hurst parameter  $H$  (see Definition 3.2). Then, there exists a finite constant  $c_H > 0$  depending only on  $H$  such that, for any  $n, d \geq 1$ ,

$$d_{\text{Wass}}(\widehat{\mathcal{W}}_{n,d}^{(s_H)}, \mathcal{R}_n^{(H)}) \leq c_H n d^{(3-4H)/2}.$$

In other words, the scaled Wishart matrix  $\widehat{\mathcal{W}}_{n,d}^{(s_H)}$  is  $\phi$ -close to the Rosenblatt-Wishart matrix  $\mathcal{R}_n^{(H)}$  for  $\phi(n, d) = n^2 d^{3-4H}$ .

The above theorem addresses the non-central limit theorems in the context of large random matrices and a crucial step in its proof is to construct explicitly a coupling of  $\widehat{\mathcal{W}}_{n,d}^{(s_H)}$  and  $\mathcal{R}_n^{(H)}$  using the *self-similarity* of fractional Brownian motion, with which we bound the Wasserstein distance by the  $L^2$ -distance.

The rest of the paper is organized as follows. In Section 2, we develop all the material needed for the proof of Theorem 1.2, and we give its proof in the end. Section 3 is devoted to the proof of Theorem 1.6 as well as the introduction to the new notion of Rosenblatt-Wishart matrix. In Section 4, we analyze the situation where the row-independence assumption is relaxed and we also look at the setting of random  $p$ -tensors ( $p \geq 3$ ), a natural extension of Wishart matrices. Finally, we propose some related open problems for future research in Section 5.

**Acknowledgement.** We thank Stéphane Chrétien for mentioning the paper [4] to one of us (Ivan), which initiated our investigation.



## 2. GAUSSIAN APPROXIMATION

The basic tools we use throughout this work is the Malliavin calculus and Stein's method, whose combination is commonly known as the Malliavin-Stein approach. Such an approach was motivated to quantify Nualart and Peccati's fourth moment theorem [17], and it has been extensively developed by the authors of [12] as well as their collaborators; see [13] for a comprehensive treatment. This Malliavin-Stein approach has turned out to be very applicable in quantifying limit theorems on a Gaussian space. More specifically, we are going to use its multidimensional version derived in [14] to investigate the high-dimensional regime concerning the Gaussian approximation of Wishart matrices. In Section 2.1, we collect several basic facts. Section 2.2 is devoted to the proof of Theorem 1.2. We refer the readers to the monograph [13] for any unexplained notation.

**2.1. Preparation of the proof of Theorem 1.2.** Let us first recall the framework put in the introduction:  $X = \{X(h), h \in \mathfrak{H}\}$  is an isonormal process over a real separable Hilbert space  $\mathfrak{H}$ , defined on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ .

For every  $p \geq 1$ , we let  $\mathcal{H}_p$  denote the  $p$ th Wiener chaos of  $X$ , that is, the closed linear subspace of  $L^2(\Omega)$  generated by the random variables of the form  $\{H_p(X(h)), h \in \mathfrak{H}, \|h\|_{\mathfrak{H}} = 1\}$ , where  $H_p$  stands for the  $p$ th Hermite polynomial<sup>3</sup>. The relation that  $I_p(h^{\otimes p}) = H_p(X(h))$  for unit vector  $h \in \mathfrak{H}$  can be extended to a linear isometry between the symmetric  $p$ th tensor product  $\mathfrak{H}^{\odot p}$  (equipped with the modified norm  $\sqrt{p!} \|\cdot\|_{\mathfrak{H}^{\otimes p}}$ ) and the  $p$ th Wiener chaos  $\mathcal{H}_p$ .

Suppose  $(h_i, i \geq 1)$  is an orthonormal basis of  $\mathfrak{H}$ , and consider  $f \in \mathfrak{H}^{\otimes p}$  and  $g \in \mathfrak{H}^{\otimes q}$  with  $p, q \geq 1$ . With  $f(i_1, \dots, i_p) = \langle f, h_{i_1} \otimes \dots \otimes h_{i_p} \rangle_{\mathfrak{H}^{\otimes p}}$  and  $g(i_1, \dots, i_q) = \langle g, h_{i_1} \otimes \dots \otimes h_{i_q} \rangle_{\mathfrak{H}^{\otimes q}}$ , we can express them as

$$f = \sum_{i_1, \dots, i_p=1}^{\infty} f(i_1, \dots, i_p) h_{i_1} \otimes \dots \otimes h_{i_p} \text{ and } g = \sum_{i_1, \dots, i_q=1}^{\infty} g(i_1, \dots, i_q) h_{i_1} \otimes \dots \otimes h_{i_q}. \quad (2.1)$$

For  $r \in \{1, \dots, p \wedge q\}$ , the  $r$ -contraction of  $f$  and  $g$  is the element in  $\mathfrak{H}^{\otimes p+q-2r}$  defined by

$$f \otimes_r g = \sum_{\substack{i_1, \dots, i_{p-r} \geq 1 \\ j_1, \dots, j_{q-r} \geq 1}} (f \star_r g)(i_1, \dots, i_{p-r}, j_1, \dots, j_{q-r}) h_{i_1} \otimes \dots \otimes h_{i_{p-r}} \otimes h_{j_1} \otimes \dots \otimes h_{j_{q-r}}$$

where

$$(f \star_r g)(i_1, \dots, i_{p-r}, j_1, \dots, j_{q-r}) = \sum_{k_1, \dots, k_r=1}^{\infty} f(k_1, \dots, k_r, i_1, \dots, i_{p-r}) g(k_1, \dots, k_r, j_1, \dots, j_{q-r}).$$

Contractions naturally appear in the product formula for multiple Wiener-Itô integrals: for  $f \in \mathfrak{H}^{\odot p}$  and  $g \in \mathfrak{H}^{\odot q}$  with  $p, q \geq 1$ , it holds that

$$I_p(f) I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \tilde{\otimes}_r g), \quad (2.2)$$

where  $f \tilde{\otimes}_r g$  stands for the symmetrization of  $f \otimes_r g$ ; see *e.g.* [13, Theorem 2.7.10].

<sup>3</sup> $H_1(x) = x$ ,  $H_2(x) = x^2 - 1$ ,  $H_3(x) = x^3 - 3x$  and  $H_{p+1}(x) = xH_p(x) - pH_{p-1}(x)$  for every  $p \geq 2$ .

We will also need the notion of Malliavin derivative  $D$  with respect to  $X$  but only its action on a fixed Wiener chaos: for  $f \in \mathfrak{H}^{\odot p}$  of the form (2.1), the Malliavin derivative of  $I_p(f)$  is the random element of  $\mathfrak{H}$  given by

$$DI_p(f) = p \sum_{i=1}^{\infty} I_{p-1}(f \otimes_1 h_i) h_i = p \sum_{i_1, \dots, i_p \geq 1} f(i_1, \dots, i_p) I_{p-1}(h_{i_2} \otimes \dots \otimes h_{i_p}) h_{i_1}.$$

Bearing all this in mind, let us go back to the rectangular matrix  $\mathcal{X}_{n,d}$  defined by (1.6). Since  $I_1(h) = X(h)$  for all  $h \in \mathfrak{H}$ , the entries of  $\mathcal{X}_{n,d}$  are realized as elements in the first Wiener chaos  $\mathcal{H}_1$ . As a consequence, due to either the very definition of  $\mathcal{H}_2$  (when  $i = j$ ) or the product formula (when  $i \neq j$ ), the  $(i, j)$ th entry  $W_{ij}$  of  $\mathcal{W}_{n,d}^{(s)}$  given by (1.7) belongs to the second Wiener chaos  $\mathcal{H}_2$ : more precisely,

$$W_{ij} = \left\{ \begin{array}{ll} \frac{1}{\sqrt{d}} \sum_{k=1}^d (X_{ik}^2 - 1) & \text{if } i = j \\ \frac{1}{\sqrt{d}} \sum_{k=1}^d X_{ik} X_{jk} & \text{if } i \neq j \end{array} \right\} = I_2(f_{ij}^{(d)}), \quad (2.3)$$

with the kernel

$$f_{ij}^{(d)} = \frac{1}{2\sqrt{d}} \sum_{k=1}^d (e_{ik} \otimes e_{jk} + e_{jk} \otimes e_{ik}).$$

Before we present the proof of Theorem 1.2, we prepare several important facts on double Wiener-Itô integrals.

**Fact 1.** For any  $f, g \in \mathfrak{H}^{\odot 2}$ , on has

$$\langle DI_2(f), DI_2(g) \rangle_{\mathfrak{H}} - \mathbb{E}[\langle DI_2(f), DI_2(g) \rangle_{\mathfrak{H}}] = 4I_2(f \tilde{\otimes}_1 g).$$

This can be verified by using the product formula (2.2).

**Fact 2.** For kernels  $f_{ij}^{(d)}$  given in (2.3), we have  $f_{ij}^{(d)} \otimes_1 f_{kl}^{(d)} = 0$ , whenever  $\{i, j\} \cap \{k, l\} = \emptyset$ . Here we may abuse the notation  $\{i, j\} = \{i\}$  if  $i = j$ . This fact follows from the specific shape of (1.5).

