Linear Algebra and its Applications 473 (2015) 14–36

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Distributions of eigenvalues of large Euclidean matrices generated from l_p balls and spheres

Tiefeng Jiang [1]

School of Statistics, University of Minnesota, 224 Church Street S. E., MN 55455, United States

ARTICLE INFO

Article history:
Received 8 March 2013
Accepted 29 September 2013
Available online 15 November 2013
Submitted by W.B. Wu

MSC:
primary 60B20, 60B10
secondary 62H20, 60F15

Keywords:
Random matrix
l_p ball
l_p sphere
L_p-norm uniform distribution
Euclidean matrix
Empirical distributions of eigenvalues
Marčenko–Pastur law
Geodesic distance

ABSTRACT

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be points randomly chosen from a set $G \subset \mathbb{R}^N$ and $f(x)$ be a function. The Euclidean random matrix is given by $\mathbf{M}_n = (f(\|\mathbf{x}_i - \mathbf{x}_j\|^2))_{n\times n}$ where $\|\cdot\|$ is the Euclidean distance. When N is fixed and $n \to \infty$ we prove that $\hat{\mu}(\mathbf{M}_n)$, the empirical distribution of the eigenvalues of $\mathbf{M}_n$, converges to δ_0 for a big class of functions of $f(x)$. Assuming both N and n go to infinity proportionally, we obtain the explicit limit of $\hat{\mu}(\mathbf{M}_n)$ when G is the l_p unit ball or sphere with $p \geqslant 1$. As corollaries, we obtain the limit of $\hat{\mu}(\mathbf{A}_n)$ with $\mathbf{A}_n = (d(\mathbf{x}_i, \mathbf{x}_j))_{n\times n}$ and d being the geodesic distance on the ordinary unit sphere in $\mathbb{R}^N$. We also obtain the limit of $\hat{\mu}(\mathbf{A}_n)$ for the Euclidean distance matrix $\mathbf{A}_n = (\|\mathbf{x}_i - \mathbf{x}_j\|)_{n\times n}$. The limits are $a + bV$ where a and b are constants and V follows the Marčenko–Pastur law. The same are also obtained for other examples appeared in physics literature including $(\exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^{\gamma}))_{n\times n}$ and $(\exp(-d(\mathbf{x}_i, \mathbf{x}_j)^{\gamma}))_{n\times n}$. Our results partially confirm a conjecture by Do and Vu [14].

1. Introduction

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be random points sampled from a set $G \subset \mathbb{R}^N$. The $n \times n$ Euclidean random matrix is defined by $(g(\mathbf{x}_i, \mathbf{x}_j))_{n\times n}$, where g is a real function. See, for example, Wun and Loring [39], Cavagna, Giardina and Parisi [10], Mézard, Parisi and Zee [22], Parisi [27]. In this paper, we will study a special class of Euclidean random matrices such that

$$\mathbf{M}_n = \big(f_n\big(\|\mathbf{x}_i - \mathbf{x}_j\|^2\big)\big)_{n\times n} \tag{1.1}$$

E-mail address: jiang040@umn.edu.
[1] The research of Tiefeng Jiang was supported in part by NSF Grants DMS-0449365 and DMS-1208982.

0024-3795/$ – see front matter
http://dx.doi.org/10.1016/j.laa.2013.09.048

where $f_n(x)$ is a real function defined on $[0, \infty)$ and $\|\cdot\|$ is the Euclidean distance with

$$\|x\| = \sqrt{x_1^2 + \cdots + x_N^2} \tag{1.2}$$

for $x = (x_1, \ldots, x_N)$. Taking $f_n(x) = \sqrt{x}$ for all $n \geqslant 2$, the matrix $\mathbf{M}_n$ becomes

$$\mathbf{D}_n := \big(\|\mathbf{x}_i - \mathbf{x}_j\|\big)_{n\times n}, \tag{1.3}$$

which is referred to as the Euclidean distance matrix in some literature. See, for example, Bogomolny, Bohigas, and Schmidt [6,5], Penrose [28] and Vershik [38]. When $\mathbf{x}_i$'s are deterministic, the so-called *negative-type property* of the matrix $(\|\mathbf{x}_i - \mathbf{x}_j\|^\alpha)_{n\times n}$ with $\alpha > 0$ was studied in as early as 1937 by Schoenberg [34,32,33]. See also Reid and Sun [31] for further research in the same direction.

The matrix $\mathbf{M}_n$ belongs to a different class of random matrices from those popularly studied where their entries are independent random variables, see, Bai [2] for a survey. The primary interest in studying Euclidean random matrices is driven by the physical models including the electronic levels in amorphous systems, very diluted impurities and the spectrum of vibrations in glasses. See, e.g., Mézard, Parisi and Zee [22] and Parisi [27] for further details.

In applications, the matrix $\mathbf{M}_n$ is related to Genomics [30], Phylogeny [17,23], the geometric random graphs [29] and Statistics [7,13,16]. A relevant study by Koltchinskii and Giné [21] is to use the matrix $(g_n(\mathbf{x}_i, \mathbf{x}_j))_{n\times n}$ to approximate the spectra of integral operators.

For an $n \times n$ symmetric matrix $\mathbf{A}$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, let $\hat{\mu}(\mathbf{A})$ be the empirical law of these eigenvalues, that is,

$$\hat{\mu}(\mathbf{A}) = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i}.$$

In this paper, we will study the limiting behavior of $\hat{\mu}(\mathbf{M}_n)$ as n goes to infinity with N fixed or N going to infinity. For fixed N, when $\mathbf{x}_i$'s have a nice moment condition, in particular, G is a compact set in $\mathbb{R}^N$, we show that $\hat{\mu}(\mathbf{M}_n)$ converges weakly to δ_0, the Dirac measure at 0 as $n \to \infty$. If $n/N \to y \in (0, \infty)$, we choose G to be the unit l_p ball or sphere for all $p \geqslant 1$, we then obtain the limiting distribution of $\hat{\mu}(\mathbf{M}_n)$. In particular, when selecting different functions of $f(x)$, the matrix $\mathbf{M}_n$ in (1.1) becomes $(\|\mathbf{x}_i - \mathbf{x}_j\|^\gamma)_{n\times n}$, $(d(\mathbf{x}_i, \mathbf{x}_j)^\gamma)_{n\times n}$, $(\exp(-\lambda^2\|\mathbf{x}_i - \mathbf{x}_j\|^\gamma))_{n\times n}$ or $(\exp(-\lambda^2 d(\mathbf{x}_i, \mathbf{x}_j)^\gamma))_{n\times n}$ where $d(\cdot,\cdot)$ is the geodesic distance on the regular unit ball in $\mathbb{R}^N$. These four matrices were considered in several literatures. In particular, Schoenberg [34,32,33] and Bogomolny, Bohigas and Schmidt [5] showed that the first two matrices have the "negative type" property: all eigenvalues, except one, are non-positive; the last two are non-negative definite. In this paper we will give their explicit limiting distributions of these matrices and others in Section 2 as corollaries of our general theorems below. In particular, our results on the four matrices are consistent with their negative type or non-negative definite property.

All of the limiting distributions we have in this paper are in the form of a linear transformation of a random variable with the Marčenko–Pastur law: given a constant $y > 0$, the Marčenko–Pastur law F_y has a density function

$$p_y(x) = \begin{cases} \frac{1}{2\pi xy}\sqrt{(b-x)(x-a)}, & \text{if } x \in [a, b]; \\ 0, & \text{otherwise} \end{cases} \tag{1.4}$$

and has a point mass $1 - y^{-1}$ at $x = 0$ if $y > 1$, where $a = (1 - \sqrt{y})^2$ and $b = (1 + \sqrt{y})^2$.

Although we are concerned on random variables taking values on a compact domain, the following is a result on general domain as N is fixed.

Theorem 1. *Let $N \geqslant 1$ be fixed and $\mathbf{M}_n$ be as in (1.1). Let $\{\mathbf{x}_i;\ i \geqslant 1\}$ be $\mathbb{R}^N$-valued random variables with $\max_{i\geqslant 1} Ee^{t_0\|\mathbf{x}_i\|^\alpha} < \infty$ for some constants $\alpha > 2$ and $t_0 > 0$. Suppose $f_n \equiv f \in C^\infty[0, \infty)$ with $\omega_m := \sup_{x\geqslant 0}|f^{(m)}(x)|$ satisfying $\log\omega_m = o(m\log m)$ as $m \to \infty$. Then, with probability one, $\hat{\mu}(\mathbf{M}_n)$ converges weakly to δ_0 as $n \to \infty$.*

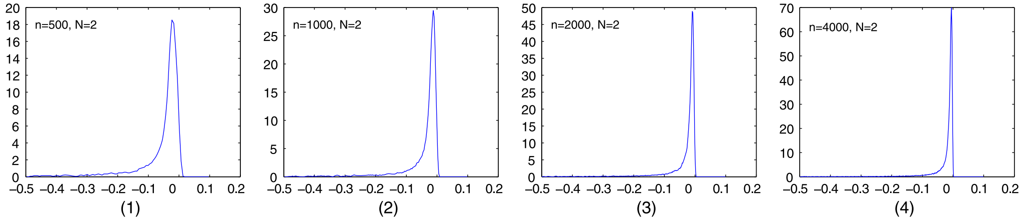


Fig. 1. Histograms of the eigenvalues of $\mathbf{D}_n = (\|\mathbf{x}_i - \mathbf{x}_j\|)_{n\times n}$ where $\mathbf{x}_1, \dots, \mathbf{x}_n$ are i.i.d. with uniform distribution on $[0,1]^2$. The curves of (1), (2), (3), (4) correspond to $n = 500, 1000, 2000, 4000$, respectively.

Assuming $\mathbf{x}_i$'s are uniformly bounded, that is, $\mathbf{x}_i$'s are sampled in a compact set such that $\max_{i\geqslant 1}\|\mathbf{x}_i\| \leqslant a$, the moment condition in Theorem 1 holds trivially. Recalling $\mathbf{M}_n = (f(\|\mathbf{x}_i - \mathbf{x}_j\|^2))_{n\times n}$, the function $f(x)$ only needs to be defined on $[0, 4a^2]$ instead of $[0,\infty)$. Making this slight change in the proof of Theorem 1, we have the following result (the proof is hence omitted). Since we do not need any correlation among $\mathbf{x}_i$'s we state it in a deterministic setting.

