Gaussian fluctuation for Gaussian Wishart matrices of overall correlation*

Ivan Nourdin and Fei Pu
University of Luxembourg

April 1, 2021

Abstract

In this note, we study the Gaussian fluctuations for the Wishart matrices $d^{-1}\mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$, where $\mathcal{X}_{n,d}$ is a $n \times d$ random matrix whose entries are jointly Gaussian and correlated with row and column covariance functions given by r and s respectively such that $r(0) = s(0) = 1$. Under the assumptions $s \in \ell^{4/3}(\mathbb{Z})$ and $\|r\|_{\ell^1(\mathbb{Z})} < \sqrt{6}/2$, we establish the $\sqrt{n^3/d}$ convergence rate for the Wasserstein distance between a normalization of $d^{-1}\mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$ and the corresponding Gaussian ensemble. This rate is the same as the optimal one computed in [3–5] for the total variation distance, in the particular case where the Gaussian entries of $\mathcal{X}_{n,d}$ are independent. Similarly, we obtain the $\sqrt{n^{2p-1}/d}$ convergence rate for the Wasserstein distance in the setting of random p-tensors of overall correlation. Our analysis is based on the Malliavin-Stein approach.

MSC 2010 subject classification. Primary: 60B20, 60F05; Secondary: 60G22,60H07.
Keywords: Stein's method; Malliavin calculus; High-dimensional regime; Wishart matrices/tensors.
Abbreviated title: Gaussian fluctuation for Wishart matrices

1 Introduction and main result

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

$$\langle e_{ij}\,, e_{i'j'} \rangle_{\mathfrak{H}} = r(i - i')\, s(j - j'), \tag{1.1}$$

where $s, r : \mathbb{Z} \to \mathbb{R}$ stand for some covariance functions satisfying $s(0) = r(0) = 1$. In particular, observe that $\|e_{ij}\|_{\mathfrak{H}} = 1$ for all $i, j \geq 1$.

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

$$\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d} = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1d} \\ X_{21} & X_{22} & \dots & X_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nd} \end{pmatrix}. \tag{1.2}$$

*Research supported in part by FNR grant APOGee (R-AGR-3585-10) at University of Luxembourg.

Our goal is to study the high-dimensional fluctuations of Gaussian Wishart matrices $d^{-1}\mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$ by considering a normalized version given by

$$\widetilde{\mathcal{W}}_{n,d} = \left(\widetilde{W}_{ij}\right)_{1\le i,j\le n}, \tag{1.3}$$

where

$$\widetilde{W}_{ij} = \frac{1}{\sqrt{d}}\sum_{k=1}^{d}\left(X_{ik}X_{jk} - r(i-j)\right). \tag{1.4}$$

Since the e_{ij}'s are not supposed to be orthogonal, see (1.1), it is important to note that the Gaussian entries X_{ij} of $\mathcal{X}_{n,d}$ are fully correlated in general.

Let $\mathcal{G}_{n,d}^{r,s} = (G_{ij})_{1\le i,j\le n}$ be the $n\times n$ random symmetric matrix such that the associated random vector $(G_{11},\dots,G_{1n},G_{21},\dots,G_{2n},\dots,G_{n1},\dots,G_{nn})$ is Gaussian with mean 0 and has the same covariance matrix as

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

Recall the definition of *Wasserstein distance* between two random variables with values in $\mathcal{M}_n(\mathbb{R})$ (the space of $n\times n$ real matrices): for $\mathcal{X},\mathcal{Y}:\Omega\to\mathcal{M}_n(\mathbb{R})$ such that $\mathrm{E}\|\mathcal{X}\|_{\mathrm{HS}}+\mathrm{E}\|\mathcal{Y}\|_{\mathrm{HS}}<\infty$,

$$d_{\mathrm{Wass}}(\mathcal{X},\mathcal{Y}) := \sup\left\{\mathrm{E}[g(\mathcal{X})]-\mathrm{E}[g(\mathcal{Y})] : \|g\|_{\mathrm{Lip}}\le 1\right\}, \tag{1.5}$$

with

$$\|g\|_{\mathrm{Lip}} := \sup_{\substack{A,B\in\mathcal{M}_n(\mathbb{R})\\ A\ne B}} \frac{|g(A)-g(B)|}{\|A-B\|_{\mathrm{HS}}} \quad \text{for } g:\mathcal{M}_n(\mathbb{R})\to\mathbb{R},$$

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

The main result of this paper is the following.

Theorem 1.1. *Assume*

$$\|r\|_{\ell^1(\mathbb{Z})} < \sqrt{6}/2. \tag{1.6}$$

Then for all $n,d\ge 1$,

$$d_{\mathrm{Wass}}\left(\widetilde{\mathcal{W}}_{n,d},\mathcal{G}_{n,d}^{r,s}\right) \le \frac{\|r\|_{\ell^1(\mathbb{Z})}^{3/2}}{3-2\|r\|_{\ell^1(\mathbb{Z})}^2}\sqrt{\frac{32n^3}{d}\left(\sum_{|k|\le d}|s(k)|^{4/3}\right)^3}. \tag{1.7}$$

Remark 1.2. (1) We will actually show that

$$d_{\mathrm{Wass}}\left(\widetilde{\mathcal{W}}_{n,d},\mathcal{G}_{n,d}^{r,s}\right) \le \frac{\|r\|_{\ell^1(\mathbb{Z})}^{3/2}}{3-2\|r\|_{\ell^1(\mathbb{Z})}^2}\sqrt{\frac{32d}{\sum_{k,\ell=1}^d s(k-\ell)^2}\times\frac{n^3}{d}\left(\sum_{|k|\le d}|s(k)|^{4/3}\right)^3}.$$

This implies (1.7) since $\sum_{k,\ell=1}^d s(k-\ell)^2 \ge d\,s(0)^2 = d$.

(2) Under the condition (1.6) and if we assume $s \in \ell^{4/3}(\mathbb{Z})$, then (1.7) leads to

$$d_{\rm Wass}\left(\widetilde{\mathcal{W}}_{n,d}\,,\mathcal{G}^{r,s}_{n,d}\right) = O(\sqrt{n^3/d}).$$

Hence in this case of overall correlation, $\mathcal{W}_{n,d}$ continues to be close to the Gaussian random matrix $\mathcal{G}^{r,s}_{n,d}$ as long as $n^3/d \to 0$, exactly like in the row independence case in [10, Theorem 1.2]; see also the full independence case considered in [3–5].

(3) An explicit example of covariance function satisfying (1.6) is $r(k) = \mathrm{e}^{-\lambda|k|^\alpha}$ for $k \in \mathbb{Z}$, with $1 \le \alpha \le 2$ and where $\lambda > 0$ is chosen large enough.