**Fact 3.** For kernels  $f_{ij}^{(d)}$  given in (2.3), we can obtain by following the same computations as in [13, Page 134-135] that

$$\begin{aligned} \|f_{ii}^{(d)} \otimes_1 f_{ii}^{(d)}\|_{\mathfrak{H}^{\otimes 2}}^2 &= \left\| \frac{1}{d} \sum_{k, \ell=1}^d (e_{ik} \otimes e_{i\ell}) s(k - \ell) \right\|_{\mathfrak{H}^{\otimes 2}}^2 \\ &= \frac{1}{d^2} \sum_{k, \ell, u, v=1}^d s(k - \ell) s(\ell - u) s(u - v) s(v - k) \leq \frac{1}{d} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^3 \end{aligned}$$



whereas, for  $i \neq j$ ,

$$\begin{aligned} \|f_{ij}^{(d)} \otimes_1 f_{ij}^{(d)}\|_{\mathfrak{H}^{\otimes 2}}^2 &= \left\| \frac{1}{4d} \sum_{k,\ell=1}^d (e_{ik} \otimes e_{i\ell} + e_{jk} \otimes e_{j\ell}) s(k-\ell) \right\|_{\mathfrak{H}^{\otimes 2}}^2 \\ &\leq \frac{1}{8} \|f_{ii}^{(d)} \otimes_1 f_{ii}^{(d)}\|_{\mathfrak{H}^{\otimes 2}}^2 + \frac{1}{8} \|f_{jj}^{(d)} \otimes_1 f_{jj}^{(d)}\|_{\mathfrak{H}^{\otimes 2}}^2 \\ &\leq \frac{1}{4d} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^3. \end{aligned} \quad (2.4)$$

Moreover, for any  $i, j, k, l$ , we have

$$\|f_{ij}^{(d)} \otimes_1 f_{kl}^{(d)}\|_{\mathfrak{H}^{\otimes 2}}^2 = \langle f_{ij}^{(d)} \otimes_1 f_{ij}^{(d)}, f_{kl}^{(d)} \otimes_1 f_{kl}^{(d)} \rangle_{\mathfrak{H}^{\otimes 2}} \leq \frac{1}{d} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^3,$$

where the equality above follows from the definition of contractions.

**Fact 4.** Finally, we state the main ingredient for our proof and we will only use it with  $p_1 = \dots = p_m = 1$  (in which case it provides a bound for the Wasserstein distance between two  $m$ -dimensional Gaussian vectors) and  $p_1 = \dots = p_m = 2$ .

**Proposition 2.1** (see Corollary 3.6 in [14]). *Fix integers  $m \geq 2$  and  $1 \leq p_1 \leq \dots \leq p_m$ . Consider a vector  $F = (F_1, \dots, F_m) = (I_{p_1}(f_1), \dots, I_{p_m}(f_m))$  with  $f_j \in \mathfrak{H}^{\odot p_j}$  for each  $j$ . On the other hand, let  $C$  be an invertible covariance matrix, and let  $Z \sim N_m(0, C)$ . Then*

$$d_{\text{Wass}}(F, Z) \leq \|C^{-1}\|_{\text{op}} \|C\|_{\text{op}}^{1/2} \left( \sum_{1 \leq i, j \leq m} \mathbb{E} \left[ (C_{ij} - p_j^{-1} \langle DF_i, DF_j \rangle_{\mathfrak{H}})^2 \right] \right)^{1/2},$$

where  $\|\cdot\|_{\text{op}}$  denotes the usual operator norm.

**2.2. Proof of Theorem 1.2.** We are now ready to give the proof of Theorem 1.2. It is divided into several steps.

*Step 1: passing from symmetric matrices to vectors.* Since the entries of  $\mathcal{W}_{n,d}^{(s)}$  are double Wiener-Itô integrals, we would like to apply Proposition 2.1. But Proposition 2.1 is stated for vectors, not for matrices. So, as a first step, we need to explain how we can reduce to vectors. If  $\mathcal{Z} = (Z_{ij})_{1 \leq i, j \leq n}$  is a  $n \times n$  random symmetric matrix, the notation  $\mathcal{Z}^{\text{half}}$  indicates the  $n(n+1)/2$ -dimensional random vector formed by the upper-triangular entries, namely:

$$\mathcal{Z}^{\text{half}} = (Z_{11}, Z_{12}, \dots, Z_{1n}, Z_{22}, Z_{23}, \dots, Z_{2n}, \dots, Z_{nn})^T. \quad (2.5)$$

**Lemma 2.2.** *Let  $\mathcal{X}$  and  $\mathcal{Y}$  be two symmetric random matrices of  $\mathcal{M}_n(\mathbb{R})$ . Then*

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

Here  $d_{\text{Wass}}(\mathcal{X}, \mathcal{Y})$  is defined according to (1.2), whereas  $d_{\text{Wass}}(\mathcal{X}^{\text{half}}, \mathcal{Y}^{\text{half}})$  stands for the Wasserstein distance between random variables with values in  $\mathbb{R}^{n(n+1)/2}$ , that is,

$$d_{\text{Wass}}(\mathcal{X}^{\text{half}}, \mathcal{Y}^{\text{half}}) = \sup \{ \mathbb{E}[g(\mathcal{X}^{\text{half}})] - \mathbb{E}[g(\mathcal{Y}^{\text{half}})] : \|g\|_{\text{Lip}} \leq 1 \},$$

where the  $\|g\|_{\text{Lip}}$  stands for the usual Lipschitz constant of a function  $g : \mathbb{R}^{n(n+1)/2} \rightarrow \mathbb{R}$ .

*Proof.* If  $x \in \mathbb{R}^{n(n+1)/2}$ , we define  $M_x$  to be the  $n \times n$  symmetric matrix such that  $M_x^{\text{half}} = x$ . Let  $g : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$  be 1-Lipschitz with respect to the Hilbert-Schmidt norm. We have

$$|\mathbb{E}[g(\mathcal{X})] - \mathbb{E}[g(\mathcal{Y})]| = \sqrt{2} |\mathbb{E}[\tilde{g}(\mathcal{X}^{\text{half}})] - \mathbb{E}[\tilde{g}(\mathcal{Y}^{\text{half}})]|,$$

where  $\tilde{g} : \mathbb{R}^{n(n+1)/2} \rightarrow \mathbb{R}$  is defined by  $\tilde{g}(x) = \frac{1}{\sqrt{2}}g(M_x)$ . Since

$$|\tilde{g}(x) - \tilde{g}(y)| = \frac{1}{\sqrt{2}} |g(M_x) - g(M_y)| \leq \frac{1}{\sqrt{2}} \|M_x - M_y\|_{\text{HS}} \leq \|x - y\|,$$

we deduce that  $|\mathbb{E}[g(\mathcal{X})] - \mathbb{E}[g(\mathcal{Y})]| \leq \sqrt{2} d_{\text{Wass}}(\mathcal{X}^{\text{half}}, \mathcal{Y}^{\text{half}})$ , thus concluding the proof by taking the supremum over  $g$ .  $\square$

*Step 2: estimating the operator norm.* Let us now look at the common covariance matrix  $C$  of  $(\mathcal{W}_{n,d}^{(s)})^{\text{half}}$  and  $(\mathcal{G}_{n,d}^{(s)})^{\text{half}}$ . It is diagonal with entries given by

$$\mathbb{E}[W_{ii}^2] = \frac{2}{d} \sum_{k,\ell=1}^d s(k-\ell)^2 \text{ for each } i \text{ and } \mathbb{E}[W_{ij}^2] = \frac{1}{d} \sum_{k,\ell=1}^d s(k-\ell)^2 \text{ for } i < j. \quad (2.6)$$

It follows immediately that

$$\|C\|_{\text{op}}^{1/2} \|C^{-1}\|_{\text{op}} = \sqrt{\frac{2d}{\sum_{k,\ell=1}^d s(k-\ell)^2}} = \sqrt{\frac{2}{\sum_{|j|\leq d} (1 - \frac{|j|}{d}) s(j)^2}}.$$

*Step 3: estimating the variance of  $\langle DW_{ij}, DW_{kl} \rangle_{\mathfrak{H}}$ .* The entries  $W_{ij} = I_2(f_{ij}^{(d)})$  being elements of second Wiener chaos, see (2.3), by **Fact 1** and isometry relation for multiple integrals we have

$$\text{Var}\left(\frac{1}{2} \langle DW_{ij}, DW_{kl} \rangle_{\mathfrak{H}}\right) = 8 \|f_{ij}^{(d)} \otimes_1 f_{kl}^{(d)}\|_{\mathfrak{H} \otimes_2}^2 \leq \frac{8}{d} \left( \sum_{|k|\leq d} |s(k)|^{4/3} \right)^3, \quad (2.7)$$

where the last inequality follows from **Fact 3**. Moreover, if  $\{i, j\} \cap \{k, \ell\} = \emptyset$ , **Fact 2** implies

$$\text{Var}\left(\frac{1}{2} \langle DW_{ij}, DW_{kl} \rangle_{\mathfrak{H}}\right) = 0. \quad (2.8)$$

*Step 4: proving (1.8).* Proposition 2.1 (with  $m = n(n+1)/2$  and  $p_1 = \dots = p_m = 2$ ) together with the conclusion of Step 2 give us the following bound

$$d_{\text{Wass}}((\mathcal{W}_{n,d}^{(s)})^{\text{half}}, (\mathcal{G}_{n,d}^{(s)})^{\text{half}}) \leq \sqrt{\frac{2}{\sum_{|j|\leq d} (1 - \frac{|j|}{d}) s(j)^2}} \left( \sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq k \leq \ell \leq n}} \text{Var}\left(\frac{1}{2} \langle DW_{ij}, DW_{kl} \rangle_{\mathfrak{H}}\right) \right)^{1/2}.$$

We deduce from (2.8) that the sum  $\sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq k \leq \ell \leq n}}$  in the previous inequality can be replaced by  $\sum_{(i,j,k,\ell) \in \mathcal{I}}$ , where the set  $\mathcal{I} := \{(i, j, k, \ell) \in \{1, \dots, n\}^4 : i, j, k, \ell \text{ are not mutually distinct}\}$

has the cardinality  $n^4 - n(n-1)(n-2)(n-3)$ , a quantity bounded by  $6n^3$ . Considering also the bound (2.7), we finally get

$$d_{\text{Wass}}((\mathcal{W}_{n,d}^{(s)})^{\text{half}}, (\mathcal{G}_{n,d}^{(s)})^{\text{half}}) \leq \sqrt{\frac{96}{\sum_{|j|\leq d} (1 - \frac{|j|}{d}) s(j)^2} \times \frac{n^3}{d} \left( \sum_{|k|\leq d} |s(k)|^{4/3} \right)^3}.$$

which, thanks to Lemma 2.2, gives exactly (1.8).