Theorem 2. *Let $N \geqslant 1$ be fixed and $\mathbf{M}_n$ be as in (1.1). Let $\{\mathbf{x}_i;\ i \geqslant 1\}$ be $\mathbb{R}^N$-valued vectors with $\max_{i\geqslant 1}\|\mathbf{x}_i\| \leqslant a$ for some constant $a > 0$. Suppose $f_n \equiv f \in C^\infty[0, 4a^2]$ with $\omega_m(t) = \sup_{x\in[0,t]}|f^{(m)}(x)|$ for all $t > 0$ and $m \geqslant 1$. If $\log\omega_m(4a^2) = o(m\log m)$ as $m \to \infty$, then, with probability one, $\hat{\mu}(\mathbf{M}_n)$ converges weakly to δ_0 as $n \to \infty$.*

The assumption "$\max_{i\geqslant 1}\|\mathbf{x}_i\| \leqslant a$ for some constant $a > 0$" holds for any points $\{\mathbf{x}_i;\ i \geqslant 1\}$ sampled from a bounded geometric shape G, say, polygons, annuli, ellipses and Yin–Yang graphs.

The condition $\log\omega_m = o(m\log m)$ in Theorem 1 roughly requires that $f^{(m)}(x)$ be of a small order of $o(m!)$ or $o(1/m!)$ as $m \to \infty$. For example, take $f(x) = (2\pi)^{-3/2}e^{-x/2}$ which appeared in Mézard, Parisi and Zee [22], then $|f^{(m)}(x)| = (2\pi)^{-3/2}2^{-m}e^{-x/2}$ for any $x \in \mathbb{R}$. So, $\log\omega_m(4a^2) = O(m) = o(m\log m)$ as $m \to \infty$. Hence, Theorem 2 holds for this $f(x)$.

Skipetrov and Goetschy [36] studied the matrix $\mathbf{M}_n$ with $f(x) = (\sin\sqrt{x})/\sqrt{x}$. Theorem 2 is true for this function, see the check of the condition "$\log\omega_m(4a^2) = O(m) = o(m\log m)$" in Section 4. The condition that $\log\omega_m(4a^2) = o(m\log m)$ is also satisfied if $f(x)$ is a polynomial. However, for $f(x) = \sqrt{x}$, the matrix $\mathbf{M}_n$ becomes the Euclidean distance matrix $\mathbf{D}_n = (\|\mathbf{x}_i - \mathbf{x}_j\|)_{n\times n}$ and

$$\liminf_{m\to\infty}\frac{1}{m\log m}\log\omega_m(t) \geqslant 1 \tag{1.5}$$

for any $t > 0$. See its verification in Section 4. This says that the condition $\log\omega_m(4a^2) = o(m\log m)$ is violated. We make some simulations on $\hat{\mu}(\mathbf{D}_n)$ for this case as shown in Fig. 1. It seems that $\hat{\mu}(\mathbf{D}_n)$ also converges weakly to δ_0 with a very slow convergent speed.

Theorems 1 and 2 study the behavior of eigenvalues of $\mathbf{M}_n$ when the sample points $\{\mathbf{x}_i\} \subset G \subset \mathbb{R}^N$ with N fixed regardless of the shape of G. When $N = N_n$ becomes large as n increases, Theorems 1 and 2 are no longer true. In particular, our simulations show that the behavior of $\hat{\mu}(\mathbf{M}_n)$ depends on the topology of G. In the following we consider two types of simple but non-trivial geometrical shapes of G: the l_p ball $B_{N,p}$ and its surface $S_{N,p}$ defined by

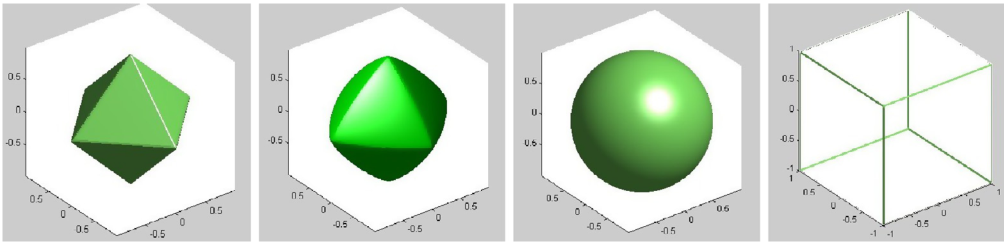
$$B_{N,p} = \left\{x \in \mathbb{R}^N;\ \|x\|_p \leqslant 1\right\} \quad \text{and} \quad S_{N,p} = \left\{x \in \mathbb{R}^N;\ \|x\|_p = 1\right\} \tag{1.6}$$

where $x = (x_1, \dots, x_N)$ and

$$\|x\|_p = \left(|x_1|^p + \cdots + |x_N|^p\right)^{1/p} \quad \text{for } 1 \leqslant p < \infty \quad \text{and} \quad \|x\|_\infty = \max_{1\leqslant i\leqslant N}|x_i|. \tag{1.7}$$

In particular, $B_{N,1}$ is the cross-polytope in $\mathbb{R}^N$; $B_{N,2}$ is the ordinary unit ball in $\mathbb{R}^N$; $B_{N,\infty}$ is the cube $[-1,1]^N$. To make our notation be consistent with (1.2), we specifically write

$$\|x\| = \|x\|_2, \qquad S^{N-1} = S_{N,2} \quad \text{and} \quad B_N(0,1) = B_{N,2} \tag{1.8}$$



(a) l_1 unit ball (b) $l_{6/5}$ unit ball (c) l_2 unit ball (d) l_∞ unit ball

Fig. 2. Comparison of l_p unit balls (surfaces) in $\mathbb{R}^3$ for $p = 1, 6/5, 2$ and ∞.

for any $x \in \mathbb{R}^N$. We give the shapes of $B_{3,p}$ and $S_{3,p}$ for $p = 1, 6/5, 2$ and ∞ in Fig. 2. This reflects the flavor of their geometries.

We next give methods to sample points from $B_{N,p}$ and $S_{N,p}$ with L_p-norm uniform distributions. Throughout the rest of the paper, for a set B in a Euclidean space, the notation $Unif_p(B)$ denotes the L_p-norm uniform distribution on B, which is explained below.

(i) *L_p-norm uniform distribution on the unit l_p-sphere.* Let $\mathbf{V}_n = (\mathbf{v}_1, \ldots, \mathbf{v}_n)_{N\times n} = (v_{ij})_{N\times n}$ where $\{v_{ij};\ i \geqslant 1,\ j \geqslant 1\}$ are i.i.d. random variables with density function

$$p(x) = \frac{p^{1-(1/p)}}{2\Gamma(1/p)} e^{-|x|^p/p}, \quad x \in \mathbb{R}. \tag{1.9}$$

Set $\mathbf{x}_i = \mathbf{v}_i/\|\mathbf{v}_i\|_p$ for $1 \leqslant i \leqslant n$. Then, by Theorem 1.1 from [37] (see also page 328 from [4] or Example 4 from [35]),

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \text{ are i.i.d. r.v.'s with } L_p\text{-norm uniform distribution on } S_{N,p} = \left\{\mathbf{x} \in \mathbb{R}^N;\ \|x\|_p = 1\right\}. \tag{1.10}$$

(ii) *L_p-norm uniform distribution on the unit l_p-ball.* Let $\{v_{ij};\ i \geqslant 1,\ j \geqslant 1\}$ and $\mathbf{v}_i$'s be as in (i). Take random variables $\{U_{n1}, \ldots, U_{nn};\ n \geqslant 1\}$ such that, for each $n \geqslant 1$, $\{U_{n1}, \ldots, U_{nn}\}$ are i.i.d. random variables taking values in $[0, 1]$ with $(U_{n1})^N \sim U[0, 1]$, and $\{U_{n1}, \ldots, U_{nn};\ n \geqslant 1\}$ are independent of $\{v_{ij};\ i \geqslant 1,\ j \geqslant 1\}$. Set $\mathbf{x}_i = U_{ni}\mathbf{v}_i/\|\mathbf{v}_i\|_p$ for $1 \leqslant i \leqslant n$. Then, by (2.16) from [4],

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \text{ are i.i.d. r.v.'s with } L_p\text{-norm uniform distribution in } B_{N,p} = \left\{\mathbf{x} \in \mathbb{R}^N;\ \|x\|_p \leqslant 1\right\}. \tag{1.11}$$

The L_p-norm uniform distribution on $S_{N,p}$ is also called the "cone probability measure", and the standard uniform distribution on $S_{N,p}$, which has a constant probability density function (pdf) equal to the reciprocal of the area of $S_{N,p}$, is also called the "surface probability measure". See, e.g., Naor and Romik [25], Naor [24] and Barthe et al. [4]. It is known from these papers that

(i) $Unif_p(B_{N,p})$ is the same as the standard uniform distribution on $B_{N,p}$ which has the pdf equal to the reciprocal of the volume of $B_{N,p}$ for all $p \geqslant 1$; (1.12)

(ii) $Unif_p(S_{N,p})$ is the same as the standard uniform distribution on $S_{N,p}$ which has the pdf equal to the reciprocal of the area of $S_{N,p}$ for $p = 1, 2, \infty$ only. (1.13)

Now, define

$$\mathbf{M}_n = \left(f\left(\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{a_N}\right)\right)_{n\times n} \quad \text{with } a_N = 2p^{\frac{2}{p}} \frac{\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})} N^{1-\frac{2}{p}}. \tag{1.14}$$

Theorem 3. *Let $p \geqslant 1$ and $\mathbf{M}_n$ be as in (1.14). Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be i.i.d. with distribution $Unif_p(S_{N,p})$ or $Unif_p(B_{N,p})$ as generated in (1.10) and (1.11). Assume that $f''(1)$ exists and $n/N \to y \in (0, \infty)$. Then, with probability one, $\hat{\mu}(\mathbf{M}_n)$ converges weakly to $a + bV$, where $a = f(0) - f(1) + f'(1)$, $b = -f'(1)$ and V has distribution F_y as in (1.4).*

Obviously, if $f'(1) = 0$, then the limiting distribution is actually the Dirac measure concentrated at constant a. The main idea of the proof of Theorem 3 follows El Karoui's decomposition of large Euclidean matrices.

By the Brunn–Minkowski inequality, the standard uniform distribution in any convex body has the log-concave property, see, for example, Gardner [19] and Pajor and Patur [26]. The conjecture by Do and Vu [14] in their paper is that Theorem 3 is always true if "distribution $Unif_p(S_{N,p})$ or $Unif_p(B_{N,p})$ as generated in (1.10) and (1.11)" there is replaced by "any probability distribution with the log-concave property". Theorem 3 partially supports the conjecture.

The reason that Theorem 3 holds for both l_p spheres and l_p balls lies in the phenomenon of the curse-of-dimensionality: when the population dimension N is large, random data tend to be around its boundary. So, if the conclusion in Theorem 3 is true for one case, it is likely to be true for the other.