The $\sqrt{n^3/d}$ convergence rate obtained in Theorem 1.1 relies on Malliavin calculus and Stein's method, precisely, Proposition 2.1 below, which has already been employed in [10] to investigate the Gaussian approximation for Wishart matrix in the row independence case (that is, $r(k) = 1_{k=0}$). In the case of overall correlation, it is not clear if the covariance matrix C in Proposition 2.1 is invertible or not. To bypass this problem, the authors of [10] made use of the bounds from [7, Theorem 6.1.2] and [9, Theorem 9.3] with, as a price to pay, the necessity to consider a smoother distance instead of Wasserstein distance; see [10, Proposition 4.1 and Theorem 4.3].

Fortunately, in the case of overall correlation, we discover that condition (1.6) guarantees that the covariance matrix C in Proposition 2.1 is strictly diagonally dominant and hence invertible. Moreover, we can bound the operator norms $\|C\|_{\rm op}$ and $\|C^{-1}\|_{\rm op}$ in terms of $\|r\|_{\ell^1(\mathbb{Z})}$. Therefore, we are able to apply the inequality (2.3) in Proposition 2.1 to derive the estimate in (1.7); see the proof of Theorem 1.1 in Section 3.

The Malliavin-Stein approach can also be applied to study Gaussian approximation of Wishart p-tensors in the case of overall correlation. In Theorem 4.1, we propose a condition on $\|r\|_{\ell^1(\mathbb{Z})}$ (see (4.3) below) under which the covariance matrix of the p-tensors is invertible. Hence we appeal to the Proposition 2.1 again and establish the $\sqrt{n^{2p-1}/d}$ convergence rate for the Wasserstein distance between the p-tensors; the same as full independence case considered in [10, Theorem 4.3]. We refer to [1, 2, 6] for some other recent applications of Malliavin calculus and Stein method in the study of high-dimensional regime of Wishart matrices/tensors.

2 Preliminaries

In this section, we collect some elements of Malliavin calculus and Stein's method and refer to [7] (see also [11, 12]) for more details. Recall the isonormal Gaussian process $X = \{X(h), h \in \mathfrak{H}\}$ over a real separable Hilbert space $\mathfrak{H}$ defined on some probability space $(\Omega, \mathcal{F}, \mathrm{P})$.

For every $p \ge 1$, we let $\mathcal{H}_p$ denote the pth 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 pth Hermite polynomial. 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 pth tensor product $\mathfrak{H}^{\odot p}$ and the pth Wiener chaos $\mathcal{H}_p$.

Consider $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$ with $p, q \ge 1$. For $j \in \{0, \dots, p \wedge q\}$, $f \otimes_j g$ denotes the j-contraction of f and g (see [7, Section B.4] for the precise definition) and $f \widetilde{\otimes}_j g$ stands for the symmetrization of $f \otimes_j g$. For $f \in \mathfrak{H}^{\odot p}$, the Malliavin derivative of $I_p(f)$ is the random element of $\mathfrak{H}$ given by $DI_p(f) = pI_{p-1}(f)$ (see [7, Proposition 2.7.4]) and we have for $f, g \in \mathfrak{H}^{\odot p}$,

$$\mathrm{E}\left[p^{-1}\langle DI_p(f)\,, DI_p(g)\rangle_{\mathfrak{H}}\right] = \mathrm{E}[I_p(f)I_p(g)] = p!\langle f\,, g\rangle_{\mathfrak{H}^{\otimes p}}. \qquad (2.1)$$

Moreover, according to the formula [7, (6.2.3)], for $f, g \in \mathfrak{H}^{\odot p}$,

$$\operatorname{Var}\left(p^{-1}\langle DI_p(f), DI_p(g)\rangle_{\mathfrak{H}}\right) = p^2\sum_{j=1}^{p-1}(j-1)!^2\binom{p-1}{j-1}^4(2p-2j)\|f\widetilde{\otimes}_j g\|^2_{\mathfrak{H}^{\otimes(2p-2j)}}. \tag{2.2}$$

The following result, the so-called Malliavin-Stein approach, provides a powerful machinery to investigate the normal approximation for the Gaussian Wishart matrix of overall correlation.

Proposition 2.1 (see [8, Corollary 3.6])**.** *Fix integers $m \ge 2$ and $1 \le p_1 \le \ldots \le p_m$. Consider a random vector $F = (F_1, \ldots, F_m) = (I_{p_1}(f_1), \ldots, I_{p_m}(f_m))$ 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 $\mathrm{Z} \in \mathrm{N}_m(0, C)$. Then*

$$d_{\mathrm{Wass}}(F, \mathrm{Z}) \le \|C^{-1}\|_{\mathrm{op}}\|C\|^{1/2}_{\mathrm{op}}\left(\sum_{1\le i,j\le m}\mathrm{E}\left[\left(C_{ij} - p_j^{-1}\langle DF_i, DF_j\rangle_{\mathfrak{H}}\right)^2\right]\right)^{1/2}, \tag{2.3}$$

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

Note that the *Wasserstein distance* $d_{\mathrm{Wass}}(F, \mathrm{Z})$ between two general m-dimensional random vectors F and Z is defined as

$$d_{\mathrm{Wass}}(F, \mathrm{Z}) := \sup\left\{\mathrm{E}\left[g(F)\right] - \mathrm{E}\left[g(\mathrm{Z})\right] : \|g\|_{\mathrm{Lip}} \le 1\right\}, \tag{2.4}$$

where $\|g\|_{\mathrm{Lip}}$ denotes the usual Lipschitz constant of a function $g : \mathbb{R}^m \to \mathbb{R}$ with respect to the Euclidean norm.

Lemma 2.2 of [10] has provided a trick to pass the high-dimensional regime for the full-size symmetric matrix to that of half-matrix. Recall that the half matrix $\mathcal{Z}^{\mathrm{half}}$ of a $n \times n$ random symmetric matrix $\mathcal{Z} = (Z_{ij})_{1\le i,j\le n}$ is the $n(n+1)/2$- dimensional random vector formed by the upper-triangular entries, namely:

$$\mathcal{Z}^{\mathrm{half}} = (Z_{11}, Z_{12}, \ldots, Z_{1n}, Z_{22}, \ldots, Z_{23}, \ldots, Z_{2n}, \ldots, Z_{nn}). \tag{2.5}$$

According to [10, Lemma 2.2], for two symmetric random matrices $\mathcal{X}, \mathcal{Y}\colon \Omega \to \mathcal{M}_n(\mathbb{R})$,

$$d_{\mathrm{Wass}}(\mathcal{X}, \mathcal{Y}) \le \sqrt{2}\, d_{\mathrm{Wass}}(\mathcal{X}^{\mathrm{half}}, \mathcal{Y}^{\mathrm{half}}), \tag{2.6}$$

where the left-hand side Wasserstein distance is defined in (1.5) while the right-hand defined in (2.4).