*Step 5: proving (1.9).* If  $C$  denotes this time the covariance matrix of  $(\mathcal{Z}_n^{(s)})^{\text{half}}$  (that is,  $C$  is diagonal with diagonal entries either equal to  $2\|s\|_{\ell^2(\mathbb{Z})}^2$  or  $\|s\|_{\ell^2(\mathbb{Z})}^2$ ), we have  $\|C\|_{\text{op}}^{1/2} = \sqrt{2}\|s\|_{\ell^2(\mathbb{Z})}$  and  $\|C^{-1}\|_{\text{op}} = \|s\|_{\ell^2(\mathbb{Z})}^{-2}$ . We deduce, according to Proposition 2.1 with  $m = n(n+1)/2$  and  $p_1 = \dots = p_m = 1$ , that

$$d_{\text{Wass}}((\mathcal{G}_{n,d}^{(s)})^{\text{half}}, (\mathcal{Z}_n^{(s)})^{\text{half}}) \leq \frac{\sqrt{2n(n+1)}}{\|s\|_{\ell^2(\mathbb{Z})}} \left( \sum_{|j|>d} s(j)^2 + \frac{1}{d} \sum_{|j|\leq d} |j|s(j)^2 \right).$$

Relying on Lemma 2.2 again, we obtain (1.9).  $\square$

### 3. ROSENBLATT APPROXIMATION

**3.1. Preliminaries on fractional Brownian motion.** We consider in this section a  $n$ -dimensional fractional Brownian motion with Hurst parameter  $H \in (0, 1)$ , that is, a centered Gaussian process  $B = \{B_t = (B_t^1, \dots, B_t^n); t \in \mathbb{R}_+\}$ , where  $B^1, \dots, B^n$  are  $n$  independent copies of a real fractional Brownian motion with covariance function  $R_H(s, t) = \frac{1}{2}(s^{2H} + t^{2H} - |t-s|^{2H})$ , for  $i = 1, \dots, n$ .

The two following fundamental properties of the fractional Brownian motion will be used throughout the sequel:

- it is *H-selfsimilar*, that is,  $(B_{ct})_{t \geq 0} \stackrel{\text{law}}{=} c^H (B_t)_{t \geq 0}$  for all  $c > 0$ ;
- it has *stationary increments*, that is,  $(B_{t+h} - B_h)_{t \geq 0} \stackrel{\text{law}}{=} (B_t)_{t \geq 0}$  for all  $h > 0$ .

We will also need a few facts about its Gaussian structure: let  $\mathcal{E}_n$  be the set of step-functions on  $\mathbb{R}_+$  with values in  $\mathbb{R}^n$  and consider the Hilbert space  $\mathfrak{H}_n$  defined as the closure of  $\mathcal{E}_n$  with respect to the scalar product induced by

$$\langle (\mathbf{1}_{[0,s_1]}, \dots, \mathbf{1}_{[0,s_n]}), (\mathbf{1}_{[0,t_1]}, \dots, \mathbf{1}_{[0,t_n]}) \rangle_{\mathfrak{H}_n} = \sum_{i=1}^n R_H(s_i, t_i). \quad (3.1)$$

Then the mapping  $(\mathbf{1}_{[0,t_1]}, \dots, \mathbf{1}_{[0,t_n]}) \in \mathcal{E}_n \mapsto \sum_{i=1}^n B_{t_i}^i$  can be extended to an isometry between  $\mathfrak{H}_n$  and the Gaussian space associated with  $B = (B^1, \dots, B^n)$ . We denote this isometry by  $\varphi \mapsto B(\varphi)$  and the process  $\{B(\varphi) : \varphi \in \mathfrak{H}_n\}$  is an isonormal Gaussian process by construction.

Eventually, for  $b > a \geq 0$  and  $i \in \{1, \dots, n\}$ , we will use the short-hand notation

$$\mathbf{1}_{[a,b]}^{i,n} := (0, \dots, 0, \mathbf{1}_{[a,b]}, 0, \dots, 0), \quad (3.2)$$

where the indicator function  $\mathbf{1}_{[a,b]}$  is located in the  $i$ th position.

**3.2. Rosenblatt-Wishart matrix.** In this section, we fix  $n \geq 1$  and we let  $\mathfrak{H}_n$  denote the Hilbert space constructed in Section 3.1, whose scalar product is defined by (3.1).

**Proposition 3.1.** *Fix  $1 \leq i, j \leq n$  and with the notation (3.2), consider*

$$f_{ij}^n(d) = \frac{d}{2} \sum_{p=0}^{d-1} \left\{ \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{i,n} \otimes \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{j,n} + \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{j,n} \otimes \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{i,n} \right\}, \quad d \geq 1. \quad (3.3)$$

Then  $\{f_{ij}^n(d)\}_{d \geq 1}$  is a Cauchy sequence in  $\mathfrak{H}_n^{\otimes 2}$ .

*Proof.* We first observe that, as  $d, d' \rightarrow \infty$ ,

$$\frac{1}{dd'} \sum_{p,q=0}^{d,d'} \left( dd' \int_{p/d}^{(p+1)/d} du \int_{q/d}^{(q+1)/d} dv |u-v|^{2H-2} \right)^2 \rightarrow \int_{[0,1]^2} |u-v|^{4H-4} dudv. \quad (3.4)$$

Now, let us compute  $\langle f_{ij}^n(d), f_{ij}^n(d') \rangle_{\mathfrak{H}_n^{\otimes 2}}$  for  $d, d' \geq 1$ :

$$\begin{aligned} \langle f_{ij}^n(d), f_{ij}^n(d') \rangle_{\mathfrak{H}_n^{\otimes 2}} &= \frac{dd'}{4} \sum_{p,q=0}^{d,d'} \langle \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{i,n} \otimes \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{j,n}, \mathbf{1}_{\left[\frac{q}{d'}, \frac{q+1}{d'}\right]}^{i,n} \otimes \mathbf{1}_{\left[\frac{q}{d'}, \frac{q+1}{d'}\right]}^{j,n} \rangle_{\mathfrak{H}_n^{\otimes 2}} \\ &+ \frac{dd'}{4} \sum_{p,q=0}^{d,d'} \langle \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{i,n} \otimes \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{j,n}, \mathbf{1}_{\left[\frac{q}{d'}, \frac{q+1}{d'}\right]}^{j,n} \otimes \mathbf{1}_{\left[\frac{q}{d'}, \frac{q+1}{d'}\right]}^{i,n} \rangle_{\mathfrak{H}_n^{\otimes 2}} \\ &+ \frac{dd'}{4} \sum_{p,q=0}^{d,d'} \langle \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{j,n} \otimes \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{i,n}, \mathbf{1}_{\left[\frac{q}{d'}, \frac{q+1}{d'}\right]}^{i,n} \otimes \mathbf{1}_{\left[\frac{q}{d'}, \frac{q+1}{d'}\right]}^{j,n} \rangle_{\mathfrak{H}_n^{\otimes 2}} \\ &+ \frac{dd'}{4} \sum_{p,q=0}^{d,d'} \langle \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{j,n} \otimes \mathbf{1}_{\left[\frac{p}{d}, \frac{p+1}{d}\right]}^{i,n}, \mathbf{1}_{\left[\frac{q}{d'}, \frac{q+1}{d'}\right]}^{j,n} \otimes \mathbf{1}_{\left[\frac{q}{d'}, \frac{q+1}{d'}\right]}^{i,n} \rangle_{\mathfrak{H}_n^{\otimes 2}}. \end{aligned}$$

We deduce

$$\begin{aligned} &\langle f_{ij}^n(d), f_{ij}^n(d') \rangle_{\mathfrak{H}_n^{\otimes 2}} \\ &= \left( \frac{1}{2} \mathbf{1}_{\{i \neq j\}} + \mathbf{1}_{\{i=j\}} \right) H^2 (2H-1)^2 \frac{1}{dd'} \sum_{p,q=0}^{d,d'} \left( dd' \int_{p/d}^{(p+1)/d} du \int_{q/d}^{(q+1)/d} dv |u-v|^{2H-2} \right)^2, \end{aligned}$$

implying in turn thanks to (3.4) that

$$\langle f_{ij}^n(d), f_{ij}^n(d') \rangle_{\mathfrak{H}_n^{\otimes 2}} \rightarrow \left( \frac{1}{2} \mathbf{1}_{\{i \neq j\}} + \mathbf{1}_{\{i=j\}} \right) H^2 (2H-1)^2 \int_{[0,1]^2} |u-v|^{4H-4} dudv \quad \text{as } d, d' \rightarrow \infty.$$

The proof of Proposition 3.1 is complete.  $\square$

We are now in a position to define the notion of Rosenblatt-Wishart matrix of size  $n$  with Hurst parameter  $H$ .