Skipetrov and Goetschy [36] studied the matrix $\mathbf{M}_n$ in (1.14) with $f(x) = (\sin\sqrt{x})/\sqrt{x}$. At Section 2, we will give the exact values of a and b in Theorem 3 for this case. In the same section, similar values will be calculated for $f(x) = (2\pi)^{-3/2}e^{-x/2}$ appearing in Mézard, Parisi and Zee [22]. Bogomolny, Bohigas and Schmidt [5] showed that the matrix $(\exp(-\lambda^2\|\mathbf{x}_i - \mathbf{x}_j\|^\gamma))_{n\times n}$ is positive definite. Theorem 3 also holds for this case. We will give the values of a and b in Section 2.

In the deterministic setting, Schoenberg [34,32,33] and Reid and Sun [31] studied the matrix $(\|\mathbf{x}_i - \mathbf{x}_j\|^\alpha)_{n\times n}$ for $\alpha > 0$. Also, Bogomolny, Bohigas and Schmidt [5] investigated the same matrix. Taking $f(x) = x^{\alpha/2}$, we have the following corollary.

Corollary 1. *Given $p \geqslant 1$ and $\alpha > 0$. Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be i.i.d. with distribution $Unif_p(S_{N,p})$ or $Unif_p(B_{N,p})$ as generated in (1.10) and (1.11). Let $\mathbf{B}_n = (\|\mathbf{x}_i - \mathbf{x}_j\|^\alpha)_{n\times n}$. If $n/N \to y \in (0, \infty)$, then, with probability one, $\hat{\mu}(N^{(\frac{2}{p}-1)\frac{\alpha}{2}}\mathbf{B}_n)$ converges weakly to the distribution of $c + dV$ where V has the distribution F_y as in (1.4),*

$$c = \left(\frac{\alpha}{2} - 1\right)\left(2p^{\frac{2}{p}}\frac{\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})}\right)^{\alpha/2} \quad \text{and} \quad d = -\frac{\alpha}{2}\left(2p^{\frac{2}{p}}\frac{\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})}\right)^{\alpha/2}.$$

Now we consider the geodesic distance on the unit sphere $S^{N-1} = S_{N,2}$ in the N-dimensional Euclidean space. Let $d(x, y)$ be the geodesic distance between x and y on the sphere S^{N-1}, i.e., the shortest distance between x and y on this unit sphere. The following corollary is about the empirical distribution of a non-Euclidean distance matrix.

Corollary 2. *Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be i.i.d. random vectors with distribution $Unif_2(S^{N-1})$. Let $\mathbf{A}_n = (d(\mathbf{x}_i, \mathbf{x}_j))_{n\times n}$. If $n/N \to y \in (0, \infty)$, then, with probability one, $\hat{\mu}(\mathbf{A}_n)$ converges weakly to $(1 - \frac{\pi}{2}) - V$, where V has the distribution F_y as in (1.4).*

It was proved by Bogomolny, Bohigas and Schmidt [5] that the eigenvalues of $\mathbf{A}_n = (d(\mathbf{x}_i, \mathbf{x}_j))_{n\times n}$ are all non-positive except one. The limiting distribution $(1 - \frac{\pi}{2}) - V$ in Corollary 2 is evidently concentrated on $(-\infty, 0]$, which is consistent with their result. Furthermore, one can see that $\hat{\mu}(\mathbf{A}_n)$ and its limiting curve match very well in picture (b) from Fig. 4. The verifications of Corollaries 1 and 2 are given at the end of Section 3.2. In Section 2, we will give the corollaries for $\mathbf{M}_n = (d(\mathbf{x}_i, \mathbf{x}_j)^\gamma)$ and $\mathbf{M}_n = (\exp(-\lambda^2 d(\mathbf{x}_i, \mathbf{x}_j)^\gamma))$ appearing in Bogomolny, Bohigas and Schmidt [5].

We simulate in Section 2 the conclusions in Corollaries 1 and 2 for $p = 1$ and 2. The empirical distributions of $\mathbf{A}_n$ and $\mathbf{D}_n$ and their corresponding limiting distributions match very well. See Figs. 3 and 4.

Now let us make some comments.

Take $f_n(x) = I(x \leqslant \epsilon)$ in (1.1) where ϵ is given. The corresponding $\mathbf{M}_n$ is called the adjacency matrix of the geometric random graphs formed by vertices $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$. See, for example, Penrose [28]. Obviously, our theorems above cannot be applied to the matrix $\mathbf{M}_n = (I(d(\mathbf{x}_i, \mathbf{x}_j) \leqslant \epsilon))_{n\times n}$ since $f(x)$ is not a smooth function. There are some studies for the spectral properties of this matrix. For example, some understanding is obtained by Preciado and Jadbabaie [29]. The limiting distribution of $\hat{\mu}(\mathbf{M}_n)$, however, is still not identified yet.

Another interesting and important problem is the matrix $\mathbf{M}_n = (m_{ij})_{n\times n}$ considered in Mézard, Parisi and Zee [22] and Parisi [27] so that

$$m_{ij} = f\left(\|\mathbf{x}_i - \mathbf{x}_j\|^2\right) - u\delta_{ij} \sum_k f\left(\|\mathbf{x}_i - \mathbf{x}_k\|^2\right)$$

where u is a constant and $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$, otherwise. We expect that the limiting distribution of $\hat{\mu}(\mathbf{M}_n)$ is different from a linear transform of the Marčenko–Pastur law as seen in our main results. See also some other discussions by Bordenave [7].

The proofs of Theorems 1 and 2 are based on the decomposition method by El Karoui in his Theorem 2.4: by the Taylor expansion, we write $\mathbf{M}_n = \mathbf{U}_n + \mathbf{V}_n$ so that the rank of $\mathbf{U}_n$ is of order $o(n)$ (by choosing a suitable number of terms in the Taylor expansion) and the eigenvalues of $\mathbf{V}_n$ are very small. The sketch of the proof of Theorem 3 is as follows. We first write by the Taylor expansion again that $\mathbf{M}_n = \mathbf{U}_n + \mathbf{V}_n + \mathbf{W}_n + \epsilon_n$ so that the rank of $\mathbf{U}_n$ is at most 2, $\mathbf{V}_n$ is proportional to $\mathbf{I}_n$, $\mathbf{W}_n = \mathbf{X}_n'\mathbf{X}_n$ as in Proposition 1, and ϵ_n is negligible, then we prove in Proposition 1 that $\hat{\mu}(\mathbf{W}_n)$ converges to the Marčenko–Pastur law.

The organization of this paper is given as follows. In Section 2, we present some corollaries from Theorems 1, 2, 3 by choosing various functions of $f(x)$ appearing in the physics literature, then conduct a simulation study to compare the empirical curves and their limit curves, and end the section with a literature study in this direction. In Section 3 we prove all the results stated in this section. In Section 4, we verify rigorously some of the statements in Section 2.

2. Examples, simulations and literature study

In this section we will first present some corollaries from Theorems 1, 2 and 3. They are based on different choices of $f(x)$ appeared in the physics literature. All of the statements in this part will be checked in Section 4. We make some simulations to compare the theoretical and the empirical curves. Finally, we will review the recent progress in the study of Euclidean random matrices.

2.1. Examples

Property of negative type of matrix $\mathbf{B}_n = (\|\mathbf{x}_i - \mathbf{x}_j\|^\alpha)_{n\times n}$. Bogomolny, Bohigas and Schmidt [5] proved that, for any $0 < \alpha \leqslant 2$ and for any points $\mathbf{x}_1, \dots, \mathbf{x}_n$, the matrix $\mathbf{B}_n = (\|\mathbf{x}_i - \mathbf{x}_j\|^\alpha)_{n\times n}$ is of *negative type*: all eigenvalues of $\mathbf{B}_n$, except one, are non-positive. Schoenberg [34,32,33] showed this for $\alpha = 1$. Our Corollary 1 is consistent with this negative-type property. In fact, recall the corollary, if $n/N \to y \in (0, \infty)$, then, with probability one, $\hat{\mu}(n^{\frac{2}{p}-1}\mathbf{B}_n)$ converges weakly to the distribution of $c + dV$ where V has the distribution F_y as in (1.4). Notice that $c \leqslant 0$ and $d \leqslant 0$ for $0 < \alpha \leqslant 2$ and $V \geqslant 0$. So the support of $c + dV$ is contained in $(-\infty, 0]$.

On the other hand, our corollary also implies that $\mathbf{B}_n$ does not necessarily have the negative-type property as $\alpha > 2$. To see this, take $p = 2$. Then, for any $\alpha > 2$, let $y > 0$ satisfy $|\sqrt{y} - 1| < (1 - 2\alpha^{-1})^{1/2}$, we see that a subinterval in the support of $a + bV$ is a subset of $(0, \infty)$.

Now we give some examples below by taking special functions of $f(x)$ in Theorems 1, 2 and 3.

Example 1. Skipetrov and Goetschy [36] discussed $\mathbf{M}_n = (f(\|\mathbf{x}_i - \mathbf{x}_j\|^2))_{n\times n}$ with $f(x) = (\sin\sqrt{x})/\sqrt{x}$ for $x \neq 0$ and $f(0) = 1$. In this case, Theorem 2 is true. Now, consider the *normalized* matrix $(f(\|\mathbf{x}_i - \mathbf{x}_j\|^2/a_N))_{n\times n}$, where a_N is as in (1.14), Theorem 3 holds for this matrix with $a = 1 + \frac{\cos 1 - 3\sin 1}{2}$ and

$b = \frac{\sin 1 - \cos 1}{2}$. When $p = 2$, by using Theorem 3 again, we know $\hat{\mu}(\mathbf{M}_n)$ converges to the law of $c_1 + d_1 V$, where V has the distribution F_y as in (1.4) and

$$c_1 = 1 + \frac{\sqrt{2}\cos\sqrt{2} - 3\sin\sqrt{2}}{2\sqrt{2}} \quad \text{and} \quad d_1 = \frac{\sin\sqrt{2} - \sqrt{2}\cos\sqrt{2}}{2\sqrt{2}}.$$

Example 2. Mézard, Parisi and Zee [22] discussed $\mathbf{M}_n = (f(\|\mathbf{x}_i - \mathbf{x}_j\|^2))_{n\times n}$ with $f(x) = (2\pi)^{-3/2}e^{-x/2}$ for $x \geqslant 0$. In this case, Theorems 1 and 2 hold. Theorem 3 also holds for the normalized matrix $(f(\|\mathbf{x}_i - \mathbf{x}_j\|^2/a_N))_{n\times n}$ with $a = (2\pi)^{-3/2}(1 - \frac{3}{2}e^{-1/2})$ and $b = \frac{1}{2}(2\pi)^{-3/2}e^{-1/2}$.