Finally, in order to apply Proposition 2.1 to obtain the rate for the Wasserstein distance stated in Theorem 1.1, we use the product formula for the multiple Wiener-Itô integrals (see [7, Theorem 2.7.10]) to realize the (i, j)th entry of $\widetilde{\mathcal{W}}_{n,d}$ defined in (1.4) as an element in the second Wiener chaos $\mathcal{H}_2$, namely:

$$\begin{aligned}\widetilde{W}_{ij} &= \frac{1}{\sqrt{d}}\sum_{k=1}^d (X_{ik}X_{jk} - r(i-j)) = \frac{1}{\sqrt{d}}\sum_{k=1}^d (I_1(e_{ik})I_1(e_{jk}) - \langle e_{ik}, e_{jk}\rangle_{\mathfrak{H}})\\ &= I_2(f_{ij}^{(d)}),\end{aligned} \tag{2.7}$$

where

$$f_{ij}^{(d)} = \frac{1}{\sqrt{d}}\sum_{k=1}^d e_{ik}\widetilde{\otimes} e_{jk} = \frac{1}{2\sqrt{d}}\sum_{k=1}^d (e_{ik}\otimes e_{jk} + e_{jk}\otimes e_{ik}). \tag{2.8}$$

3 Proof of Theorem 1.1

We prove Theorem 1.1 in this section and assume that (1.6) holds throughout this section. We begin to establish the following supporting lemmas.

Lemma 3.1. *For all $n \geq 1$ and for all fixed (i,j) with $1 \leq i \leq j \leq n$,*

$$\sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}} |r(i-u)r(v-j)+r(i-v)r(u-j)| \leq 2\|r\|^2_{\ell^1(\mathbb{Z})} - 2 < 1. \tag{3.1}$$

Proof. The second inequality in (3.1) is clearly true by (1.6). In order to prove the first one, we write

$$\begin{aligned}\sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}} &|r(i-u)r(v-j)+r(i-v)r(u-j)| \\ &\leq \sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}} |r(i-u)r(v-j)| + \sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}} |r(i-v)r(u-j)|.\end{aligned} \tag{3.2}$$

For the first sum on the right-hand side of (3.2), we have

$$\begin{aligned}\sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}} |r(i-u)r(v-j)| &\leq r(0)\sum_{v\neq j}|r(v-j)| + \sum_{u\neq i}\sum_{v\in\mathbb{Z}}|r(i-u)|\,|r(v-j)| \\ &= \left(r(0)+\|r\|_{\ell^1(\mathbb{Z})}\right)\|r\|_{\ell^1(\mathbb{Z}\setminus\{0\})} \\ &= \|r\|^2_{\ell^1(\mathbb{Z})} - 1.\end{aligned} \tag{3.3}$$

For the second sum on the right-hand side of (3.2),

$$\sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}} |r(i-v)r(u-j)| \leq |r(i-j)|\sum_{v\neq j}|r(i-v)| + \sum_{\substack{1\le u\le v\le n\\u\neq i}} |r(i-v)r(u-j)| \tag{3.4}$$

We observe that

$$|r(i-j)|\sum_{v\neq j}|r(i-v)| \leq \begin{cases} r(0)\|r\|_{\ell^1(\mathbb{Z}\setminus\{0\})} = \|r\|_{\ell^1(\mathbb{Z})} - 1, & \text{if } i=j, \\ \frac{1}{2}\|r\|_{\ell^1(\mathbb{Z}\setminus\{0\})}\|r\|_{\ell^1(\mathbb{Z})} = \|r\|_{\ell^1(\mathbb{Z})}(\|r\|_{\ell^1(\mathbb{Z})}-1)/2, & \text{if } i<j. \end{cases} \tag{3.5}$$

Moreover, since $i \leq j$,

$$\begin{aligned}\sum_{\substack{1\le u\le v\le n\\u\neq i}} |r(i-v)r(u-j)| &= \sum_{\substack{1\le u\le v\le n\\u<i}} |r(i-v)r(u-j)| + \sum_{i<u\le v\le n} |r(i-v)r(u-j)| \\ &\leq \|r\|_{\ell^1(\mathbb{Z})}\sum_{1\le u<i}|r(u-j)| + \sum_{i<v\le n}|r(i-v)|\|r\|_{\ell^1(\mathbb{Z})} \\ &\leq \|r\|_{\ell^1(\mathbb{Z})}\|r\|_{\ell^1(\mathbb{Z}\setminus\{0\})} = \|r\|_{\ell^1(\mathbb{Z})}(\|r\|_{\ell^1(\mathbb{Z})}-1).\end{aligned} \tag{3.6}$$

By (3.4)–(3.6), it yields that

$$\sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}} |r(i-v)r(u-j)| \leq (\|r\|_{\ell^1(\mathbb{Z})}-1)\left(\left(1\vee(\|r\|_{\ell^1(\mathbb{Z})}/2)\right)+\|r\|_{\ell^1(\mathbb{Z})}\right). \tag{3.7}$$

Therefore, we combine (3.3) and (3.7) to obtain

$$\sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}} |r(i-u)r(v-j)+r(i-v)r(u-j)| \le (\|r\|_{\ell^1(\mathbb{Z})}-1)\left(\left(1\vee(\|r\|_{\ell^1(\mathbb{Z})}/2)\right)+2\|r\|_{\ell^1(\mathbb{Z})}+1\right), \tag{3.8}$$

Elementary calculation on quadratic inequality shows that

$$(\|r\|_{\ell^1(\mathbb{Z})}-1)\left(\left(1\vee(\|r\|_{\ell^1(\mathbb{Z})}/2)\right)+2\|r\|_{\ell^1(\mathbb{Z})}+1\right)<1 \Leftrightarrow \|r\|_{\ell^1(\mathbb{Z})}<\sqrt{6}/2,$$

whence

$$(\|r\|_{\ell^1(\mathbb{Z})}-1)\left(\left(1\vee(\|r\|_{\ell^1(\mathbb{Z})}/2)\right)+2\|r\|_{\ell^1(\mathbb{Z})}+1\right)=2\|r\|^2_{\ell^1(\mathbb{Z})}-2 \tag{3.9}$$

provided $\|r\|_{\ell^1(\mathbb{Z})}<\sqrt{6}/2$.