**Definition 3.2.** *For each  $1 \leq i, j \leq n$ , let  $g_{ij}^n \in \mathfrak{H}_n^{\otimes 2}$  be the limit of  $\{f_{ij}^n(d)\}_{d \geq 1}$  given in (3.3). The  $n \times n$  Rosenblatt-Wishart matrix with Hurst parameter<sup>4</sup>  $H$  is the random symmetric matrix  $\mathcal{R}_n^{(H)} = (R_{ij})_{1 \leq i, j \leq n}$  with its entries given by  $R_{ij} = I_2(g_{ij}^n)$ .*

<sup>4</sup>It is clear that the limiting kernels  $g_{ij}^n$ ,  $1 \leq i \leq j \leq n$ , depend on the Hurst parameter  $H$ .

Equivalently, one can also define  $\mathcal{R}_n^{(H)}$  as the entrywise  $L^2(\Omega)$ -limit (as  $d \rightarrow \infty$ ) of

$$\mathcal{S}_{n,d} = \begin{pmatrix} S_{11}(d) & \dots & S_{1n}(d) \\ \vdots & & \vdots \\ S_{n1}(d) & \dots & S_{nn}(d) \end{pmatrix}, \quad (3.5)$$

where

$$S_{ij}(d) = \begin{cases} d \sum_{p=0}^{d-1} (B_{\frac{p+1}{d}}^i - B_{\frac{p}{d}}^i)(B_{\frac{p+1}{d}}^j - B_{\frac{p}{d}}^j) & \text{if } i \neq j \\ d \sum_{p=0}^{d-1} \left\{ (B_{\frac{p+1}{d}}^i - B_{\frac{p}{d}}^i)^2 - d^{-2H} \right\} & \text{if } i = j \end{cases}. \quad (3.6)$$

Indeed, bearing in mind the notation (3.3) and (3.6), one obtains

$$S_{ij}(d) = I_2(f_{ij}^n(d)) \xrightarrow{L^2(\Omega)} R_{ij} \quad \text{as } d \rightarrow \infty, \text{ for any } 1 \leq i, j \leq n, \quad (3.7)$$

where the existence of the previous  $L^2(\Omega)$ -limit is a consequence of Proposition 3.1 and the isometry property of double Wiener-Itô integrals. It is clear from (3.7) and (3.6) that  $\mathcal{R}_n^{(H)}$  satisfies the following *compatibility* relation: if we delete the last row and last column of  $\mathcal{R}_{n+1}^{(H)}$ , then we obtain a matrix that is distributed as  $\mathcal{R}_n^{(H)}$ .

Another consequence of both (3.7) and the explicit expression of  $S_{ij}(d)$  is that the diagonal entries  $R_{ii}$  of the Rosenblatt-Wishart matrix are all independent from each other and distributed according to the Rosenblatt distribution. We refer the reader to the survey [20] and the references therein for the definition of the Rosenblatt distribution (one of them being precisely that it is the distributional limit of  $S_{11}(d)$  as  $d \rightarrow \infty$ ) together with its main properties (cumulants, characteristic function, etc.). For the non-diagonal entries, one first observes that  $(B^i, B^j) \stackrel{\text{law}}{=} (\frac{B^i+B^j}{\sqrt{2}}, \frac{B^i-B^j}{\sqrt{2}})$  as a process if  $i \neq j$ , so that

$$\{S_{ij}(d)\}_{d \geq 1} \stackrel{\text{law}}{=} \left\{ \frac{d}{2} \sum_{p=0}^{d-1} \left[ (B_{\frac{p+1}{d}}^i - B_{\frac{p}{d}}^i)^2 - d^{-2H} \right] - \frac{d}{2} \sum_{p=0}^{d-1} \left[ (B_{\frac{p+1}{d}}^j - B_{\frac{p}{d}}^j)^2 - d^{-2H} \right] \right\}_{d \geq 1},$$

implying in turn, by letting  $d$  go to infinity, that

$$R_{ij} \stackrel{\text{law}}{=} \frac{1}{2} [R_{ii} + R_{jj}]. \quad (3.8)$$

Note however that the previous identity in law (3.8) *only* holds for *fixed*  $i, j$ , that is, the corresponding identity in law at the matrix level *does not* hold true.

**3.3. Proof of Theorem 1.6.** We let the notation of Theorem 1.6 prevail, as well as the notation introduced in the previous section. Before bounding the Wasserstein distance between  $\widehat{\mathcal{W}}_{n,d}$  and  $\mathcal{R}_n^{(H)}$  (with  $\widehat{\mathcal{W}}_{n,d}$  given by (1.13) and  $\mathcal{R}_n^{(H)}$  being the  $n \times n$  Rosenblatt-Wishart matrix with Hurst parameter  $H$ ), we observe the following two facts:

(a) Given (1.5) and (1.10), one has  $\widehat{\mathcal{W}}_{n,d} \stackrel{\text{law}}{=} d^{1-2H} (W_{ij}(d))_{1 \leq i, j \leq n}$ , where

$$W_{ij}(d) = \begin{cases} \sum_{p=0}^{d-1} (B_{p+1}^i - B_p^i)(B_{p+1}^j - B_p^j) & \text{if } i \neq j \\ \sum_{p=0}^{d-1} \{(B_{p+1}^i - B_p^i)^2 - 1\} & \text{if } i = j \end{cases}.$$

By the *selfsimilarity* property of fractional Brownian motion, we deduce that  $\widehat{\mathcal{W}}_{n,d} \stackrel{\text{law}}{=} \mathcal{S}_{n,d}$ , with  $\mathcal{S}_{n,d}$  given by (3.5). As a result, with the notation of Definition 3.2, we get

$$d_{\text{Wass}}(\widehat{\mathcal{W}}_{n,d}, \mathcal{R}_n^{(H)}) = d_{\text{Wass}}(\mathcal{S}_{n,d}, \mathcal{R}_n^{(H)}).$$

(b) By its very definition, the Wasserstein distance is bounded by the  $L^2(\Omega)$ -distance, that is,

$$d_{\text{Wass}}(\mathcal{S}_{n,d}, \mathcal{R}_n^{(H)}) \leq \sqrt{\sum_{1 \leq i, j \leq n} \mathbb{E}[(S_{ij}(d) - R_{ij})^2]}. \quad (3.9)$$

We are thus left to estimate the right-hand side of (3.9). For this, we refer to [2]: in the inequality (17) therein, the existence of a finite constant  $c_H > 0$ , depending only on  $H$ , satisfying

$$\mathbb{E}[(S_{ij}(d) - R_{ij})^2] \leq c_H d^{3-4H}$$

is shown. Plugging this into (3.9) completes the proof of Theorem 1.6.  $\square$

## 4. FURTHER RESULTS

**4.1. Relaxing the row independence: overall correlation.** In this section, we consider a more general setting where we no longer assume the row independence. That is, the relation (1.5) will be replaced by a more general one, namely:

$$\langle e_{ij}, e_{i'j'} \rangle_{\mathcal{S}} = r(i - i')s(j - j') \quad (4.1)$$

where  $r$  is another correlation function also satisfying  $r(0) = 1$ . Recall the definition (1.3) of  $\mathcal{W}_{n,d} = (W_{ij})_{1 \leq i, j \leq n}$ . Since  $\mathbb{E}[X_{ik}X_{jk}] = r(i - j) \neq \mathbf{1}_{\{i=j\}}$  in general, its entries  $W_{ij}$  are no more centered in general, so we should modify the corresponding Gaussian ensemble by shifting a little bit. Equivalently, by keeping the corresponding Gaussian ensemble centered, we can modify the Wishart ensemble accordingly, that is, we will consider the following *shifted* Wishart matrix

$$\widetilde{\mathcal{W}}_{n,d} = (\widetilde{W}_{ij})_{1 \leq i, j \leq n} = (I_2(f_{ij}^{(d)}))_{1 \leq i, j \leq n} \quad (4.2)$$

with kernels  $f_{ij}^{(d)}$  given as in (2.3). That is,

$$\widetilde{W}_{ij} = \frac{1}{\sqrt{d}} \sum_{k=1}^d (X_{ik}X_{jk} - r(i - j)).$$

Now let  $\mathcal{G}_{n,d}^{(r,s)} = (G_{ij})_{1 \leq i,j \leq n}$  be the  $n \times n$  random symmetric matrix such that the associated random vector  $(G_{11}, \dots, G_{1n}, G_{21}, \dots, G_{2n}, \dots, G_{n1}, \dots, G_{nn})^T$  of  $\mathbb{R}^{n^2}$  is Gaussian and has the same covariance matrix as

$$(\widetilde{W}_{11}, \dots, \widetilde{W}_{1n}, \widetilde{W}_{21}, \dots, \widetilde{W}_{2n}, \dots, \widetilde{W}_{n1}, \dots, \widetilde{W}_{nn})^T.$$

It is again routine to check that for  $1 \leq i \leq j \leq n$  and  $1 \leq u \leq v \leq n$ ,

$$\mathbb{E}[G_{ij}G_{uv}] = \frac{r(i-u)r(v-j) + r(i-v)r(u-j)}{d} \sum_{k,\ell=1}^d s(k-\ell)^2, \quad (4.3)$$

and regardless of the integrability of  $r$ , the covariance in (4.3) is uniformly bounded by  $2\|s\|_{\ell^2(\mathbb{Z})}^2$ , since  $|r(k)| \leq 1$  for all  $k$  and

$$\frac{1}{d} \sum_{k,\ell=1}^d s(k-\ell)^2 = \sum_{|j| \leq d} \left(1 - \frac{|j|}{d}\right) s(j)^2. \quad (4.4)$$

However, it seems highly nontrivial to decide whether the covariance matrix of  $\mathcal{G}_{n,d}^{(r,s)}$  half is invertible or not. Therefore, we will not be able to apply Proposition 2.1 for the Gaussian approximation as we did in the proof of Theorem 1.2. Instead, we shall use the following bounds from [13, Theorem 6.1.2] and [16, Theorem 9.3], whose main interest for us is that the covariance matrix of the underlying Gaussian vector may not be invertible. The price to pay, however, is that one can no longer deal with the Wasserstein distance, and we have to replace it by a smoother distance.