Example 3. Let $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ be i.i.d. r.v.'s with distribution $Unif_2(S^{N-1})$. Recall $d(x, y)$ is the geodesic distance on S^{N-1}. Bogomolny, Bohigas and Schmidt [5] investigated the following three matrices for the signs of their eigenvalues: $(d(\mathbf{x}_i, \mathbf{x}_j)^{\gamma})_{n\times n}$ is of negative type for all $0 < \gamma \leqslant 1$; $(\exp(-\lambda^2 d(\mathbf{x}_i, \mathbf{x}_j)^{\gamma}))_{n\times n}$ is non-negative definite for $0 < \gamma \leqslant 1$; $(\exp(-\lambda^2\|\mathbf{x}_i - \mathbf{x}_j\|^{\gamma}))_{n\times n}$ is non-negative definite for $0 < \gamma \leqslant 2$ (Mézard, Parisi and Zee [22] also studied this for $\gamma = 2$). The parameter $\lambda \in \mathbb{R}$ is given. Now we present their limiting spectral distributions as corollaries of Theorem 3.

(i) For $\mathbf{M}_n = (d(\mathbf{x}_i, \mathbf{x}_j)^{\gamma})_{n\times n}$, $\hat{\mu}(\mathbf{M}_n)$ converges weakly to the distribution of $a + bV$, where V has distribution F_y as in (1.4) and

$$a = \frac{2\gamma - \pi}{2}\left(\frac{\pi}{2}\right)^{\gamma-1} \quad \text{and} \quad b = -\gamma\left(\frac{\pi}{2}\right)^{\gamma-1}$$

for all $0 < \gamma \leqslant 1$. Evidently, $a < 0$ and $b < 0$. So the support of the limiting distribution of $a + bV$ is a subset of $(-\infty, 0)$ since $V \geqslant 0$. This is consistent with the negative-type property. Also, taking $\gamma = 1$, we recover Corollary 2.

(ii) For $\mathbf{M}_n = (\exp(-\lambda^2 d(\mathbf{x}_i, \mathbf{x}_j)^{\gamma}))_{n\times n}$, we have $\hat{\mu}(\mathbf{M}_n)$ converges weakly to $a + bV$, where V has distribution F_y as in (1.4) and

$$a = 1 - e^{-\lambda^2(\pi/2)^{\gamma}} - \gamma\lambda^2\left(\frac{\pi}{2}\right)^{\gamma-1} e^{-\lambda^2(\pi/2)^{\gamma}} > 0 \quad \text{and} \quad b = \gamma\lambda^2\left(\frac{\pi}{2}\right)^{\gamma-1} e^{-\lambda^2(\pi/2)^{\gamma}} > 0.$$

Hence, the support of the limiting distribution of $a + bV$ is contained in $(0, \infty)$. This is consistent with the property that $(\exp(-\lambda^2\|\mathbf{x}_i - \mathbf{x}_j\|^{\gamma}))_{n\times n}$ is non-negative definite.

(iii) For $\mathbf{M}_n = (\exp(-\lambda^2\|\mathbf{x}_i - \mathbf{x}_j\|^{\gamma}))_{n\times n}$ with $0 < \gamma \leqslant 2$, we have $\hat{\mu}(\mathbf{M}_n)$ converges weakly to the distribution of $a + bV$, where V has distribution F_y as in (1.4) and

$$a = 1 - e^{-\lambda^2 2^{\gamma/2}} - \frac{\gamma}{2}(\lambda^2 2^{\gamma/2})e^{-\lambda^2 2^{\gamma/2}} > 0 \quad \text{and} \quad b = \frac{\gamma}{2}(\lambda^2 2^{\gamma/2})e^{-\lambda^2 2^{\gamma/2}} > 0.$$

The same is true for the non-negative definite property as discussed at the end of (ii).

Example 4. Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be i.i.d. random vectors with the L_p-norm uniform distribution on $S_{N,p}$ or $B_{N,p}$ as generated by (1.10) and (1.11) with $p \geqslant 1$. Consider $\mathbf{M}_n = (f(\|\mathbf{x}_i - \mathbf{x}_j\|^2))_{n\times n}$ with $f(x) = x^m + \alpha_1 x^{m-1} + \cdots + \alpha_m$ for fixed integer $m \geqslant 1$ and coefficients $\{\alpha_1, \dots, \alpha_m\}$. Easily, $\omega_k(t) = 0$ for all $k > m$ and $t \geqslant 0$. Thus, Theorems 1 and 2 hold.

Further, assume $p > 4m/(2m - 1)$ and $n/N \to y \in (0, \infty)$. Then, with probability one, $\hat{\mu}(N^{(\frac{2}{p}-1)m}\mathbf{M}_n)$ converges weakly to $c + dV$, where V has the law F_y as in (1.4) and

$$c = (m-1)\left(2p^{\frac{2}{p}}\frac{\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})}\right)^m \quad \text{and} \quad d = -m\left(2p^{\frac{2}{p}}\frac{\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})}\right)^m.$$

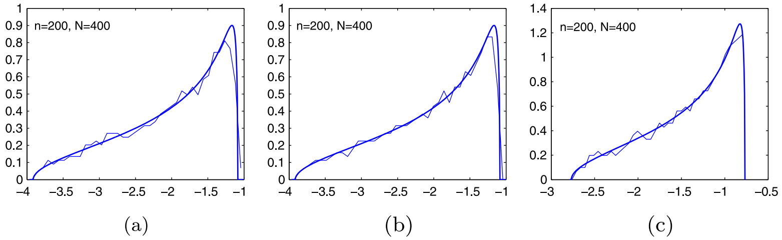


Fig. 3. Comparisons between limiting and empirical distributions for $n = 200$, $N = 400$: (a) and (b) correspond to (2.1) for the cross-polytope and its surface ($p = 1$); (c) corresponds to (2.2) for the ordinary sphere ($p = 2$). The lighter curves are the smoothed curves by taking 3-point average of the original histograms. The blacker ones are the limiting densities.

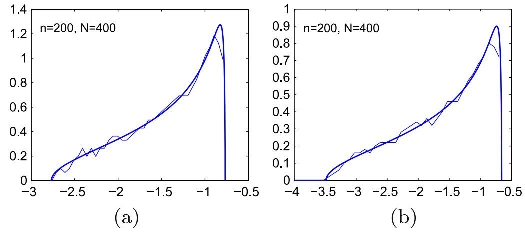


Fig. 4. Comparisons between limiting and empirical distributions for $n = 200$, $N = 400$: (a) corresponds to (2.2) for the ordinary ball ($p = 2$); (b) corresponds to Corollary 2 for the geodesic distance. The lighter and blacker curves are the same as explained in Fig. 3.

2.2. Simulations

In this section we compare the empirical curves and their limiting curves by simulation for the Euclidean distance matrix $\mathbf{D}_n = (\|\mathbf{x}_i - \mathbf{x}_j\|)_{n\times n}$ for the two special cases with $p = 1$ and 2 and for the geodesic matrix $\mathbf{A}_n = (d(\mathbf{x}_i, \mathbf{x}_j))_{n\times n}$. We first state the theoretical results case by case.

(1) *Cross-polytope and its surface*. Take $p = 1$ and $\alpha = 1$ in Corollary 1 to see that $c = d = -1$. Recall (1.12) and (1.13). Then we have the following situation: let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be i.i.d. with the standard uniform distribution on the cross-polytope $B_{N,1} = \{(x_1, \dots, x_N);\ |x_1| + \cdots + |x_N| \leqslant 1\}$ or its surface $S_{N,1} = \{(x_1, \dots, x_N);\ |x_1| + \cdots + |x_N| = 1\}$. If $n/N \to y \in (0, \infty)$, then, with probability one,

$$\hat{\mu}\big(N^{1/2}\mathbf{D}_n\big) \text{ converges weakly to the distribution of } -(V+1) \tag{2.1}$$

where V has the distribution F_y as in (1.4).

(2) *Ordinary ball and sphere*. Take $p = 2$ and $\alpha = 1$ in Corollary 1 to see that $c = d = -1/\sqrt{2}$. Then we have the following situation: let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be i.i.d. with distribution $Unif_2(S^{N-1})$ or $Unif_2(B_N(0,1))$. If $n/N \to y \in (0, \infty)$, then, with probability one,

$$\hat{\mu}(\mathbf{D}_n) \text{ converges weakly to the distribution of } -\frac{V+1}{\sqrt{2}} \tag{2.2}$$

where V has the distribution F_y as in (1.4).

(3) *Ordinary sphere with geodesic distance*. The limiting result is stated in Corollary 2.

In Figs. 3 and 4, the results stated in (1)–(3) above are simulated. We take $n = 200$ and $N = 400$ for each case. Thus, $y = n/N = 1/2$. From (1.4), we see that the limiting distribution F_y does not have the point pass at 0. It is easy to see that the empirical curve (the rugged one) and its limiting curve (the smooth one) match very well in each case.

2.3. Literature study

In this paper, we derive the limiting distributions of various Euclidean random matrices. Theorem 3 is proved by using the spirit of Theorem 2.4 from [16]. Our emphasis is the examples appeared in the physics literature. There are several recent research papers related to our study. The common parts and differences are stated next.

At the time the author writing this paper, Cheng and Singer [11] and other two authors Do and Vu [14] obtained nice results on $\mathbf{M}_n$ in the same context. The differences between their results and ours are summarized as follows:

(1) Cheng and Singer [11] assume that the distribution of $\mathbf{x}_1$ is a Gaussian random vector. Our assumption is that $\mathbf{x}_1$ follows the L_p-norm uniform distribution in the l_p ball and sphere for all $p \geqslant 1$. The two are obviously different.

(2) Do and Vu [14] give a general principle to get the limiting spectral distributions of $\mathbf{M}_n$ and $\mathbf{K}_n$ by assuming that the spectral distribution of $\mathbf{X}_n'\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)'(\mathbf{x}_1, \dots, \mathbf{x}_n)$ converges to the Marčenko–Pastur law. In our setting, we spend considerable efforts in Proposition 1 to prove that $\mathbf{X}_n'\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)'(\mathbf{x}_1, \dots, \mathbf{x}_n)$ satisfy the Marčenko–Pastur law asymptotically. However, their results do not imply ours. In fact, recall the probability measures $\hat{\mu}(\mathbf{M}_n)$ and F_y in Theorem 3. Let $m_n(z)$ and $m_y(z)$ be their Stieltjes transforms, respectively. Do and Vu showed that $\lim_{n\to\infty} E|m_n(z) - m(z)| = 0$ for all complex z with $\mathrm{Im}(z) > 0$, which is equivalent to that $\tau(\hat{\mu}(\mathbf{M}_n), F_y) \to 0$ in probability as $n \to \infty$, where $\tau(\cdot, \cdot)$ is the Prohorov distance characterizing the weak convergence of probability measures. Our Theorem 3 says that $\tau(\hat{\mu}(\mathbf{M}_n), F_y) \to 0$ *almost surely* as $n \to \infty$, which is stronger than the previous convergence in probability. The cost of this and the derivation of the limit law of $\mathbf{X}_n'\mathbf{X}_n$ is the more subtle concentration inequalities developed in Lemmas 3.1–3.4 and Corollary 3.