Therefore, under the assumption (1.6), the estimate (3.1) follows from (3.8) and (3.9). □

Recall the $n\times n$ symmetric Gaussian random matrix $\mathcal{G}^{r,s}_{n,d}=(G_{ij})_{1\le i,j\le n}$ from (1.4). Let C denote the covariance matrix of $(\mathcal{G}^{r,s}_{n,d})^{\text{half}}$. Notice that the matrix norms $\|C\|_1$ and $\|C\|_\infty$ of the symmetric matrix C are equal and given by

$$\|C\|_1=\|C\|_\infty=\sup_{1\le i\le j\le n}\sum_{1\le u\le v\le n}|\mathrm{E}[G_{ij}G_{uv}]|\,. \tag{3.10}$$

Lemma 3.2. *The matrix C is invertible and the following estimates on operator norms hold:*

$$\|C^{-1}\|_{\text{op}}\le\frac{d}{\sum_{k,\ell=1}^d s(k-\ell)^2}\left(3-2\|r\|^2_{\ell^1(\mathbb{Z})}\right)^{-1}\quad \textit{and}\quad \|C\|_{\text{op}}\le\frac{2\|r\|^2_{\ell^1(\mathbb{Z})}}{d}\sum_{k,\ell=1}^d s(k-\ell)^2. \tag{3.11}$$

Proof. According to [10, (4.3)], the entries of C are given by

$$\mathrm{E}[G_{ij}G_{uv}]=\frac{(r(i-u)r(v-j)+r(i-v)r(u-j))}{d}\sum_{k,\ell=1}^d s(k-\ell)^2 \tag{3.12}$$

for $1\le i\le j\le n$ and $1\le u\le v\le n$. Letting $(u,v)=(i,j)$ in (3.12), we obtain the following lower and upper bounds on the diagonal entries of C:

$$\frac{1}{d}\sum_{k,\ell=1}^d s(k-\ell)^2=\inf_{1\le i\le j\le n}\mathrm{E}[G^2_{ij}]\le\sup_{1\le i\le j\le n}\mathrm{E}[G^2_{ij}]=\frac{2}{d}\sum_{k,\ell=1}^d s(k-\ell)^2. \tag{3.13}$$

Moreover, under the condition (1.6), we have

$$\|C\|_1=\|C\|_\infty=\sup_{1\le i\le j\le n}\sum_{1\le u\le v\le n}|\mathrm{E}[G_{ij}G_{uv}]|\le\frac{2\|r\|^2_{\ell^1(\mathbb{Z})}}{d}\sum_{k,\ell=1}^d s(k-\ell)^2 \tag{3.14}$$

thanks to (3.13) and Lemma 3.1. Hence the second inequality in (3.11) follows from (3.14) and the Hölder's inequality for matrix norms: $\|C\|_{\text{op}}\le\sqrt{\|C\|_1\|C\|_\infty}$ (see [13, Theorem 4.3.1]).

Furthermore, by (3.12) and (3.13) in the first equality,

$$\inf_{1\le i\le j\le n}\left(\mathrm{E}[G_{ij}^2]-\sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}}|\mathrm{E}[G_{ij}G_{uv}]|\right)\ge\inf_{1\le i\le j\le n}\mathrm{E}[G_{ij}^2]-\sup_{1\le i\le j\le n}\sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}}|\mathrm{E}[G_{ij}G_{uv}]|$$

$$=\frac{1}{d}\sum_{k,\ell=1}^{d}s(k-\ell)^2\left(1-\sup_{1\le i\le j\le n}\sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}}|r(i-u)r(v-j)+r(i-v)r(u-j)|\right)$$

$$\ge\frac{1}{d}\sum_{k,\ell=1}^{d}s(k-\ell)^2\left(3-2\|r\|_{\ell^1(\mathbb{Z})}^2\right)>0,\quad\text{by Lemma 3.1 and under (1.6)},\tag{3.15}$$

which implies that the symmetric matrix C is strictly diagonally dominant and hence invertible.

Now we apply [14, Corollary 2] and (3.15) to see that

$$\|C^{-1}\|_{\mathrm{op}}^{-1}\ge\inf_{1\le i\le j\le n}\left(\mathrm{E}[G_{ij}^2]-\sum_{\substack{1\le u\le v\le n\\(u,v)\neq(i,j)}}|\mathrm{E}[G_{ij}G_{uv}]|\right)$$

$$\ge\frac{1}{d}\sum_{k,\ell=1}^{d}s(k-\ell)^2\left(3-2\|r\|_{\ell^1(\mathbb{Z})}^2\right),\tag{3.16}$$

which proves the first inequality in (3.11). □

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. Recall that $(\widetilde{\mathcal{W}}_{n,d})^{\mathrm{half}}$ is the half matrix of the Wishart matrix $\widetilde{\mathcal{W}}_{n,d}$ defined in (1.3) and (1.4), whose entries can be represented as the elements in the second Wiener chaos $\mathcal{H}_2$; see (2.7) and (2.8). By Lemma 3.2, the covariance matrix of $(\widetilde{\mathcal{W}}_{n,d})^{\mathrm{half}}$ is invertible and hence we can apply Proposition 2.1 with $m=n(n+1)/2$, $p_1=\ldots=p_m=2$ and $F=(\widetilde{\mathcal{W}}_{n,d})^{\mathrm{half}}$. Indeed, we have

$$d_{\mathrm{Wass}}\left((\widetilde{\mathcal{W}}_{n,d})^{\mathrm{half}},(\mathcal{G}_{n,d}^{r,s})^{\mathrm{half}}\right)=d_{\mathrm{Wass}}\left(F,(\mathcal{G}_{n,d}^{r,s})^{\mathrm{half}}\right)$$

$$\le\|C^{-1}\|_{\mathrm{op}}\|C\|_{\mathrm{op}}^{1/2}\left(\sum_{1\le i,j\le m}\mathrm{E}\left[\left(C_{ij}-\frac{1}{2}\langle DF_i,DF_j\rangle_{\mathfrak{H}}\right)^2\right]\right)^{1/2}.$$

Using the identities (2.1) and (2.2), the proceeding yields that

$$d_{\text{Wass}}\left((\widetilde{\mathcal{W}}_{n,d})^{\text{half}},(\mathcal{G}^{r,s}_{n,d})^{\text{half}}\right)\le \|C^{-1}\|_{\text{op}}\|C\|_{\text{op}}^{1/2}\left(\sum_{\substack{1\le i\le j\le n\\1\le p\le q\le n}}\operatorname{Var}\left(\frac{1}{2}\langle D\widetilde{W}_{ij}\,,D\widetilde{W}_{pq}\rangle_{\mathfrak{H}}\right)\right)^{1/2}$$

$$=\|C^{-1}\|_{\text{op}}\|C\|_{\text{op}}^{1/2}\left(8\sum_{\substack{1\le i\le j\le n\\1\le p\le q\le n}}\|f_{ij}^{(d)}\widetilde{\otimes}_1 f_{pq}^{(d)}\|^2_{\mathfrak{H}^{\otimes 2}}\right)^{1/2}$$

$$\le\frac{\|r\|_{\ell^1(\mathbb{Z})}}{3-2\|r\|^2_{\ell^1(\mathbb{Z})}}\sqrt{\frac{16d}{\sum_{k,\ell=1}^d s(k-\ell)^2}\sum_{\substack{1\le i\le j\le n\\1\le p\le q\le n}}\|f_{ij}^{(d)}\otimes_1 f_{pq}^{(d)}\|^2_{\mathfrak{H}^{\otimes 2}}},\quad(3.17)$$

where the second inequality follows from (3.11) and the fact that $\|\widetilde{h}\|_{\mathfrak{H}^{\otimes r}}\le\|h\|_{\mathfrak{H}^{\otimes r}}$.