**Proposition 4.1.** *Fix integers  $m \geq 2$  and  $1 \leq p_1 \leq \dots \leq p_m$ . Let  $F = (F_1, \dots, F_m)$  be a random vector such that  $F_i = I_{p_i}(f_i)$ , with some  $f_i \in \mathfrak{H}^{\circ p_i}$  for each  $i$ . Assume that  $Z$  is a centered Gaussian vector in  $\mathbb{R}^m$  with the same covariance matrix  $C$  as  $F$ . Then,*

(i) *for any  $h : \mathbb{R}^m \rightarrow \mathbb{R}$  belonging to  $C^2(\mathbb{R}^m)$  such that  $\|h''\|_\infty < +\infty$ , we have*

$$|\mathbb{E}[h(F)] - \mathbb{E}[h(Z)]| \leq \frac{1}{2} \|h''\|_\infty \sum_{i,j=1}^m \mathbb{E} \left[ |C(i,j) - p_i^{-1} \langle DF_i, DF_j \rangle| \right] \quad (4.5)$$

$$\leq \frac{m}{2} \|h''\|_\infty \sqrt{\sum_{i,j=1}^m \text{Var} \left( \frac{1}{p_i} \langle DF_i, DF_j \rangle \right)}, \quad (4.6)$$

where  $\|h''\|_\infty := \sup \left\{ \left| \frac{\partial^2 h}{\partial x_i \partial x_j} (x) \right| : x \in \mathbb{R}^m, 1 \leq i, j \leq m \right\}$ .

(ii) *for every  $h \in C^2(\mathbb{R}^m)$  with  $M_2(h) := \sup \{ \|D^2 h(x)\|_{\text{op}} : x \in \mathbb{R}^m \} < +\infty$ ,*

$$\left| \mathbb{E}[h(F) - h(Z)] \right| \leq \frac{\sqrt{m} M_2(h)}{2p_1} \left( \sum_{i,j=1}^m \text{Var}(\langle DF_i, DF_j \rangle) \right)^{1/2}.$$

<sup>5</sup>See (2.5) for the definition of the ‘half’ of a symmetric matrix.

<sup>6</sup>The equation (4.5) is clear from the proof of Theorem 6.1.2 in [13], while the inequality (4.6) follows easily from the Cauchy-Schwarz: notice that there is a typo in the display (6.1.3) of [13] and our version is correct.



**Remark 4.2.** With the above result, it is natural to consider the following distances

$$d_2(X, Y) := \sup_{\|h''\|_\infty \leq 1} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]| \quad \text{and} \quad \tilde{d}_2(X, Y) := \sup_{M_2(h) \leq 1} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|$$

for any  $m$ -dimensional random vectors  $X, Y$  with square-integrable components. Given  $h \in C^2(\mathbb{R}^m)$ , it is easy to check that  $\|h''\|_\infty \leq \sqrt{m}M_2(h)$  and  $M_2(h) \leq m\|h''\|_\infty$ , so if  $M_2(h) < +\infty$ , then  $h$  has at most quadratic growth so that the random variables  $h(X)$  and  $h(Y)$  are integrable. It follows from the previous discussion that

$$\frac{1}{\sqrt{m}} \tilde{d}_2(X, Y) \leq d_2(X, Y) \leq m \tilde{d}_2(X, Y).$$

Now we are ready to state the main result of this section.

**Theorem 4.3.** *Let the above notation prevail, and assume that  $s \in \ell^2(\mathbb{Z})$  and  $r(0) = s(0) = 1$ . Then, (recalling the ‘half’ notation from (2.5),)*

(1) *we have*

$$d_2((\tilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) = O\left\{ \frac{n^4}{\sqrt{d}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \right\}$$

and

$$\tilde{d}_2((\tilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) = O\left\{ \frac{n^3}{\sqrt{d}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \right\};$$

(2) *if in addition  $r \in \ell^2(\mathbb{Z})$ , we have*

$$d_2((\tilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) = O\left\{ \frac{n^{7/2}}{\sqrt{d}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \right\}$$

and

$$\tilde{d}_2((\tilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) = O\left\{ \frac{n^{5/2}}{\sqrt{d}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \right\};$$

(3) *if in addition  $r \in \ell^1(\mathbb{Z})$ , we have*

$$d_2((\tilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) = O\left\{ \frac{n^3}{\sqrt{d}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \right\}$$

and

$$\tilde{d}_2((\tilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) = O\left\{ \frac{n^{5/2}}{\sqrt{d}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \right\}.$$

*Proof.* For  $i, j, p, q \in \{1, \dots, n\}$ , we have, similarly to (2.4), that

$$\begin{aligned}
 & \left\| f_{ij}^{(d)} \otimes_1 f_{pq}^{(d)} \right\|_{\mathfrak{H}^{\otimes 2}}^2 \\
 &= \frac{1}{16d^2} \left\| \sum_{k, \ell=1}^d (e_{ik} \otimes e_{jk} + e_{jk} \otimes e_{ik}) \otimes_1 (e_{p\ell} \otimes e_{q\ell} + e_{q\ell} \otimes e_{p\ell}) \right\|_{\mathfrak{H}^{\otimes 2}}^2 \\
 &= \frac{1}{16d^2} \left\| \sum_{k, \ell=1}^d \left( (e_{ik} \otimes e_{p\ell})r(j-q)s(k-\ell) + (e_{ik} \otimes e_{q\ell})r(p-j)s(k-\ell) \right. \right. \\
 & \quad \left. \left. + (e_{jk} \otimes e_{p\ell})r(i-q)s(k-\ell) + e_{jk} \otimes e_{q\ell}r(i-p)s(k-\ell) \right) \right\|_{\mathfrak{H}^{\otimes 2}}^2 \\
 &= \frac{\mathfrak{X}_{i,j,p,q}}{16d^2} \sum_{k, \ell, u, v=1}^d s(k-u)s(\ell-v)s(k-\ell)s(u-v) \leq \frac{\mathfrak{X}_{i,j,p,q}}{16d} \left( \sum_{|k| \leq d} |s(k)|^{\frac{4}{3}} \right)^3, \quad (4.7)
 \end{aligned}$$

where the last inequality follows from **Fact 3** in Section 2.1 and  $\mathfrak{X}_{i,j,p,q}$  is a sum of the following sixteen terms:

$$\begin{aligned}
 & r(j-q)^2 + r(j-q)r(p-j)r(p-q) + r(j-i)r(q-j)r(i-q) \\
 & + r(i-j)r(p-q)r(i-p)r(j-q) + r(p-j)^2 + r(p-q)r(j-q)r(p-j) \\
 & + r(i-j)r(p-q)r(i-q)r(p-j) + r(i-j)r(i-p)r(p-j) + r(i-q)^2 + r(i-p)^2 \\
 & + r(i-q)r(i-j)r(j-q) + r(i-q)r(i-j)r(p-q)r(p-j) + r(i-q)r(p-q)r(i-p) \\
 & + r(i-p)r(i-j)r(p-q)r(j-q) + r(i-p)r(i-j)r(p-j) + r(i-p)r(p-q)r(q-i).
 \end{aligned}$$

Using the fact that  $|r(k)| \leq 1$  for each  $k \in \mathbb{Z}$  and  $2ab \leq a^2 + b^2$  for  $a, b \in \mathbb{R}$ , we have the following estimate:

$$\mathfrak{X}_{i,j,p,q} = O\left(r(j-q)^2 + r(i-j)^2 + r(i-p)^2 + r(i-q)^2 + r(j-p)^2 + r(p-q)^2\right), \quad (4.8)$$

while the following two crude estimates also hold:

$$\mathfrak{X}_{i,j,p,q} = O\left[|r(j-q)| + |r(i-j)| + |r(i-p)| + |r(i-q)| + |r(j-p)| + |r(p-q)|\right], \quad (4.9)$$

$$|\mathfrak{X}_{i,j,p,q}| \leq 16. \quad (4.10)$$

Note that if  $r \in \ell^2(\mathbb{Z})$ ,

$$\sum_{i,j,p,q=1}^n r(j-q)^2 = n^2 \sum_{1 \leq j,q \leq n} r(j-q)^2 = n^3 \frac{1}{n} \sum_{k,\ell=1}^n r(k-\ell)^2 \leq n^3 \|r\|_{\ell^2(\mathbb{Z})}^2, \quad (4.11)$$

where the last inequality follows from (4.4). The same argument will give us that under the assumption  $r \in \ell^1(\mathbb{Z})$ ,

$$\sum_{i,j,p,q=1}^n |r(j-q)| \leq n^3 \|r\|_{\ell^1(\mathbb{Z})}. \quad (4.12)$$

Now let us show the bounds in the assertion (1): it follows first from Proposition 4.1-(i), then from **Fact 1** and isometry relation that

$$\begin{aligned} d_2((\widetilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) &\leq \frac{1}{2} \sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \mathbb{E}[|2I_2(f_{ij}^{(d)} \otimes_1 f_{pq}^{(d)})|] \\ &\leq \sqrt{2} \sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \|f_{ij}^{(d)} \otimes_1 f_{pq}^{(d)}\|_{\mathfrak{H}^{\otimes 2}}. \end{aligned} \quad (4.13)$$