(3) All of Cheng and Singer [11], Do and Vu [14] and the author study $\mathbf{M}_n$ when $\mathbf{x}_1$ has the standard uniform distribution on the unit sphere S^{N-1}. In this paper we go further in this direction to obtain the spectral limits of the non-Euclidean matrices $(d(\mathbf{x}_i, \mathbf{x}_j)^\gamma)_{n\times n}$ and $(\exp(-\lambda^2 d(\mathbf{x}_i, \mathbf{x}_j)^\gamma))_{n\times n}$ appeared in physics literature, where $d(x, y)$ is the geodesic distance on the sphere.

Finally our Theorem 3 partially confirms the conjecture posed by Do and Vu [14] in their paper. The detail is given below the theorem.

3. Proofs of main results

Let $\mathbf{A}$ be an $n \times n$ symmetric matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, let $F^{\mathbf{A}}(x)$ be the empirical cumulative distribution function of these eigenvalues, that is,

$$F^{\mathbf{A}}(x) = \frac{1}{n}\sum_{i=1}^{n} I\{\lambda_i \leqslant x\}, \quad x \in \mathbb{R}. \tag{3.1}$$

For a sequence of Borel probability measures $\{\mu_n;\ n = 0, 1, 2, \dots\}$ on $\mathbb{R}$, set $F_n(x) := \mu_n((-\infty, x])$ for $n \geqslant 0$. It is well known that the following are equivalent:

(i) μ_n converges weakly to μ_0 as $n \to \infty$.
(ii) $\lim_{n\to\infty} F_n(x) = F_0(x)$ for all continuous point x of $F_0(x)$.
(iii) $\lim_{n\to\infty} \int_{\mathbb{R}} g(x)\mu_n(dx) \to \int_{\mathbb{R}} g(x)\mu_0(dx)$ for any bounded continuous function $g(x)$ defined on $\mathbb{R}$.
(iii)′ The limit in (iii) holds for any bounded Liptschitz function $g(x)$ defined on $\mathbb{R}$.
(iv) $\lim_{n\to\infty} L(F_n, F_0) = 0$, where $L(\cdot, \cdot)$ is the Lévy distance with

$$L(F_1, F_2) \leqslant \|F_1 - F_2\|_\infty := \sup_{x\in\mathbb{R}} \big|F_1(x) - F_2(x)\big|. \tag{3.2}$$

See, for example, Exercise 2.15 from [15]. In the proofs next, we will use the above equivalences from time to time.

3.1. The proof of Theorem 1

Proof of Theorem 1. Let $\eta = \alpha/(2\alpha - 4)$. For $n \geqslant \exp(e^e)$, set $\log_3 n = \log\log\log n$ and $m = m_n = [\eta(\log n)/\log_3 n] + 1$. Then, for any sequence of numbers $\{h_n;\ n \geqslant 1\}$ with $h_n = O(m)$ as $n \to \infty$, it is trivial to check that

$$m_n \to \infty, \qquad m_n = o(\log n) \quad \text{and} \quad \frac{\eta^{-1} m \log m - \log n + h_n}{\log n} \to +\infty \tag{3.3}$$

as $n \to \infty$.

Step 1. By the Taylor expansion

$$f(x) = f(0) + \sum_{k=1}^{m-1} \frac{f^{(k)}(0)}{k!} x^k + \frac{f^{(m)}(\xi)}{m!} x^m \tag{3.4}$$

where $\xi > 0$ is between 0 and x. Note that

$$\begin{aligned} \mathbf{M}_n &= \big(f\big(\|\mathbf{x}_i - \mathbf{x}_j\|^2\big)\big)_{n\times n} \\ &= f(0)\mathbf{e}\mathbf{e}' + \sum_{k=1}^{m-1} \frac{f^{(k)}(0)}{k!} \big(\|\mathbf{x}_i - \mathbf{x}_j\|^{2k}\big)_{n\times n} + \mathbf{E}_n \end{aligned} \tag{3.5}$$

where $\mathbf{e} = (1, \dots, 1)' \in \mathbb{R}^n$, $\mathbf{E}_n := \frac{1}{m!}(f^{(m)}(\xi_{ij})\|\mathbf{x}_i - \mathbf{x}_j\|^{2m})_{n\times n}$ with $0 \leqslant \xi_{ij} \leqslant \|\mathbf{x}_i - \mathbf{x}_j\|$ for all $1 \leqslant i, j \leqslant n$. Write $\|\mathbf{x}_i - \mathbf{x}_j\|^2 = \|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2 - 2\mathbf{x}_i'\mathbf{x}_j$ for all $1 \leqslant i \leqslant j \leqslant n$. Then

$$\mathbf{H}_n := \big(\|\mathbf{x}_i - \mathbf{x}_j\|^2\big)_{n\times n} = \big(\|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2\big)_{n\times n} - 2(\mathbf{x}_1, \dots, \mathbf{x}_n)'(\mathbf{x}_1, \dots, \mathbf{x}_n).$$

Since $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ is an $N \times n$ matrix, its rank and the rank of $(\mathbf{x}_1, \dots, \mathbf{x}_n)'(\mathbf{x}_1, \dots, \mathbf{x}_n)$ are both less than or equal to N. Besides, it is easy to check that the rank of $(\|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2)_{n\times n}$ is at most 2. It follows that $rank(\mathbf{H}_n) \leqslant N + 2 := q$. Notice $(\|\mathbf{x}_i - \mathbf{x}_j\|^{2k})_{n\times n} = \mathbf{H}_n \circ \mathbf{H}_n \circ \dots \circ \mathbf{H}_n$, where there are k's many $\mathbf{H}_n$ in the Hadamard product. Theorem 5.1.7 from [20] says that, if $\mathbf{A} \circ \mathbf{B}$ is the Hadamard product of $\mathbf{A} = (a_{ij})_{m\times n}$ and $\mathbf{B} = (b_{ij})_{m\times n}$, that is, $\mathbf{A} \circ \mathbf{B} = (a_{ij}b_{ij})_{m\times n}$, then $rank(\mathbf{A} \circ \mathbf{B}) \leqslant rank(\mathbf{A}) \cdot rank(\mathbf{B})$. Thus, the rank of $(\|\mathbf{x}_i - \mathbf{x}_j\|^{2k})_{n\times n}$ is at most q^k. Therefore, use the inequality $rank(\mathbf{U} + \mathbf{V}) \leqslant rank(\mathbf{U}) + rank(\mathbf{V})$ for any matrices $\mathbf{U}$ and $\mathbf{V}$ to obtain that

$$\text{the rank of } f(0)\mathbf{e}\mathbf{e}' + \sum_{k=1}^{m-1} \frac{f^{(k)}(0)}{k!} \big(\|\mathbf{x}_i - \mathbf{x}_j\|^{2k}\big)_{n\times n} \leqslant 1 + \sum_{k=1}^{m-1} q^k = \frac{q^m - 1}{q - 1} \leqslant q^m. \tag{3.6}$$

Thus, by Lemma 2.2 from [2] we have from (3.5) and (3.6) that

$$L\big(F^{\mathbf{M}_n}, F^{\mathbf{E}_n}\big) \leqslant \big\|F^{\mathbf{M}_n} - F^{\mathbf{E}_n}\big\|_\infty \leqslant \frac{q^m}{n} \to 0 \tag{3.7}$$

as $n \to \infty$ since $m = o(\log n)$ where $\|v\|_\infty = \sup_{x\in\mathbb{R}} |v(x)|$ for any function $v(x)$ defined on $\mathbb{R}$.

Step 2. We now estimate $\mathbf{E}_n$. Review $\mathbf{E}_n = \frac{1}{m!}(f^{(m)}(\xi_{ij})\|\mathbf{x}_i - \mathbf{x}_j\|^{2m})_{n\times n}$. Let $\mathbf{O}_n$ be an $n \times n$ matrix whose entries are all equal to zero. Then, by Lemma 2.3 from [2] (see also (2.16) from [8]),

$$\begin{aligned} L^3\big(F^{\mathbf{E}_n}, F^{\mathbf{O}_n}\big) &\leqslant \frac{1}{n} \sum_{1\leqslant i,j\leqslant n} (\mathbf{E}_n)_{ij}^2 \\ &\leqslant \frac{\omega_m^2}{n(m!)^2} \sum_{1\leqslant i<j\leqslant n} \|\mathbf{x}_i - \mathbf{x}_j\|^{4m} \\ &\leqslant \frac{C^m\omega_m^2}{n(m!)^2} \sum_{1\leqslant i<j\leqslant n} \big(\|\mathbf{x}_i\|^{4m} + \|\mathbf{x}_j\|^{4m}\big) \leqslant \frac{C^m\omega_m^2}{(m!)^2} \sum_{i=1}^{n} \|\mathbf{x}_i\|^{4m} \end{aligned}$$

where the constant C is chosen such that $(x + y)^{4m} \leqslant C^m(x^{4m} + y^{4m})$ for all $x \geqslant 0$ and $y \geqslant 0$. For any $\epsilon > 0$, by the Markov inequality,

$$P\big(L^3\big(F^{\mathbf{E}_n}, F^{\mathbf{O}_n}\big) > \epsilon\big) \leqslant \frac{1}{\epsilon} \cdot \frac{\omega_m^2 C^m}{(m!)^2} \sum_{i=1}^{n} E\big(\|\mathbf{x}_i\|^{4m}\big). \tag{3.8}$$

Recall the assumption $\max_{i\geqslant 1} Ee^{t_0\|\mathbf{x}_i\|^\alpha} = C' < \infty$ for constants $\alpha > 2$ and $t_0 > 0$. Set $\beta = 4m/\alpha$. We claim that there exists a constant $C_1 \geqslant 1$ satisfying

$$\sum_{i=1}^{n} E\big(\|\mathbf{x}_i\|^{4m}\big) \leqslant C_1(mn)(C_1\beta)^\beta \tag{3.9}$$

as n is sufficiently large. If so, by (3.8) we get

$$P\big(L^3\big(F^{\mathbf{E}_n}, F^{\mathbf{O}_n}\big) > \epsilon\big) \leqslant \frac{C_1}{\epsilon} \cdot \frac{(mn)\omega_m^2 C^m (C_1\beta)^\beta}{(m!)^2} \tag{3.10}$$

as n is sufficiently large. The Stirling formula (see, for example, Gamelin [18]) says that

$$\log \Gamma(z) = z\log z - z - \frac{1}{2}\log z + \log\sqrt{2\pi} + \frac{1}{12z} + O\left(\frac{1}{x^3}\right) \tag{3.11}$$

as $x = \mathrm{Re}(z) \to +\infty$. Remember $\Gamma(m+1) = m!$. Take $z = m+1$ in (3.11) and use the assumption $\log\omega_m = o(m\log m)$ to have that the logarithm of the RHS of (3.10) is equal to

$$\begin{aligned}
&\log m + \log n + o(m\log m) + m(\log C) + \frac{4m}{\alpha}\log\frac{4m}{\alpha} - \frac{4m}{\alpha}\log C_1 - 2m\log m \\
&\quad + 2m + \log m + O(1) \\
&= \log n - \left(2 - \frac{4}{\alpha} + o(1)\right) m\log m + O(m) = r_n\log n
\end{aligned}$$

such that $r_n \to -\infty$ as $n \to \infty$ by (3.3). It follows that $P(L^3(F^{\mathbf{E}_n}, F^{\mathbf{O}_n}) > \epsilon) = O(n^{-2})$ as $n \to \infty$. By the Borel–Cantelli lemma, we obtain

$$L\big(F^{\mathbf{E}_n}, F^{\mathbf{O}_n}\big) \to 0 \quad \text{a.s.}$$

as $n \to \infty$. This and (3.7) conclude that $\lim_{n\to\infty} L(F^{\mathbf{M}_n}, F^{\mathbf{O}_n}) = 0$ a.s. Thus, $F^{\mathbf{M}_n}$ converges weakly to δ_0 since $F^{\mathbf{O}_n}$ is equal to the cumulative distribution function of δ_0.