It remains to estimate the last term in (3.17). Appealing to [10, (4.7)],

$$\|f_{ij}^{(d)}\otimes_1 f_{pq}^{(d)}\|^2_{\mathfrak{H}^{\otimes 2}}\le\frac{\mathfrak{X}_{i,j,p,q}}{16d}\left(\sum_{|k|\le d}|s(k)|^{4/3}\right)^3,\quad(3.18)$$

where $\mathfrak{X}_{i,j,p,q}$ is a sum of sixteen terms given by the expression below (4.7) in [10]. Moreover, we have

$$\mathfrak{X}_{i,j,p,q}\le 7|r(j-q)|+5|r(p-j)|+3|r(i-q)|+|r(i-p)|,\quad(3.19)$$

which together with [10, (4.12)] implies that

$$\sum_{\substack{1\le i\le j\le n\\1\le p\le q\le n}}\mathfrak{X}_{i,j,p,q}\le 16n^3\|r\|_{\ell^1(\mathbb{Z})}.\quad(3.20)$$

Taking into account (3.17), (3.18) and (3.20), we conclude that

$$d_{\text{Wass}}\left((\widetilde{\mathcal{W}}_{n,d})^{\text{half}},(\mathcal{G}^{r,s}_{n,d})^{\text{half}}\right)\le\frac{\|r\|^{3/2}_{\ell^1(\mathbb{Z})}}{3-2\|r\|^2_{\ell^1(\mathbb{Z})}}\sqrt{\frac{16d}{\sum_{k,\ell=1}^d s(k-\ell)^2}\times\frac{n^3}{d}\left(\sum_{|k|\le d}|s(k)|^{4/3}\right)^3}.\quad(3.21)$$

Finally, we combine (3.21) with (2.6) to obtain the estimate in Remark 1.2(1), which leads to (1.7) □

4 Random p-tensors

The result of Theorem 1.1 for Gaussian Wishart matrix can be extended to random p-tensors ($p\ge 2$). We first introduce some notations of p-tensors. Let $\mathbb{X}_i=(X_{1i},\dots,X_{ni})^T$ be the ith column of the random matrix $\mathcal{X}_{n,d}$ defined in (1.2). We write

$$\mathbb{X}_i=\sum_{j=1}^n X_{ji}\,\varepsilon_j,$$

where $\{\varepsilon_j, j = 1, \ldots, n\}$ is the canonical basis of $\mathbb{R}^n$. Then the p-tensor product of $\mathbb{X}_i$ is given by

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

so that

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

A repeated application of the product formula for the multiple Wiener-Itô integrals (see [7, Theorem 2.7.10]) ensures that

$$\prod_{k=1}^{p} X_{j_k i} = \prod_{k=1}^{p} I_1(e_{j_k i}) = I_p \left(\mathrm{sym} \left(e_{j_1 i} \otimes \cdots \otimes e_{j_p i} \right) \right) + \text{lower order terms},$$

where sym denotes the canonical symmetrization.

Analogous to the choice of $\widetilde{W}_{ij}$ defined in (2.7) and (2.8), in the case of overall correlation, we consider the following normalized version of p-tensor of $\mathcal{X}_{n,d}$:

$$\left(\widetilde{\mathbf{Y}}_{\mathbf{j}} = I_p(f_{\mathbf{j}}^{(d)}),\, \mathbf{j} = (j_1, \ldots, j_p) \in \{1, \ldots, n\}^p \right),$$

where

$$f_{\mathbf{j}}^{(d)} = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} \mathrm{sym} \left(e_{j_1 k} \otimes \cdots \otimes e_{j_p k} \right). \tag{4.1}$$

Moreover, we remove the diagonal terms and focus on the Gaussian approximation of

$$\widetilde{\mathcal{Y}}_{n,d} = \left(\widetilde{\mathbf{Y}}_{\mathbf{j}} = I_p(f_{\mathbf{j}}^{(d)}),\, \mathbf{j} \in \Delta_p \right), \tag{4.2}$$

where $f_{\mathbf{j}}^{(d)}$ is defined in (4.1) and $\Delta_p = \{(j_1, \ldots, j_p) \in \{1, \ldots, n\}^p : j_1, \ldots, j_p \text{ are mutually distinct}\}$.

The following result extends the Gaussian approximation of random p-tensors of full independence in [10, Theorem 4.6] to the case of overall correlation.

Theorem 4.1. *Let $\widetilde{\mathcal{Y}}_{n,d}$ be defined in (4.2) and $\mathbf{Z} = (\mathbf{Z}_{\mathbf{j}} : \mathbf{j} \in \Delta_p)$ a centered Gaussian vector in $\mathbb{R}^{p!\binom{n}{p}}$ which has the same covariance matrix as $\widetilde{\mathcal{Y}}_{n,d}$. Assume that the covariance function r satisfies*

$$\left(1 - \left(\|r\|_{\ell^1(\mathbb{Z})} - 1 \right) \left(p! \|r\|_{\ell^1(\mathbb{Z})}^{p-1} + (p! - 1)/2 \right) \right) > 0. \tag{4.3}$$

Then there exists a positive constant C_p depending on p and $\|r\|_{\ell^1(\mathbb{Z})}$ (see (4.21) below) such that

$$d_{Wass} \left(\widetilde{\mathcal{Y}}_{n,d}\,, \mathbf{Z} \right) \leq C_p \sqrt{\frac{d}{\left| \sum_{k,\ell=1}^{d} s(k-\ell)^p \right|} \left(\sum_{|k| \leq d} |s(k)|^{4/3} \right)^3 \frac{n^{2p-1}}{d}}. \tag{4.4}$$

Remark 4.2. (1) Note that the above Wasserstein distance is for $\mathbb{R}^{p!\binom{n}{p}}$-valued random vectors; as defined in (2.4).

(2) The estimate in (4.4) is trivial if $\sum_{k,\ell=1}^d s(k-\ell)^p = 0$. Hence we assume $\sum_{k,\ell=1}^d s(k-\ell)^p \neq 0$ in the following.

(3) Under the condition (4.3), if we assume $s \in \ell^{4/3}(\mathbb{Z})$ and in addition that p is even or $s(k) \geq 0$ for all $k \in \mathbb{Z}$, (4.4) leads to $d_{\text{Wass}}(\widetilde{\mathcal{Y}}_{n,d}, \mathbf{Z}) = O(\sqrt{n^{2p-1}/d})$; the same as the full independence case considered in [10, Theorem 4.6].