Then the crude estimate (4.10) and the bound (4.7) imply that

$$d_2((\widetilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) \leq \frac{n^4}{\sqrt{d}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2}.$$

And similarly, we can apply Proposition 4.1-(ii) to get

$$\begin{aligned} \widetilde{d}_2((\widetilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) &\leq \frac{n}{4} \sqrt{\sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \text{Var}(4I_2(f_{ij}^{(d)} \otimes_1 f_{pq}^{(d)}))} \\ &\leq \sqrt{2}n \sqrt{\sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \|f_{ij}^{(d)} \otimes_1 f_{pq}^{(d)}\|_{\mathfrak{H}^{\otimes 2}}^2} \\ &\leq \sqrt{2}n \sqrt{\frac{n^4}{d} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^3}. \end{aligned} \quad (4.14)$$

Now let us assume additionally that  $r \in \ell^2(\mathbb{Z})$ , then with similar arguments and using in particular the estimates (4.8), (4.11), we have

$$\begin{aligned} &d_2((\widetilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) \\ &\leq \frac{n^2}{2} \sqrt{\sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \|f_{ij}^{(d)} \otimes_1 f_{pq}^{(d)}\|_{\mathfrak{H}^{\otimes 2}}^2} = O(n^2) \sqrt{\sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \frac{\mathfrak{X}_{i,j,p,q}}{16d} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^3} \\ &= O(n^2) \sqrt{\frac{n^3 \|r\|_{\ell^2(\mathbb{Z})}^2}{d} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^3} \quad \text{by (4.11);} \end{aligned}$$

and

$$\begin{aligned} \tilde{d}_2((\widetilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) &\leq \sqrt{2}n \sqrt{\sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \|f_{ij}^{(d)} \otimes_1 f_{pq}^{(d)}\|_{\mathfrak{H}^{\otimes 2}}^2} \quad \text{see (4.14);} \\ &= O \left\{ \frac{n^{\frac{5}{2}}}{\sqrt{d}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{\frac{3}{2}} \right\}. \end{aligned}$$

Now let us assume additionally that  $r \in \ell^1(\mathbb{Z})$ , then with similar arguments and using in particular the estimates (4.10), (4.12), we have

$$\begin{aligned} d_2((\widetilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) &\leq \sqrt{2} \sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \|f_{ij}^{(d)} \otimes_1 f_{pq}^{(d)}\|_{\mathfrak{H}^{\otimes 2}} \quad \text{see (4.13);} \\ &\leq \sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \frac{\sqrt{\mathfrak{X}_{i,j,p,q}}}{\sqrt{d}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \quad \text{by (4.7);} \\ &= O \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \times \frac{1}{\sqrt{d}} \sum_{i,j,p,q=1}^n |r(j-q)| \quad \text{by (4.8);} \\ &= O \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \times \frac{n^3}{\sqrt{d}} \|r\|_{\ell^1(\mathbb{Z})} \quad \text{by (4.12)}. \end{aligned}$$

Finally,

$$\begin{aligned} \tilde{d}_2((\widetilde{\mathcal{W}}_{n,d})^{\text{half}}, (\mathcal{G}_{n,d}^{(r,s)})^{\text{half}}) &\leq \sqrt{2}n \sqrt{\sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \|f_{ij}^{(d)} \otimes_1 f_{pq}^{(d)}\|_{\mathfrak{H}^{\otimes 2}}^2} \quad \text{see (4.14);} \\ &= \sqrt{2}n \sqrt{\sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq p \leq q \leq n}} \frac{\mathfrak{X}(i,j,p,q)}{d} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^3} \quad \text{by (4.7);} \\ &= O(n) \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \times \sqrt{\frac{1}{d} \sum_{i,j,p,q=1}^n |r(i-j)|} \quad \text{by (4.9);} \\ &= O(n) \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/2} \times \sqrt{\frac{n^3}{d} \|r\|_{\ell^1(\mathbb{Z})}} \quad \text{by (4.12)}. \quad \square \end{aligned}$$

Keeping in mind that our goal is to obtain high-dimensional regime for the distributional convergence, we compare the distances  $d_2, \tilde{d}_2$  with the Wasserstein distance in the following remark.

**Proposition 4.4.** *Let  $X, Y$  be two random vectors in  $\mathbb{R}^m$  with square-integrable components. Then, by standard smoothing argument, we have*

$$d_{\text{Wass}}(X, Y) \leq 2\sqrt{2} m^{1/4} \sqrt{d_2(X, Y)}. \quad (4.15)$$

*Proof.* Indeed, consider any  $h : \mathbb{R}^m \rightarrow \mathbb{R}$  1-Lipschitz function. Define, for  $\varepsilon > 0$ , the smoothed version  $h_\varepsilon(x) := \mathbb{E}[h(x + \varepsilon N)]$  for every  $x \in \mathbb{R}^m$ , with  $N$  standard Gaussian independent of  $X$  and  $Y$ . It is routine to check via a simple Gaussian integration by parts that for each  $i, j \in \{1, \dots, m\}$

$$\frac{\partial^2}{\partial x_i \partial x_j} h_\varepsilon(x) = \frac{1}{\varepsilon} \mathbb{E} \left[ N_i \frac{\partial}{\partial x_j} h(x + \varepsilon N) \right],$$

from which we obtain  $\|h_\varepsilon''\| \leq \varepsilon^{-1}$ . Therefore, we have  $|\mathbb{E}[h_\varepsilon(X) - h_\varepsilon(Y)]| \leq \varepsilon^{-1} d_2(X, Y)$  while  $|\mathbb{E}[h(X) - h_\varepsilon(X)]| \leq \mathbb{E}[|h(X) - h(X + \varepsilon N)|] \leq \varepsilon \mathbb{E}[\|N\|_{\mathbb{R}^m}] \leq \varepsilon \sqrt{m}$ , thus we have

$$|\mathbb{E}[h(X) - h(Y)]| \leq |\mathbb{E}[h(X) - h_\varepsilon(X)]| + |\mathbb{E}[h_\varepsilon(X) - h_\varepsilon(Y)]| + |\mathbb{E}[h_\varepsilon(Y) - h(Y)]| \quad (4.16)$$

$$\leq 2\varepsilon \sqrt{m} + \varepsilon^{-1} d_2(X, Y) \text{ for any } \varepsilon > 0. \quad (4.17)$$

Optimizing over  $\varepsilon > 0$  in (4.17) first, then taking supremum on the left-hand side of (4.16) give us (4.15).  $\square$

As an easy consequence of Proposition 4.4 combined with Lemma 2.2 and Theorem 4.3, we have the following bounds in Wasserstein distance.

**Corollary 4.5.** *Let the assumptions of Theorem 4.3 prevail, that is,  $s \in \ell^2(\mathbb{Z})$  and  $r(0) = s(0) = 1$ . Then,*

$$d_{\text{Wass}}(\widetilde{\mathcal{W}}_{n,d}, \mathcal{G}_{n,d}^{(r,s)}) = \begin{cases} O \left\{ \frac{n^{5/2}}{d^{1/4}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/4} \right\}; \\ O \left\{ \frac{n^{9/4}}{d^{1/4}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/4} \right\} & \text{if in addition } r \in \ell^2(\mathbb{Z}); \\ O \left\{ \frac{n^2}{d^{1/4}} \left( \sum_{|k| \leq d} |s(k)|^{4/3} \right)^{3/4} \right\} & \text{if in addition } r \in \ell^1(\mathbb{Z}). \end{cases}$$

In particular, if in addition  $s \in \ell^{4/3}(\mathbb{Z})$ , then  $\widetilde{\mathcal{W}}_{n,d}$  is  $\phi$ -close to  $\mathcal{G}_{n,d}^{(r,s)}$  for

$$\phi(n, d) = \begin{cases} n^{10}/d; \\ n^9/d & \text{if in addition } r \in \ell^2(\mathbb{Z}); \\ n^8/d & \text{if in addition } r \in \ell^1(\mathbb{Z}). \end{cases}$$

**4.2. Random  $p$ -tensors.** In this section, we investigate the Gaussian approximation of random  $p$ -tensors, a natural extension of the Wishart matrices. The distribution of the random  $p$ -tensors has recently gained interest in the context of machine learning, see for instance [1].