Step 3. Now we turn to prove (3.9). In fact, set $C' = \max_{i\geqslant 1} Ee^{t_0\|\mathbf{x}_i\|^\alpha}$. Then

$$\begin{aligned}
E\big(\|\mathbf{x}_i\|^{4m}\big) = t_0^{-\beta}\cdot E\big(\big(t_0\|\mathbf{x}_i\|^\alpha\big)^\beta\big) &= \beta t_0^{-\beta}\int_0^\infty t^{\beta-1}P\big(t_0\|\mathbf{x}_i\|^\alpha \geqslant t\big)\,dt \\
&\leqslant \big(\beta C'\big)t_0^{-\beta}\int_0^\infty t^{\beta-1}e^{-t}\,dt = \beta C' t_0^{-\beta}\cdot\Gamma(\beta),
\end{aligned} \tag{3.12}$$

where the formula $E(Z^\beta) = \beta\int_0^\infty t^{\beta-1}P(Z \geqslant t)\,dt$ for $Z \geqslant 0$ is used above. Recall (3.11). We know $\Gamma(\beta) \leqslant \beta^\beta e^{-\beta}$ as n is large enough (note $\beta = 4m/\alpha$ and $m = m_n$ defined above (3.3)). This and (3.12) yield (3.9). □

3.2. The proof of Theorem 3

Lemma 3.1. *Let ξ be a random variable with density function $p(x)$ as in (1.9). Then*

$$E\big(|\xi|^t\big) = \frac{\Gamma(\frac{t+1}{p})}{\Gamma(\frac{1}{p})}p^{t/p}, \quad t > 0.$$

In particular, $E(|\xi|^p) = 1$.

Proof. By symmetry,

$$E(|\xi|^t) = \frac{p^{1-(1/p)}}{\Gamma(\frac{1}{p})} \int_0^\infty x^t e^{-x^p/p}\,dx.$$

Set $y = x^p/p$, then $x = p^{1/p}y^{1/p}$ and $dx = p^{\frac{1}{p}-1}y^{\frac{1}{p}-1}\,dy$. Thus the above integral is equal to

$$\frac{1}{\Gamma(\frac{1}{p})}p^{t/p}\int_0^\infty y^{\frac{t+1}{p}-1}e^{-y}\,dy = \frac{\Gamma(\frac{t+1}{p})}{\Gamma(\frac{1}{p})}p^{t/p}. \qquad \square$$

Lemma 3.2. *Let $p \geqslant 1$. Let U_{ij}'s, v_{ij}'s and $\mathbf{v}_i$'s be as in (1.10) and (1.11). Assume $n/N \to y \in (0,\infty)$. Then, as $n \to \infty$,*

(i) $$\frac{N^{\frac{2}{p}-2}}{(\log N)^2}\sum_{i=1}^n \frac{\|\mathbf{v}_i\|^2}{\|\mathbf{v}_i\|_p^2} \to 0 \quad a.s.;$$

(ii) $$\frac{\sqrt{N}}{\log N}\max_{1\leqslant i\leqslant n}\left|1 - \frac{\|\mathbf{v}_i\|_p}{N^{1/p}}\right| \to 0 \quad a.s.$$

The convergence rates given in the lemma, for instance, $(\log N)^2/N^{\frac{2}{p}-2}$ in (i), may not be the best ones. However, we make them precise enough to prove Theorem 3 rather than pursue the exact speeds with lengthy arguments.

Proof of Lemma 3.2. (i) Easily,

$$H_n := \frac{N^{\frac{2}{p}-2}}{(\log N)^2}\sum_{i=1}^n \frac{\|\mathbf{v}_i\|^2}{\|\mathbf{v}_i\|_p^2} \leqslant 2y\frac{N^{\frac{2}{p}-1}}{(\log N)^2}\cdot\max_{1\leqslant i\leqslant n}\frac{\|\mathbf{v}_i\|^2}{\|\mathbf{v}_i\|_p^2}$$

as n is sufficiently large. Therefore, for any $\epsilon > 0$,

$$\begin{aligned} P(H_n \geqslant 2y\epsilon) &\leqslant nP\left(\frac{\|\mathbf{v}_1\|^2}{\|\mathbf{v}_1\|_p^2} \geqslant \epsilon\frac{(\log N)^2}{N^{\frac{2}{p}-1}}\right) \\ &\leqslant nP\left(\|\mathbf{v}_1\|^2 \geqslant \epsilon N\log N\right) + nP\left(\|\mathbf{v}_1\|_p^2 \leqslant N^{2/p}(\log N)^{-1}\right) \end{aligned} \tag{3.13}$$

as n is sufficiently large. From (1.9) and Lemma 3.1, we know that

$$E\left(|v_{11}|^p\right) = 1 \quad\text{and}\quad Ee^{t_0|v_{11}|^p} < \infty \tag{3.14}$$

where $t_0 = 1/(2p) > 0$. By the Cramér large deviation (see, e.g., Dembo and Zeitouni [12]), there exists $\delta > 0$ such that

$$\begin{aligned} P\left(\|\mathbf{v}_1\|_p^2 \leqslant N^{2/p}(\log N)^{-1}\right) &= P\left(\frac{1}{N}\sum_{k=1}^N |v_{k1}|^p \leqslant (\log N)^{-p/2}\right) \\ &\leqslant P\left(\frac{1}{N}\sum_{k=1}^N |v_{k1}|^p \leqslant \frac{1}{2}\right) \leqslant e^{-N\delta} \end{aligned} \tag{3.15}$$

as n is large enough. By Lemma 6.4 from [9], there exists $C > 0$ such that

$$P\left(\|\mathbf{v}_1\|^2 \geqslant \epsilon N\log N\right) \leqslant P\left(\frac{|\sum_{k=1}^N(v_{k1}^2 - Ev_{k1}^2)|}{\sqrt{N}\log N} \geqslant 1\right) \leqslant e^{-C(\log N)^2} \tag{3.16}$$

as n is sufficiently large. Combining the above we see that $\sum_{n\geqslant 1} P(H_n \geqslant 2y\epsilon) < \infty$ for any $\epsilon > 0$. Then, conclusion (i) follows from the Borel–Cantelli lemma.

(ii) By the inequality $|1-t^{\alpha}| \leqslant |1-t|$ for all $t \geqslant 0$ and $0 < \alpha \leqslant 1$, we get

$$\frac{\sqrt{N}}{\log N} \cdot \max_{1\leqslant i\leqslant n}\left|1-\frac{\|\mathbf{v}_i\|_p}{N^{1/p}}\right| \leqslant \max_{1\leqslant i\leqslant n}\frac{|N-\|\mathbf{v}_i\|_p^p|}{\sqrt{N}\log N}.$$

Now, use (3.14) to write $\|\mathbf{v}_i\|_p^p - N = \sum_{k=1}^N(|v_{ki}|^p - E|v_{ki}|^p)$. Then, replace "$\|\mathbf{v}_1\|^2$" and "$|v_{k1}|^2$" with "$\|\mathbf{v}_1\|_p^p$" and "$|v_{k1}|^p$" in (3.16), respectively, the conclusion (ii) is obtained by using the same argument as in (3.16) and the union bound. □

Lemma 3.3. *Assume $p \geqslant 1$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d. with distribution $Unif_p(B_{N,p})$ or $Unif_p(S_{N,p})$. Assume $n/N \to y \in (0,\infty)$. Then, for any $t > 0$, there exists a constant $\delta > 0$ such that $P(\max_{1\leqslant i<j\leqslant n}|\mathbf{x}_i'\mathbf{x}_j| \geqslant tN^{\frac{1}{2}-\frac{2}{p}}\log N) \leqslant e^{-\delta(\log N)^2}$ as n is sufficiently large.*

The bound "$e^{-\delta(\log N)^2}$" given in the lemma may not be tight. However it is precise enough for the proof of Proposition 1. The same is true for Lemma 3.4.