Similar to the proof of Theorem 1.1, we reduce the estimate of Wasserstein distance between $\widetilde{\mathcal{Y}}_{n,d}$ and $\mathbf{Z}$ to that of their "half matrices", given by the following $\mathbb{R}^{\binom{n}{p}}$-valued random vectors

$$\widetilde{\mathcal{Y}}_{n,d}^{\uparrow} = \left(\widetilde{\mathbf{Y}}_{\mathbf{j}} = I_p(f_{\mathbf{j}}^{(d)}),\, \mathbf{j} \in \Delta_p^{\uparrow}\right) \quad \text{and} \quad \mathbf{Z}^{\uparrow} = \left(\mathbf{Z}_{\mathbf{j}} : \mathbf{j} \in \Delta_p^{\uparrow}\right), \tag{4.5}$$

where $\Delta_p^{\uparrow} = \{\mathbf{j} \in \{1,\ldots,n\}^p : j_1 < j_2 < \ldots < j_p\}$. Denote $S(p)$ the collection of all permutations of $\{1,\ldots,p\}$.

Lemma 4.3. *Let $\widetilde{C}$ be the covariance matrix of $\widetilde{\mathcal{Y}}_{n,d}^{\uparrow}$. Under the condition (4.3), $\widetilde{C}$ is invertible and we have*

$$\|\widetilde{C}^{-1}\|_{\text{op}} \leq d \left|\sum_{k,\ell=1}^d s(k-\ell)^p\right|^{-1} \left(1 - \left(\|r\|_{\ell^1(\mathbb{Z})} - 1\right)\left(p!\|r\|_{\ell^1(\mathbb{Z})}^{p-1} + (p!-1)/2\right)\right)^{-1} \tag{4.6}$$

and

$$\|\widetilde{C}\|_{\text{op}} \leq \frac{1}{d}\left|\sum_{k,\ell=1}^d s(k-\ell)^p\right| \left(1 + \left(\|r\|_{\ell^1(\mathbb{Z})} - 1\right)\left(p!\|r\|_{\ell^1(\mathbb{Z})}^{p-1} + (p!-1)/2\right)\right). \tag{4.7}$$

Proof. The proof is similar to that of Lemma 3.2. We will see that the condition (4.3) guarantees that the symmetric matrix $\widetilde{C}$ is strictly diagonally dominant and hence invertible. We first compute the entries of $\widetilde{C}$. For $\mathbf{j} = (j_1,\ldots,j_p)$, $\mathbf{j}' = (j_1',\ldots,j_p') \in \Delta_p^{\uparrow}$, using (4.2), (4.1) and isometry,

$$\begin{aligned}
\mathrm{E}[\widetilde{\mathbf{Y}}_{\mathbf{j}}\widetilde{\mathbf{Y}}_{\mathbf{j}'}] &= \mathrm{E}[I_p(f_{\mathbf{j}}^{(d)})I_p(f_{\mathbf{j}'}^{(d)})] \\
&= \frac{p!}{d}\sum_{k,\ell=1}^d \left\langle \operatorname{sym}\left(e_{j_1k}\otimes\cdots\otimes e_{j_pk}\right), \operatorname{sym}\left(e_{j_1'\ell}\otimes\cdots\otimes e_{j_p'\ell}\right)\right\rangle_{\mathfrak{H}^{\otimes p}} \\
&= \frac{1}{p!d}\sum_{k,\ell=1}^d \sum_{\sigma,\tau\in S(p)} \left\langle e_{j_{\sigma(1)}k}\otimes\cdots\otimes e_{j_{\sigma(p)}k}, e_{j_{\tau(1)}'\ell}\otimes\cdots\otimes e_{j_{\tau(p)}'\ell}\right\rangle_{\mathfrak{H}^{\otimes p}} \\
&= \frac{1}{p!d}\sum_{k,\ell=1}^d s(k-\ell)^p \sum_{\sigma,\tau\in S(p)}\prod_{m=1}^p r(j_{\sigma(m)} - j_{\tau(m)}').
\end{aligned} \tag{4.8}$$

As a consequence of (4.8), we have the following lower and upper bounds on the diagonal entries of $\widetilde{C}$: for all $\mathbf{j} \in \Delta_p^{\uparrow}$

$$\begin{aligned}
\mathrm{E}[\widetilde{\mathbf{Y}}_{\mathbf{j}}^2] &\geq \frac{1}{p!d} \left| \sum_{k,\ell=1}^{d} s(k-\ell)^p \right| \left(p! r(0)^p - \sum_{\substack{\sigma,\tau \in S(p) \\ \sigma \neq \tau}} \prod_{m=1}^{p} \left| r(j_{\sigma(m)} - j_{\tau(m)}) \right| \right) \\
&\geq \frac{1}{p!d} \left| \sum_{k,\ell=1}^{d} s(k-\ell)^p \right| \left(p! - ((p!)^2 - p!) \frac{1}{2} \|r\|_{\ell^1(\mathbb{Z}\setminus\{0\})} \right) \\
&= \frac{1}{d} \left| \sum_{k,\ell=1}^{d} s(k-\ell)^p \right| \left(1 - (p!-1) \frac{1}{2} \left(\|r\|_{\ell^1(\mathbb{Z})} - 1 \right) \right), \qquad (4.9)
\end{aligned}$$

and similarly

$$\mathrm{E}[\widetilde{\mathbf{Y}}_{\mathbf{j}}^2] \leq \frac{1}{d} \left| \sum_{k,\ell=1}^{d} s(k-\ell)^p \right| \left(1 + (p!-1) \frac{1}{2} \left(\|r\|_{\ell^1(\mathbb{Z})} - 1 \right) \right). \qquad (4.10)$$

Moreover, the identity (4.8) implies the following estimate on the off-diagonal entries of $\widetilde{C}$: for all $\mathbf{j} \in \Delta_p^{\uparrow}$,

$$\begin{aligned}
\sum_{\substack{\mathbf{j}' \in \Delta_p^{\uparrow} \\ \mathbf{j}' \neq \mathbf{j}}} \left| \mathrm{E}[\widetilde{\mathbf{Y}}_{\mathbf{j}} \widetilde{\mathbf{Y}}_{\mathbf{j}'}] \right| &\leq \frac{1}{p!d} \left| \sum_{k,\ell=1}^{d} s(k-\ell)^p \right| \sum_{\sigma,\tau \in S(p)} \sum_{\substack{\mathbf{j}' \in \Delta_p^{\uparrow} \\ \mathbf{j}' \neq \mathbf{j}}} \prod_{m=1}^{p} \left| r(j_{\sigma(m)} - j'_{\tau(m)}) \right| \\
&\leq \frac{1}{p!d} \left| \sum_{k,\ell=1}^{d} s(k-\ell)^p \right| (p!)^2 \|r\|_{\ell^1(\mathbb{Z}\setminus\{0\})} \|r\|_{\ell^1(\mathbb{Z})}^{p-1} \\
&= \frac{p!}{d} \left| \sum_{k,\ell=1}^{d} s(k-\ell)^p \right| \left(\|r\|_{\ell^1(\mathbb{Z})} - 1 \right) \|r\|_{\ell^1(\mathbb{Z})}^{p-1}. \qquad (4.11)
\end{aligned}$$