Let us suppose  $\{\varepsilon_i, i = 1, \dots, n\}$  is the canonical basis of  $\mathbb{R}^n$ . Then for  $p \geq 2$ , the  $p$ -tensor product of  $x \in \mathbb{R}^n$  is given by

$$x^{\otimes p} = \left( \sum_{i=1}^n \langle x, \varepsilon_i \rangle_{\mathbb{R}^n} \varepsilon_i \right)^{\otimes p} = \sum_{i_1, \dots, i_p=1}^n \left( \prod_{k \in [p]} \langle x, \varepsilon_{i_k} \rangle_{\mathbb{R}^n} \right) \varepsilon_{i_1} \otimes \dots \otimes \varepsilon_{i_p},$$

which can be identified as a vector in  $\mathbb{R}^{n^p}$ . For what we care most in the present work,  $\mathbb{X}_i = (X_{1i}, \dots, X_{ni})^T \in \mathbb{R}^n$  is the  $i$ th column of the rectangular random matrix  $\mathcal{X}_{n,d}$ , then

$$\mathbb{X}_i^{\otimes p} = \sum_{j_1, \dots, j_p=1}^n \left( \prod_{k=1}^p X_{j_k i} \right) \varepsilon_{j_1} \otimes \dots \otimes \varepsilon_{j_p}$$

so that

$$\frac{1}{\sqrt{d}} \sum_{i=1}^d \mathbb{X}_i^{\otimes p} = \sum_{j_1, \dots, j_p=1}^n \frac{1}{\sqrt{d}} \sum_{i=1}^d \left( \prod_{k=1}^p X_{j_k i} \right) \varepsilon_{j_1} \otimes \dots \otimes \varepsilon_{j_p},$$

which can also be seen as the random vector  $\mathbf{Y} \in \mathbb{R}^{n^p}$  given below

$$\left( \mathbf{Y}_{\mathbf{j}} = \frac{1}{\sqrt{d}} \sum_{i=1}^d \prod_{k=1}^p X_{j_k i}, 1 \leq j_1, \dots, j_p \leq n \right).$$

Note that 2-tensor case corresponds to the matrix, which we have dealt with in the previous sections. To simplify the matter and without losing the essence, we will remove the diagonal terms, that is, we will only consider the Gaussian approximation of

$$\mathcal{Y}_{n,d} = \left( \mathbf{Y}_{\mathbf{j}} := \frac{1}{\sqrt{d}} \sum_{i=1}^d \prod_{k=1}^p X_{j_k i}, \mathbf{j} \in \Delta_p \right), \quad (4.18)$$

where  $\Delta_p := \{\{j_1, \dots, j_p\} \in \{1, \dots, n\}^p : j_1, \dots, j_p \text{ are mutually distinct}\}$ .

Although our approach can extend to a more general setting, we focus on the full independence case to illustrate the ideas. That is, from now on, we assume that

*the entries of  $\mathcal{X}_{n,d}$  are i.i.d. standard Gaussian.*

And in the following result, we present the high-dimensional regime (as  $n, d$  both tend to infinity), in which the random  $p$ -tensor (4.18) is close to a Gaussian distribution.

**Theorem 4.6.** *Fix an integer  $p \geq 2$  and let the above assumptions and notation prevail. Let  $Z_\nu$  be a standard Gaussian vector in  $\mathbb{R}^\nu$  with  $\nu = p! \binom{n}{p}$ . Then we have*

$$d_{\text{Wass}}(\mathcal{Y}_{n,d}, Z_\nu) = O(\sqrt{n^{2p-1}/d}). \quad (4.19)$$

*In particular,  $\mathcal{Y}_{n,d}$  is close to a standard Gaussian vector with the same dimension as soon as  $n^{2p-1}/d \rightarrow 0$ .*

*Proof.* We split the proof into two main steps. The first step is to restrict our attention to the random vector in  $\mathbb{R}^{\nu/p!}$

$$\mathcal{Y}_{n,d}^\uparrow = (\mathbf{Y}_{\mathbf{j}}, \mathbf{j} \in \Delta_p^\uparrow) \quad (4.20)$$

and the standard Gaussian vector  $Z_\nu^\uparrow$  in  $\mathbb{R}^{\nu/p!}$ , here  $\Delta_p^\uparrow = \{\mathbf{j} \in \Delta_p : j_1 < j_2 < \dots < j_p\}$ .

Step 1: *Gaussian approximation of  $\mathcal{Y}_{n,d}^\uparrow$ .* This step can be established easily using the multivariate bound in Proposition 2.1. To do so, we first write for  $\mathbf{j} \in \Delta_p^\uparrow$ ,

$$Y_{\mathbf{j}} = \frac{1}{\sqrt{d}} \sum_{i=1}^d \prod_{k=1}^p X_{j_k i} = I_p(f_{\mathbf{j}}^{(d)})$$

with the kernel  $f_{\mathbf{j}}^{(d)} = \frac{1}{\sqrt{d}} \sum_{i=1}^d \text{sym}(e_{j_1 i} \otimes \cdots \otimes e_{j_p i})$ , where  $\text{sym}(\cdot)$  stands for the canonical symmetrization. Then for any  $r \in \{1, \dots, p-1\}$  and  $\mathbf{j}, \mathbf{j}' \in \Delta_p^\uparrow$ , it is immediate to verify by using the orthogonality of  $\{e_{ij}, i, j \geq 1\}$  that

$$\|f_{\mathbf{j}}^{(d)} \otimes_r f_{\mathbf{j}'}^{(d)}\|_{\mathfrak{S}^{\otimes 2p-2r}}^2 = O(1/d)$$

and  $\|f_{\mathbf{j}}^{(d)} \otimes_r f_{\mathbf{j}'}^{(d)}\|_{\mathfrak{S}^{\otimes 2p-2r}}^2 = 0$  if additionally  $\mathbf{j}$  and  $\mathbf{j}'$  have no common index (*i.e.*  $\mathbf{j} \cap \mathbf{j}' = \emptyset$ ).

Therefore, in view of the equation (6.2.3) in [13], we have

$$\text{Var}\left(\langle DI_p(f_{\mathbf{j}}^{(d)}), DI_p(f_{\mathbf{j}'}^{(d)}) \rangle_{\mathfrak{S}}\right) = O\left(\sum_{r=1}^{p-1} \|f_{\mathbf{j}}^{(d)} \otimes_r f_{\mathbf{j}'}^{(d)}\|_{\mathfrak{S}^{\otimes 2p-2r}}^2\right) = O(1/d)$$

and  $\text{Var}(\langle DI_p(f_{\mathbf{j}}^{(d)}), DI_p(f_{\mathbf{j}'}^{(d)}) \rangle_{\mathfrak{S}}) = 0$  if additionally  $\mathbf{j} \cap \mathbf{j}' = \emptyset$ . Thus, it follows from Proposition 2.1 that

$$d_{\text{Wass}}(\mathcal{Y}_{n,d}^\uparrow, Z_\nu^\uparrow) \leq \left( \sum_{\mathbf{j}, \mathbf{j}' \in \Delta_p^\uparrow} \text{Var}\left(p^{-1} \langle DI_p(f_{\mathbf{j}}^{(d)}), DI_p(f_{\mathbf{j}'}^{(d)}) \rangle_{\mathfrak{S}}\right) \right)^{1/2}$$

and the sum  $\sum_{\mathbf{j}, \mathbf{j}' \in \Delta_p^\uparrow}$  can be replaced by the sum  $\sum_{\mathbf{j}, \mathbf{j}' \in \Delta_p^\uparrow: \mathbf{j} \cap \mathbf{j}' \neq \emptyset}$ . It is easy to see that the cardinality of the set  $\{(j_1, \dots, j_p, j'_1, \dots, j'_p) \in \{1, \dots, n\}^{2p} : \mathbf{j}, \mathbf{j}' \in \Delta_p^\uparrow \text{ and } \mathbf{j} \cap \mathbf{j}' \neq \emptyset\}$  is  $O(n^{2p-1})$ . Hence we conclude our Step 1 with  $d_{\text{Wass}}(\mathcal{Y}_{n,d}^\uparrow, Z_\nu^\uparrow) = O(\sqrt{n^{2p-1}/d})$ .

Step 2: *The passage from  $\mathcal{Y}_{n,d}^\uparrow$  (4.20) to  $\mathcal{Y}_{n,d}$  (4.18).* This can be done by using an easy extension of Lemma 2.2. For the sake of completeness, we sketch the main arguments below (Recall  $\nu = p! \binom{n}{p}$ ).

- Consider  $\mathbf{x} = (\mathbf{x}_{\mathbf{j}}, \mathbf{j} \in \Delta_p)$ , let  $h : \mathbb{R}^\nu \rightarrow \mathbb{R}$  be 1-Lipschitz;
- set  $h^\uparrow(\mathbf{x}^\uparrow) = \frac{1}{\sqrt{p!}} h(\mathbf{x})$  for any  $\mathbf{x} = (\mathbf{x}_{\mathbf{j}}, \mathbf{j} \in \Delta_p)$  and it is easy to verify that  $h^\uparrow$  is a 1-Lipschitz function defined on  $\mathbb{R}^{\nu/p!}$ .

Then  $|\mathbb{E}[h(\mathcal{Y}_{n,d}) - h(Z_\nu)]| = \sqrt{p!} |\mathbb{E}[h^\uparrow(\mathcal{Y}_{n,d}^\uparrow) - h^\uparrow(Z_\nu^\uparrow)]| \leq \sqrt{p!} d_{\text{Wass}}(\mathcal{Y}_{n,d}^\uparrow, Z_\nu^\uparrow) = O(\sqrt{n^{2p-1}/d})$ , hence the desired bound (4.19) follows immediately.  $\square$

Note that the above theorem is a substantial extension of the regime in [3, 9].

## 5. CONCLUSION AND OPEN PROBLEMS

In our work, we answered some questions meanwhile we also raised some questions. To motivate further research, we provide a summary in this section.



In a large complex system formed by many independent components, many universal pictures can appear, for example the classical central limit theorem or Gaussian fluctuation is one of them. In this work, the large complex object we are considering is the Wishart matrix  $\mathcal{W}_{n,d}$  (1.3), which appears naturally as the sample covariance matrix in the context of multivariate analysis. It is also closely related to the so-called “*principal component analysis*”, see [10]. Several previous papers [3, 4, 9] have been devoted to the high dimensional limit and their methodology consists of random matrix techniques or information-theoretical tools. However, these articles only considered the case of full independence, that is, the entries of the rectangular random matrix  $\mathcal{X}_{n,d}$  that forms  $\mathcal{W}_{n,d}$  are independent. Such a setting gives arise to several advantages, for example,

- in the Gaussian setting, the authors of [3] were able to directly compute the total variation distance between  $\mathcal{W}_{n,d}$  and corresponding Gaussian ensemble using the available density formulae;
- the authors of [4] were able to perform an induction argument in order to use the entropic CLT as well as some other tools to bound the total variation distance.