Proof of Lemma 3.3. Recall the sampling schemes in (1.10) and (1.11). For both the case of $Unif_p(B_{N,p})$ and that of $Unif_p(S_{N,p})$, we have that

$$\max_{1\leqslant i<j\leqslant n}\left|\mathbf{x}_i'\mathbf{x}_j\right| \leqslant \max_{1\leqslant i<j\leqslant n}\frac{|\mathbf{v}_i'\mathbf{v}_j|}{\|\mathbf{v}_i\|_p\cdot\|\mathbf{v}_j\|_p}$$

where $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are i.i.d. $\mathbb{R}^N$-dimensional random vectors whose nN entries are i.i.d. random variables with the density function $p(x)$ as in (1.9). Thus,

$$\begin{aligned} &P\Big(\max_{1\leqslant i<j\leqslant n}\left|\mathbf{x}_i'\mathbf{x}_j\right| > tN^{\frac{1}{2}-\frac{2}{p}}\log N\Big) \\ &\leqslant n^2 P\left(\frac{|\mathbf{v}_1'\mathbf{v}_2|}{\|\mathbf{v}_1\|_p\cdot\|\mathbf{v}_2\|_p} > tN^{\frac{1}{2}-\frac{2}{p}}\log N\right) \\ &\leqslant 2n^2 P\left(\|\mathbf{v}_1\|_p^p \leqslant \frac{1}{2}N\right) + P\big(\left|\mathbf{v}_1'\mathbf{v}_2\right| \geqslant C_p\sqrt{N}\log N\big) \end{aligned} \tag{3.17}$$

where C_p is a constant depending on p only. Note that $\mathbf{v}_1'\mathbf{v}_2 = \sum_{i=1}^N v_{i1}v_{i2}$ where $\{v_{ij};\ i \geqslant 1,\ j \geqslant 1\}$ are i.i.d. random variables with density function as in (1.9). Evidently, there is a constant $C_p > 0$ depending on p only such that

$$|v_{11}v_{12}|^{p/2} \leqslant \left(\frac{v_{11}^2+v_{12}^2}{2}\right)^{p/2} \leqslant C_p\big(|v_{11}|^p + |v_{12}|^p\big).$$

This joint with (3.14) implies that $Ee^{t_0|v_{11}v_{12}|^{p/2}} < \infty$ for some $t_0 > 0$. By Lemma 6.4 from [9], for some constant $C_p' > 0$, we have $P(|\mathbf{v}_1'\mathbf{v}_2| \geqslant C_p\sqrt{N}\log N) \leqslant e^{-C_p'(\log N)^2}$ as n is sufficiently large. This combining with (3.15) and (3.17) leads to the desired conclusion. □

Lemma 3.4. *Given $p \geqslant 1$. Let a_N be as in (1.14) and $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be as in Lemma 3.3. Assume $n/N \to y \in (0,\infty)$. Then, for any $t > 0$, there exists a constant $\delta = \delta_{p,t} > 0$ such that*

$$P\left(\frac{\sqrt{N}}{\log N}\cdot\max_{1\leqslant i\leqslant n}\left|\frac{\|\mathbf{x}_i\|^2}{a_N}-\frac{1}{2}\right| \geqslant t\right) \leqslant e^{-\delta(\log N)^2} \tag{3.18}$$

as n is sufficiently large.

Proof. First,

$$P\left(\frac{\sqrt{N}}{\log N}\cdot\max_{1\leqslant i\leqslant n}\left|\frac{\|\mathbf{x}_i\|^2}{a_N}-\frac{1}{2}\right|>t\right)\leqslant nP\left(\frac{\sqrt{N}}{\log N}\cdot\left|\frac{\|\mathbf{x}_1\|^2}{a_N}-\frac{1}{2}\right|>t\right) \tag{3.19}$$

where $\mathbf{x}_1\in\mathbb{R}^N$ follows the L_p-norm uniform distribution on $S_{N,p}$ or $B_{N,p}$.

Case (i): $\mathbf{x}_1$ *follows the L_p-norm uniform distribution on $S_{N,p}$*. From (1.10), we know $\mathbf{x}_1=\mathbf{v}/\|\mathbf{v}\|_p$ for some $\mathbf{v}=(v_1,\dots,v_N)$ where v_i's are i.i.d. with the density function as in (1.9). By (3.14),

$$\frac{\|\mathbf{v}\|_p^2}{N^{2/p}}=\left(\frac{\sum_{i=1}^N|v_i|^p}{N}\right)^{2/p}=\left(1+\frac{\sum_{i=1}^N(|v_i|^p-E|v_i|^p)}{N}\right)^{2/p}.$$

From Lemma 6.4 from [9], there exists a constant $\delta_p>0$ such that, for any $s>0$,

$$P\Bigg(\underbrace{\frac{|\sum_{i=1}^N(|v_i|^p-E|v_i|^p)|}{N}\geqslant s\frac{\log N}{2\sqrt{N}}}_{E_{n,1}}\Bigg)\leqslant e^{-\delta_p s^2(\log N)^2} \tag{3.20}$$

as n is sufficiently large. Trivially, there exists a constant $C_p>0$ such that $|(1+x)^{2/p}-1|\leqslant C_p x$ for $x>0$ small enough. Hence, for any $s>0$,

$$\left|\frac{\|\mathbf{v}\|_p^2}{N^{2/p}}-1\right|\leqslant(C_p s)\frac{\log N}{2\sqrt{N}}$$

on $E_{n,1}^c$ as n is sufficiently large. Consequently,

$$\left|\frac{N^{2/p}}{\|\mathbf{v}\|_p^2}-1\right|\leqslant(C_p s)\frac{\log N}{\sqrt{N}} \tag{3.21}$$

since $|1-x^{-1}|\leqslant 2x$ for all x close to 1 enough. By Lemma 6.4 from [9] again, there exists $\delta_p'>0$ such that, for any $s>0$,

$$P\Bigg(\underbrace{\left|\frac{\|\mathbf{v}\|^2}{NE(v_1^2)}-1\right|\geqslant s\frac{\log N}{\sqrt{N}}}_{E_{n,2}}\Bigg)\leqslant e^{-\delta_p' s^2(\log N)^2} \tag{3.22}$$

as n is sufficiently large. From Lemma 3.1, we know $E(v_1^2)=(\Gamma(3/p)/\Gamma(1/p))p^{2/p}$. By the definition of a_N as in (1.14), we see that

$$\left|\frac{\|\mathbf{x}_1\|^2}{a_N}-\frac{1}{2}\right|=\frac{1}{2}\cdot\left|\frac{N^{2/p}}{\|\mathbf{v}\|_p^2}\cdot\frac{\|\mathbf{v}\|^2}{NE(v_1^2)}-1\right|.$$

Using the fact $|ab-1|\leqslant|a-1|+|b-1|+|a-1|\cdot|b-1|$ for any $a,b\in\mathbb{R}$, we have from (3.21) and (3.22) that

$$\begin{aligned}\left|\frac{\|\mathbf{x}_1\|^2}{a_N}-\frac{1}{2}\right|&\leqslant(C_p s)\frac{\log N}{\sqrt{N}}+s\frac{\log N}{\sqrt{N}}+C_p s^2\frac{\log N}{\sqrt{N}}\cdot\frac{\log N}{\sqrt{N}}\\&\leqslant(C_p+2)s\frac{\log N}{\sqrt{N}}\end{aligned}$$

on $E_{n,1}^c\cap E_{n,2}^c$ as n is sufficiently large, where $E_{n,1}$ and $E_{n,2}$ are as in (3.20) and (3.22). This gives that

$$P\left(\frac{\sqrt{N}}{\log N}\cdot\left|\frac{\|\mathbf{x}_1\|^2}{a_N}-\frac{1}{2}\right|>(C_p+2)s\right)\leqslant 2e^{-\delta_p''^2 s^2(\log N)^2}$$

for any $s>0$ as n is large enough, where $\delta_p''=\min\{\delta_p,\delta_p'\}>0$. Take $s=t(C_p+2)^{-1}$ in the above inequality and use (3.19) to yield (3.18).

Case (ii): $\mathbf{x}_1$ *follows the* L_p*-norm uniform distribution on* $B_{N,p}$. By (1.11), for some random variable $U_N\in[0,1]$ with $(U_N)^N\sim Unif([0,1])$, $\mathbf{x}_1=U_N\mathbf{v}/\|\mathbf{v}\|_p$. Set $\mathbf{y}=\mathbf{v}/\|\mathbf{v}\|_p$. From the conclusion in case (i), for any $t>0$, there exists a constant $\delta=\delta_{p,t}>0$ such that

$$P\left(\frac{\sqrt{N}}{\log N}\cdot\left|\frac{\|\mathbf{y}\|^2}{a_N}-\frac{1}{2}\right|\geqslant t\right)\leqslant e^{-\delta(\log N)^2} \tag{3.23}$$

as n is sufficiently large. On the other hand,

$$P\left(1-U_N\geqslant\frac{(\log N)^2}{N}\right)=\left(1-\frac{(\log N)^2}{N}\right)^N\leqslant e^{-(\log N)^2} \tag{3.24}$$

as n is large enough. If $\frac{\sqrt{N}}{\log N}\cdot|\frac{\|\mathbf{y}\|^2}{a_N}-\frac{1}{2}|<t$ and $1-U_N<\frac{(\log N)^2}{N}$, then by a similar discussion in the case (i), $\frac{\sqrt{N}}{\log N}\cdot|U_N\frac{\|\mathbf{y}\|^2}{a_N}-\frac{1}{2}|<2t$ as n is sufficiently large. It follows from (3.23) and (3.24) that

$$P\left(\frac{\sqrt{N}}{\log N}\cdot\left|\frac{\|\mathbf{x}_1\|^2}{a_N}-\frac{1}{2}\right|\geqslant 2t\right)=P\left(\frac{\sqrt{N}}{\log N}\cdot\left|U_N\frac{\|\mathbf{y}\|^2}{a_N}-\frac{1}{2}\right|\geqslant 2t\right)\leqslant 2e^{-K(\log N)^2}$$

as n is sufficiently large, where $K=\min\{\delta,1\}$. This inequality and (3.19) yield the desired conclusion. □

Corollary 3. *Given $p\geqslant 1$. Let a_N be as in (1.14) and $\mathbf{x}_1,\dots,\mathbf{x}_n$ be as in Lemma 3.3. Assume $n/N\to y\in(0,\infty)$. For any $\delta>0$, the following hold.*

(i) *Set $E_n=\{\max_{1\leqslant i<j\leqslant n}|\frac{1}{a_N}\|\mathbf{x}_i-\mathbf{x}_j\|^2-1|<\delta\}$ for $n\geqslant 2$. Then $\sum_{n=2}^{\infty}P(E_n^c)<\infty$;*
(ii) *As $n\to\infty$,*

$$\sum_{i=1}^n\left(\frac{\|\mathbf{x}_i\|^2}{a_N}-\frac{1}{2}\right)^4\to 0\quad a.s.\quad and\quad \frac{1}{n}\sum_{1\leqslant i<j\leqslant n}\left(a_N^{-1}\mathbf{x}_i'\mathbf{x}_j\right)^4\to 0\quad a.s.$$

Proof. (i) Write $\|\mathbf{x}_i-\mathbf{x}_j\|^2=\|\mathbf{x}_i\|^2+\|\mathbf{x}_j\|^2-2\mathbf{x}_i'\mathbf{x}_j$. Then

$$\max_{1\leqslant i<j\leqslant n}\left|\frac{1}{a_N}\|\mathbf{x}_i-\mathbf{x}_j\|^2-1\right|\leqslant 2\cdot\max_{1\leqslant i\leqslant n}\left|\frac{\|\mathbf{x}_i\|^2}{a_N}-\frac{1}{2}\right|+2\cdot\max_{1\leqslant i<j\leqslant n}\frac{1}{a_N}\left|\mathbf{x}_i'\mathbf{x}_j\right|.$$

Recall $a_N=2p^{\frac{2}{p}}\Gamma(\frac{3}{p})/\Gamma(\frac{1}{p})N^{1-\frac{2}{p}}$. It follows that

$$E_n^c\subset\left\{\max_{1\leqslant i\leqslant n}\left|\frac{\|\mathbf{x}_i\|^2}{a_N}-\frac{1}{2}\right|\geqslant\frac{\delta}{4}\right\}\cup\left\{\max_{1\leqslant i<j\leqslant n}\left|\mathbf{x}_i'\mathbf{x}_j\right|\geqslant\frac{\delta}{4}a_N\right\}.$$

Evidently, for any $t>0$, the last event is smaller than $\{\max_{1\leqslant i<j\leqslant n}|\mathbf{x}_i'\mathbf{x}_j|\geqslant tN^{\frac{1}{2}-\frac{2}{p}}\log N\}$ as n is large enough. We then get (i) from Lemmas 3.3 and 3.4.