Therefore, we obtain from (4.9) and (4.11) that

$$\begin{aligned}
&\inf_{\mathbf{j} \in \Delta_p^{\uparrow}} \left(\mathrm{E}[\widetilde{\mathbf{Y}}_{\mathbf{j}}^2] - \sum_{\mathbf{j}' \in \Delta_p^{\uparrow}, \mathbf{j}' \neq \mathbf{j}} \left| \mathrm{E}[\widetilde{\mathbf{Y}}_{\mathbf{j}} \widetilde{\mathbf{Y}}_{\mathbf{j}'}] \right| \right) \\
&\qquad \geq \frac{1}{d} \left| \sum_{k,\ell=1}^{d} s(k-\ell)^p \right| \left(1 - \left(\|r\|_{\ell^1(\mathbb{Z})} - 1 \right) \left(p! \|r\|_{\ell^1(\mathbb{Z})}^{p-1} + (p!-1)/2 \right) \right) > 0, \qquad (4.12)
\end{aligned}$$

thanks to (4.3). Hence the symmetric matrix $\widetilde{C}$ is strictly diagonally dominant and invertible. One more appeal to [14, Corollary 2] yields that

$$\|\widetilde{C}^{-1}\|_{\mathrm{op}}^{-1} \geq \inf_{\mathbf{j} \in \Delta_p^{\uparrow}} \left(\mathrm{E}[\widetilde{\mathbf{Y}}_{\mathbf{j}}^2] - \sum_{\mathbf{j}' \in \Delta_p^{\uparrow}, \mathbf{j}' \neq \mathbf{j}} \left| \mathrm{E}[\widetilde{\mathbf{Y}}_{\mathbf{j}} \widetilde{\mathbf{Y}}_{\mathbf{j}'}] \right| \right), \qquad (4.13)$$

which together with (4.12) proves (4.6).

Furthermore, we deduce from (4.10) and (4.11) that

$$\|\widetilde{C}\|_1 = \|\widetilde{C}\|_\infty = \sup_{\mathbf{j}\in\Delta_p^{\uparrow}} \sum_{\mathbf{j}'\in\Delta_p^{\uparrow}} \left|\mathrm{E}[\widetilde{\mathbf{Y}}_{\mathbf{j}}\widetilde{\mathbf{Y}}_{\mathbf{j}'}]\right|$$

$$\leq \frac{1}{d}\left|\sum_{k,\ell=1}^{d} s(k-\ell)^p\right| \left(1 + \left(\|r\|_{\ell^1(\mathbb{Z})} - 1\right)\left(p!\|r\|_{\ell^1(\mathbb{Z})}^{p-1} + (p!-1)/2\right)\right). \tag{4.14}$$

Hence (4.7) follows from (4.14) and the Hölder's inequality for matrix norms. □

Lemma 4.4. *Let $f_{\mathbf{j}}^{(d)}, \mathbf{j}\in\Delta_p^{\uparrow}$ be defined in (4.1). For all $1\leq q\leq p-1$,*

$$\sum_{\mathbf{j},\mathbf{j}'\in\Delta_p^{\uparrow}} \left\|f_{\mathbf{j}}^{(d)} \otimes_q f_{\mathbf{j}'}^{(d)}\right\|^2_{\mathfrak{H}^{\otimes 2p-2q}} \leq \|r\|_{\ell^1(\mathbb{Z})}\frac{n^{2p-1}}{d}\left(\sum_{|k|\leq d}|s(k)|^{4/3}\right)^3. \tag{4.15}$$

Proof. We first compute the norm on the left-hand side of (4.15). For $\mathbf{j}=(j_1,\ldots,j_p)$ and $\mathbf{j}' = (j'_1,\ldots,j'_p)$, using the definition of $f_{\mathbf{j}}^{(d)}$ and $f_{\mathbf{j}'}^{(d)}$,

$$f_{\mathbf{j}}^{(d)} \otimes_q f_{\mathbf{j}'}^{(d)} = \frac{1}{d}\sum_{k,\ell=1}^{d} \mathrm{sym}\left(e_{j_1k}\otimes\cdots\otimes e_{j_pk}\right)\otimes_q \mathrm{sym}\left(e_{j'_1\ell}\otimes\cdots\otimes e_{j'_p\ell}\right)$$

$$= \frac{1}{d(p!)^2}\sum_{k,\ell=1}^{d}\sum_{\sigma,\tau\in S(p)}\left(e_{j_{\sigma(1)}k}\otimes\cdots\otimes e_{j_{\sigma(p)}k}\right)\otimes_q\left(e_{j'_{\tau(1)}\ell}\otimes\cdots\otimes e_{j'_{\tau(p)}\ell}\right)$$

$$= \frac{1}{d(p!)^2}\sum_{k,\ell=1}^{d}\sum_{\sigma,\tau\in S(p)} e_{j_{\sigma(q+1)}k}\otimes\cdots\otimes e_{j_{\sigma(p)}k}\otimes e_{j'_{\tau(q+1)}\ell}\otimes\cdots\otimes e_{j'_{\tau(p)}\ell}$$

$$\times\, s(k-\ell)^q\prod_{m=1}^{q} r(j_{\sigma(m)} - j'_{\tau(m)}). \tag{4.16}$$

Now we take the square norm and it yields that

$$\left\|f_{\mathbf{j}}^{(d)}\otimes_q f_{\mathbf{j}'}^{(d)}\right\|^2_{\mathfrak{H}^{\otimes 2p-2q}}$$

$$= \frac{1}{d^2(p!)^4}\sum_{k,k',\ell,\ell'=1}^{d}\sum_{\sigma,\sigma',\tau,\tau'\in S(p)} s(k-\ell)^q s(k'-\ell')^q\prod_{m=1}^{q} r(j_{\sigma(m)}-j'_{\tau(m)})r(j_{\sigma'(m)}-j'_{\tau'(m)})$$

$$\times\, s(k-k')^{p-q}s(\ell-\ell')^{p-q}\prod_{m=q+1}^{p} r(j_{\sigma(m)}-j_{\sigma'(m)})r(j'_{\tau(m)}-j'_{\tau'(m)}). \tag{4.17}$$

For $1\leq q\leq p-1$, since $|s(k)|\leq 1$ for all $k\in\mathbb{Z}$,

$$\frac{1}{d^2}\sum_{k,k',\ell,\ell'=1}^{d}\left|s(k-\ell)^q s(k'-\ell')^q s(k-k')^{p-q}s(\ell-\ell')^{p-q}\right|$$

$$\leq \frac{1}{d^2}\sum_{k,k',\ell,\ell'=1}^{d}\left|s(k-\ell)s(k'-\ell')s(k-k')s(\ell-\ell')\right| \leq \frac{1}{d}\left(\sum_{|k|\leq d}|s(k)|^{4/3}\right)^3, \tag{4.18}$$

where the second inequality follows from the same computations as in [7, page 134-135].