The lack of independence breaks the above strategies, while it is well known that Stein’s method of distributional approximation is very powerful for investigating situations in presence of dependence. This motivated us to apply Stein’s method for studying the high-dimensional limit of large Gaussian correlated Wishart matrices. In the present work, we not only recover known high-dimensional regimes but also provide new phenomena in the correlated case, see our Theorem 1.2 and many interesting examples in Corollary 1.4. We are also able to deal with non-central limit in high-dimensional regime, see Theorem 1.6, where notably a new probabilistic object called the Rosenblatt-Wishart matrix shows up as the highlight in our Section 3. Our bounds are described in terms of the Wasserstein distance, different from the total variation distance in [3, 4]. As we simply consider the symmetric matrix of size  $n$  as the  $n^2$ -dimensional vector, the random vector  $\mathcal{W}_{n,d}$  has many repeated components so that it has no density on  $\mathbb{R}^{n^2}$ , preventing us from obtaining the bounds in the total variation distance; on the other hand, (even for the Gaussian approximation of  $\mathcal{W}_{n,d}^{\text{half}}$ ) it is extremely difficult to get a *multivariate* total variation bound by using Stein’s method.

In Section 4, we obtain high-dimensional regimes in the case of overall correlation. It covers the case of column independence where the column vectors  $\mathbb{X}_i = (X_{1i}, \dots, X_{ni})^T$ ,  $i = 1, \dots, d$ , of the rectangular random matrix  $\mathcal{X}_{n,d}$  are independent and identically distributed. As pointed out in the paper [4], a straightforward application of the Stein’s method (see *e.g.*[6]) gives us the following bound in the *full independence case*:

$$\tilde{d}_2(\mathcal{W}_{n,d}, \mathcal{Z}_n) = O(n^3/\sqrt{d}), \quad (5.1)$$

where  $\mathcal{W}_{n,d}$  is given by (1.3) (with general entry distribution) and  $\mathcal{Z}_n$  belongs to the GOE described in (1.4). The authors of [4] then stated an intriguing open question: whether there is a way to use Stein’s method to recover their regime ( $d$  much larger than  $n^3$ ) in any reasonable metric (total variation metric, Wasserstein metric, etc.)? So in the present paper, we answered this open question in the *Gaussian setting* and moreover, what we achieved is beyond the full independence case: as already mentioned in [4], the induction argument therein (at the core of their strategy) breaks without the full independence assumption.

It is also worth mentioning that the authors of [4] claimed that the case of row independence is probably much harder than column independence. This claim is reasonable in view of how they applied Stein's method to get (5.1): they first express the Wishart matrix as the normalized partial sum of i.i.d 2-tensors

$$\mathcal{W}_{n,d} = \frac{1}{\sqrt{d}} \sum_{i=1}^d \mathbb{Y}_i := \frac{1}{\sqrt{d}} \sum_{i=1}^d (\mathbb{X}_i^{\otimes 2} - \text{diag}(\mathbb{X}_i^{\otimes 2}))$$

and then they directly applied the bound in [6] to obtain (5.1). For us, it is much easier to deal with the case of row independence and more importantly, we get better regimes. Such a difference is not a contradiction to the claim from [4], but instead it stems from our different strategy of applying Stein's method: we first consider the half-matrix or the random vector formed by the upper-triangular entries, which, in the case of row-independence (not in the case of column independence), has invertible covariance matrix, so we are able to use a powerful machinery — the so-called Malliavin-Stein approach, to get the high-dimensional regimes for half-matrix. And the high-dimensional regime for the full-size matrix can be easily passed from that of half-matrix in view of our easy Lemma 2.2. This trick has also been applied in Section 4.2 to obtain the high-dimensional regime of random  $p$ -tensors, a natural extension of Wishart matrices.

To conclude this article, we propose several open questions:

**Q.1:** We introduced the notion of Rosenblatt-Wishart matrix in our Section 3.2 and we listed several basic properties of such a new object.

Is it the right candidate for the *random-matrix approximation* (in the sense of [18]) of the so-called *non-commutative Tchebycheff process* introduced in [15] ?

**Q.2:** In Corollary 4.5, we provide the high-dimensional regimes with respect to the Gaussian approximation in the case of overall correlation. Similar to Corollary 1.4, one may be able to construct many interesting examples of the correlation functions  $r$  and  $s$  that arise from some generalization of fractional Brownian sheets. This is left for interested readers.

**Q.3:** Following Q.2, one may ask the following question:

what does the non-central high-dimensional limit look like in the case of overall correlation? Can one obtain some generalization of our Rosenblatt-Wishart matrix ?

**Q.4:** In the case of full independence, we obtain the high-dimensional regime for the random  $p$ -tensors with respect to the Gaussian approximation. So one may ask the following reasonable question:

Can one formulate a natural correlated setting for random  $p$ -tensors? Can one still obtain nice high-dimensional regimes therein? Concerning the non-central high-dimensional limit, what is the generalization of Rosenblatt-Wishart matrix in the  $p$ -tensor setting?

## REFERENCES

- [1] A. Anandkumar, R. Ge, D. Hsu, S. M. Kakade, and M. Telgarsky (2014). Tensor decompositions for learning latent variable models. *Journal of Machine Learning Research*, 15: 2773–2832.
- [2] J.-C. Breton and I. Nourdin (2008). Error bounds on the non-normal approximation of Hermite power variations of fractional Brownian motion. *Electron. Comm. Probab.* **13**, pp. 482–493.
- [3] S. Bubeck, J. Ding, R. Eldan and M.Z. Rácz (2016). Testing high-dimensional geometry in random graphs. *Random Structures and Algorithms* **49**, no.3, pp. 503-532.
- [4] S. Bubeck and S. Ganguly (2016). Entropic CLT and phase transition in high-dimensional Wishart matrices. *Int. Math. Res. Not.*, pp. 1-19
- [5] S. Chatterjee (2009). Fluctuations of eigenvalues and second order Poincaré inequalities. *Probability Theory and Related Fields* **143**, no. 1–2, pp. 1–40.
- [6] S. Chatterjee and E. Meckes (2008). Multivariate normal approximation using exchangeable pairs. *ALEA* **4**, pp. 257-283.
- [7] R. M. Dudley (2010). *Real Analysis and Probability*. 2nd edition, Cambridge University Press.
- [8] A. Guntuboyina and H. Leeb (2009). Concentration of the spectral measure of large Wishart matrices with dependent entries. *Elect. Comm. in Probab.* **14**, pp. 334–342.
- [9] T. Jiang and D. Li (2015). Approximation of rectangular beta-laguerre ensembles and large deviations. *J. Theor. Probab.* **28**, no. 3, pp. 804 - 847
- [10] I. M. Johnstone (2006). High Dimensional Statistical Inference and Random Matrices. *Proceedings of the International Congress of Mathematicians*, Madrid, Spain.
- [11] V. A. Marchenko and L. A. Pastur (1967). Distribution of eigenvalues for some sets of random matrices, *Math. USSR Sbornik* **72**, no. 4, pp. 507–536
- [12] I. Nourdin and G. Peccati (2009). Stein’s method on Wiener chaos. *Probab. Theory Relat. Fields* **145**, no. 1, pp. 75-118.
- [13] I. Nourdin and G. Peccati (2012). *Normal approximations with Malliavin calculus: from Stein’s method to universality*. Cambridge tracts in Mathematics **192**, Cambridge University Press.
- [14] I. Nourdin, G. Peccati and A. Réveillac (2010). Multivariate normal approximation using Stein’s method and Malliavin calculus. *Ann. Institut. H. Poincaré - Probab. Statist.* **46**, no. 1, pp. 45–58.
- [15] I. Nourdin and M.S. Taqqu (2014). Central and non-central limit theorems in a free probability setting. *J. Theoret. Probab.*, **27**, no. 1, pp 220–248.
- [16] I. Nourdin and G. Zheng (2017). Exchangeable pairs on Wiener chaos. arXiv preprint.
- [17] D. Nualart and G. Peccati (2005). Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33**, no. 1, pp. 177-193.
- [18] J.C. Pardo, J.-L. Pérez and V. Pérez-Abreu (2016). A random matrix approximation for the non-commutative fractional Brownian motion. *J. Theoret. Probab.*, **29**, no. 4, pp 1581–1598.
- [19] M. Z. Rácz and J. Richey (2016). A smooth transition from Wishart to GOE. *J. Theoret. Probab.*, to appear.
- [20] M. S. Veillette and M. S. Taqqu (2013). Properties and numerical evaluation of the Rosenblatt distribution. *Bernoulli* **19**, no. 3, pp. 982–1005
- [21] J. Wishart (1928). The generalized product moment distribution in samples from a normal multivariate population. *Biometrika* **20 A**, pp. 32–52

IVAN NOURDIN: UNIVERSITÉ DU LUXEMBOURG, UNITÉ DE RECHERCHE EN MATHÉMATIQUES, MAISON DU NOMBRE, 6 AVENUE DE LA FONTE, L-4364 ESCH-SUR-ALZETTE, GRAND DUCHY OF LUXEMBOURG

*E-mail address:* [ivan.nourdin@uni.lu](mailto:ivan.nourdin@uni.lu)

GUANGQU ZHENG: UNIVERSITY OF MELBOURNE, SCHOOL OF MATHEMATICS AND STATISTICS, PETER HALL BUILDING, MONASH ROAD, PARKVILLE VIC 3010, AUSTRALIA

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