Moreover, for any $\sigma, \sigma', \tau, \tau' \in S(p)$,

$$\sum_{\mathbf{j},\mathbf{j}'\in\Delta_p^{\uparrow}} \left| \left(\prod_{m=1}^{q} r(j_{\sigma(m)} - j'_{\tau(m)}) r(j_{\sigma'(m)} - j'_{\tau'(m)}) \right) \left(\prod_{m=q+1}^{p} r(j_{\sigma(m)} - j_{\sigma'(m)}) r(j'_{\tau(m)} - j'_{\tau'(m)}) \right) \right|$$

$$\leq \sum_{\mathbf{j},\mathbf{j}'\in\Delta_p^{\uparrow}} \left| r(j_{\sigma(1)} - j'_{\tau(1)}) \right| \leq n^{2p-2} \sum_{k,\ell=1}^{n} |r(k-\ell)| \leq n^{2p-1} \|r\|_{\ell^1(\mathbb{Z})}. \tag{4.19}$$

Therefore, we combine (4.18), (4.19) and (4.17) to obtain (4.15). □

We are now at the position to prove Theorem 4.1

Proof of Theorem 4.1. Recall the random vectors $\widetilde{Y}_{n,d}$ and $\mathbf{Z}^{\uparrow}$ defined in (4.5). By Lemma 4.3, the covariance matrix $\widetilde{C}$ of $\mathbf{Z}^{\uparrow}$ is invertible. Hence, according to Proposition 2.1 and the identity (2.1), we have

$$\begin{aligned} d_{\text{Wass}}\left(\tilde{\mathcal{Y}}_{n,d}^{\uparrow}, \mathbf{Z}^{\uparrow}\right) &\leq \|\widetilde{C}^{-1}\|_{\text{op}} \|\widetilde{C}\|_{\text{op}}^{1/2} \left(\sum_{\mathbf{j},\mathbf{j}'\in\Delta_p^{\uparrow}} \text{Var}\left(p^{-1} \left\langle DI_p(f_{\mathbf{j}}^{(d)}), DI_p(f_{\mathbf{j}'}^{(d)}) \right\rangle_{\eta} \right) \right)^{1/2} \\ &= \|\widetilde{C}^{-1}\|_{\text{op}} \|\widetilde{C}\|_{\text{op}}^{1/2} \left(\sum_{\mathbf{j},\mathbf{j}'\in\Delta_p^{\uparrow}} O\left(\sum_{q=1}^{p-1} \|f_{\mathbf{j}}^{(d)} \otimes_q f_{\mathbf{j}'}^{(d)}\|^2_{\eta^{\otimes 2p-2r}} \right) \right)^{1/2} \\ &= \|\widetilde{C}^{-1}\|_{\text{op}} \|\widetilde{C}\|_{\text{op}}^{1/2} \|r\|_{\ell^1(\mathbb{Z})}^{1/2} \left(\sum_{|k|\leq d} |s(k)|^{4/3} \right)^{3/2} O\left(\sqrt{\frac{n^{2p-1}}{d}} \right), \end{aligned} \tag{4.20}$$

where the first equality holds by (2.2) and the second follows from Lemma 4.4. Moreover, we apply Lemma 4.3 to see that there exists a constant $c_p > 0$ such that

$$\begin{aligned} d_{\text{Wass}}\left(\tilde{\mathcal{Y}}_{n,d}^{\uparrow}, \mathbf{Z}^{\uparrow}\right) &\leq c_p \sqrt{\frac{d}{\left|\sum_{k,\ell=1}^{d} s(k-\ell)^p\right|}} \frac{\left(1 - \left(\|r\|_{\ell^1(\mathbb{Z})} - 1\right)\left(p!\|r\|_{\ell^1(\mathbb{Z})}^{p-1} + (p!-1)/2\right)\right)^{1/2}}{1 - \left(\|r\|_{\ell^1(\mathbb{Z})} - 1\right)\left(p!\|r\|_{\ell^1(\mathbb{Z})}^{p-1} + (p!-1)/2\right)} \\ &\quad \times \|r\|_{\ell^1(\mathbb{Z})}^{1/2} \sqrt{\frac{n^{2p-1}}{d} \times \left(\sum_{|k|\leq d} |s(k)|^{4/3} \right)^3}. \end{aligned} \tag{4.21}$$

Applying the argument of Step 2 on page 22 of [10] (similar to (2.6)), we conclude the proof of (4.4). □

References

[1] Bourguin, S. and Dang, T.: High dimensional regimes of non-stationary Gaussian correlated Wishart matrices.arXiv:2011.01199 (2020)

[2] Bourguin, S., Diez, C.P. and Tudor, C.A.: Limiting behavior of large correlated Wishart matrices with chaotic entries. Bernoulli **27**, no. 2, 1077–1102 (2021)

[3] Bubeck, S., Ding, J., Eldan, R. and Rácz, M.Z.: Testing high-dimensional geometry in random graphs. *Random Structures and Algorithms* **49**, no.3, 503–532 (2016)

[4] Bubeck, S. and Ganguly, S.: Entropic CLT and phase transition in high-dimensional Wishart matrices. *Int. Math. Res. Not.* no. 2, 588–606 (2018)

[5] Jiang, T. and Li, D.: Approximation of rectangular beta-laguerre ensembles and large deviations. *J. Theor. Probab.* **28**, no. 3, 804–847 (2015)

[6] Mikulincer, D.: A CLT in Stein's distance for generalized Wishart matrices and higher order tensors. arXiv:2002.10846 (2020)

[7] Nourdin, I. and Peccati, G.: Normal approximations with Malliavin calculus. From Stein's method to universality. Cambridge University Press, Cambridge, UK (2012)

[8] Nourdin, I., Peccati, G. and Réveillac, R.: Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Instit. H. Poincaré - Probab. Statist.* **46**, no. 1, 45–58 (2010).

[9] Nourdin, I. and Zheng, G.: Exchangeable pairs on Wiener chaos. High-Dimensional Probability VIII Proceedings, 277-303 (2019)

[10] Nourdin, I. and Zheng, G.: Asymptotic behavior of large Gaussian correlated Wishart matrices. arXiv:1804.06220 (2018)

[11] Nualart, D.: *The Malliavin Calculus and Related Topics.* Springer, New York. (2006)

[12] Nualart, D.and Nualart, E.: *An introduction to Malliavin calculus.* Cambridge University Press, Cambridge, UK. (2018)

[13] Serre, D.: *Matrices: Theory and applications* (second edition), Graduate Texts in Mathematics, vol. 216, Springer, New York. (2010)

[14] Varah, J. M.: A lower bound for the smallest singular value of a matrix. *Linear Algebra Appl.* **11** 3–5 (1975)

Ivan Nourdin and **Fei Pu.** University of Luxembourg, Department of Mathematics, Maison du Nombre, 6 avenue de la Fonte, L-4364 Esch-sur-Alzette, Grand Duchy of Luxembourg
Emails: `ivan.nourdin@uni.lu` and `fei.pu@uni.lu`