(ii) From the Borel–Cantelli lemma and Lemma 3.4, we see that

$$\frac{\sqrt{N}}{\log N}\cdot\max_{1\leqslant i\leqslant n}\left|\frac{\|\mathbf{x}_i\|^2}{a_N}-\frac{1}{2}\right|\to 0\quad\text{a.s.}$$

as $n\to\infty$. Consequently,

$$\sum_{i=1}^n\left(\frac{\|\mathbf{x}_i\|^2}{a_N}-\frac{1}{2}\right)^4\leqslant\left(n^{1/4}\max_{1\leqslant i\leqslant n}\left|\frac{\|\mathbf{x}_i\|^2}{a_N}-\frac{1}{2}\right|\right)^4=O\left(\frac{(\log N)^4}{N}\right)\to 0\quad\text{a.s.}$$

as $n \to \infty$. So the first limit in (ii) holds. Furthermore, by the Borel–Cantelli lemma and Lemma 3.3, we obtain

$$\frac{N^{\frac{2}{p}-\frac{1}{2}}}{\log N} \cdot \max_{1 \leqslant i < j \leqslant n} \left|\mathbf{x}_i'\mathbf{x}_j\right| \to 0 \quad \text{a.s.}$$

as $n \to \infty$. Thus, we use $a_N = 2p^{\frac{2}{p}}\Gamma(\frac{3}{p})/\Gamma(\frac{1}{p})N^{1-\frac{2}{p}}$ to have

$$\frac{1}{n}\sum_{1 \leqslant i < j \leqslant n} \left(a_N^{-1}\mathbf{x}_i'\mathbf{x}_j\right)^4 \leqslant C_p\Big(N^{\frac{2}{p}-\frac{3}{4}} \max_{1 \leqslant i < j \leqslant n} \left|\mathbf{x}_i'\mathbf{x}_j\right|\Big)^4 = O\left(\frac{(\log N)^4}{N}\right) \to 0 \quad \text{a.s.}$$

where C_p is a constant not depending on n. We then obtain the second limit in (ii). □

For integer $p \geqslant 1$, define

$$c_y = \frac{\Gamma(\frac{1}{p})}{\Gamma(\frac{3}{p})} p^{-2/p} y^{1-\frac{2}{p}}. \tag{3.25}$$

One of the important parts in proving Theorem 3 is the following result.

Proposition 1. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d. random variables with distribution $Unif_p(S_{N,p})$ or $Unif_p(B_{N,p})$ for $p \geqslant 1$ as generated in (1.10) and (1.11), respectively. Write $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$. If $n/N \to y \in (0, +\infty)$ then, with probability one, $\hat{\mu}(c_y n^{\frac{2}{p}-1}\mathbf{X}_n'\mathbf{X}_n)$ converges weakly to F_y as in (1.4), where c_y is defined as in (3.25).*

The case for $Unif_p(B_{N,p})$ in Proposition 1 is due to Aubrun [1] and Pajor and Patur [26]. The case for $Unif_p(S_{N,p})$ is new.

Proof of Proposition 1. Recall $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are i.i.d. random vectors with the L_p-norm uniform distribution in $S_{N,p}$ as in (1.10). We have that

$$\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)_{N\times n} \quad \text{with } \mathbf{x}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|_p}$$

for $i = 1, \ldots, n$. Define

$$\tilde{\mathbf{X}}_n = N^{-1/p}(\mathbf{v}_1, \ldots, \mathbf{v}_n)_{N\times n}.$$

Set $b_n = n^{\frac{2}{p}-1}$. By Lemma 2.7 from [2],

$$\begin{aligned} &L^4\big(F^{b_n\mathbf{X}_n'\mathbf{X}_n}, F^{b_n\tilde{\mathbf{X}}_n'\tilde{\mathbf{X}}_n}\big) \\ &\quad \leqslant \frac{2b_n^2}{n^2} \cdot tr\left(\left(\mathbf{X}_n - \frac{\mathbf{V}_n}{N^{1/p}}\right)'\left(\mathbf{X}_n - \frac{\mathbf{V}_n}{N^{1/p}}\right)\right) tr\left(\mathbf{X}_n'\mathbf{X}_n + \frac{\mathbf{V}_n'\mathbf{V}_n}{N^{2/p}}\right). \end{aligned} \tag{3.26}$$

Further, by the standard law of large numbers, $(nN)^{-1}\sum_{i=1}^n \|\mathbf{v}_i\|^2 \to E(v_{11}^2)$ a.s. Thus, from Lemma 3.2 again, we have that

$$tr\left(\mathbf{X}_n'\mathbf{X}_n + \frac{\mathbf{V}_n'\mathbf{V}_n}{N^{2/p}}\right) = \sum_{i=1}^n \frac{\|\mathbf{v}_i\|^2}{\|\mathbf{v}_i\|_p^2} + \frac{1}{N^{2/p}}\sum_{i=1}^n \|\mathbf{v}_i\|^2 = o\left(\frac{(\log N)^2}{N^{\frac{2}{p}-2}}\right) \quad \text{a.s.} \tag{3.27}$$

as $n \to \infty$. Note

$$\mathbf{X}_n - \frac{\mathbf{V}_n}{N^{1/p}} = \left(\left(1 - \frac{\|\mathbf{v}_1\|_p}{N^{1/p}}\right)\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|_p}, \ldots, \left(1 - \frac{\|\mathbf{v}_n\|_p}{N^{1/p}}\right)\frac{\mathbf{v}_n}{\|\mathbf{v}_n\|_p}\right).$$

It follows that

$$\begin{aligned} tr\left(\left(\mathbf{X}_n-\frac{\mathbf{V}_n}{N^{1/p}}\right)'\left(\mathbf{X}_n-\frac{\mathbf{V}_n}{N^{1/p}}\right)\right) &= \sum_{i=1}^{n}\left(1-\frac{\|\mathbf{v}_i\|_p}{N^{1/p}}\right)^2\frac{\|\mathbf{v}_i\|^2}{\|\mathbf{v}_i\|_p^2} \\ &\leqslant \max_{1\leqslant i\leqslant n}\left(1-\frac{\|\mathbf{v}_i\|_p}{N^{1/p}}\right)^2\cdot\sum_{i=1}^{n}\frac{\|\mathbf{v}_i\|^2}{\|\mathbf{v}_i\|_p^2} \\ &= o\left(\frac{(\log N)^4}{N^{\frac{2}{p}-1}}\right) \end{aligned}$$

by Lemma 3.2. This joint with (3.26) and (3.27) leads to

$$L^4\left(F^{b_n\mathbf{X}_n'\mathbf{X}_n}, F^{b_n\tilde{\mathbf{X}}_n'\tilde{\mathbf{X}}_n}\right) = O\left(\frac{2b_n^2}{n^2}\cdot\frac{(\log N)^4}{N^{\frac{2}{p}-1}}\cdot\frac{(\log N)^2}{N^{\frac{2}{p}-2}}\right) = O\left(\frac{(\log N)^6}{N}\right)\to 0 \quad \text{a.s.} \tag{3.28}$$

as $n\to\infty$.

Now let us look at the asymptotic distribution of $F^{b_n\tilde{\mathbf{X}}_n'\tilde{\mathbf{X}}_n}$. Observe that $b_n\tilde{\mathbf{X}}_n'\tilde{\mathbf{X}}_n = (n/N)^{2/p-1}(\mathbf{V}_n'\mathbf{V}_n/N)$ with $\mathbf{V}_n=(\mathbf{v}_1,\dots,\mathbf{v}_n)=(v_{ij})_{N\times n}$. Since $Ev_{ij}=0$, $h_\sigma := E(v_{ij}^2)=\frac{\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})}p^{2/p}$ and $(n/N)^{2/p-1}\to y^{2/p-1}$. By Theorem 3.6 from [3], with probability one, $F^{\mathbf{V}_n'\mathbf{V}_n/(Nh_\sigma)}$ converges weakly to F_y as $n\to\infty$. This implies that, with probability one, $F^{b_n\tilde{\mathbf{X}}_n'\tilde{\mathbf{X}}_n}$ converges weakly to $\mathcal{L}(c_y^{-1}T)$, the distribution of $c_y^{-1}T$, where T has law F_y. Equivalently, $L(F^{b_n\tilde{\mathbf{X}}_n'\tilde{\mathbf{X}}_n}, \mathcal{L}(c_y^{-1}T))\to 0$ a.s. This and (3.28) yield that $L(F^{n^{\frac{2}{p}-1}(\mathbf{X}_n'\mathbf{X}_n)}, \mathcal{L}(c_y^{-1}T))\to 0$ a.s. We then get the conclusion in the proposition. □

Proof of Theorem 3. By the Taylor expansion, since $f''(1)$ exists, there are constants $\delta\in(0,1)$ and $C>0$ such that

$$\left|f(x+1)-\left(f(1)+f'(1)x\right)\right|\leqslant Cx^2 \tag{3.29}$$

for all $|x|<\delta$. Set $E_n=\{\max_{1\leqslant i<j\leqslant n}|\frac{1}{a_N}\|\mathbf{x}_i-\mathbf{x}_j\|^2-1|<\delta\}$ for $n\geqslant 2$. Then, by (i) of Corollary 3, $E\sum_{n=2}^{\infty}I_{E_n^c}=\sum_{n=2}^{\infty}P(E_n^c)<\infty$, where $I_{E_n^c}$ is the indicator function of set E_n^c. This implies $\sum_{n=2}^{\infty}I_{E_n^c}<\infty$ a.s. Thus, $P(\Omega_1)=1$ where

$$\Omega_1 := \left\{\omega\text{: there exists } N=N(\omega) \text{ such that } \omega\in E_n \text{ for all } n\geqslant N\right\}. \tag{3.30}$$

Define

$$\begin{aligned} &\mathbf{Z}_n = f(0)\mathbf{I}_n+(z_{ij})_{n\times n}, \quad \text{where} \\ &z_{ii}=0 \quad\text{and}\quad z_{ij}=f(1)+f'(1)\left(a_N^{-1}\|\mathbf{x}_i-\mathbf{x}_j\|^2-1\right) \end{aligned} \tag{3.31}$$

for all $i\neq j$. Write $a_N^{-1}\|\mathbf{x}_i-\mathbf{x}_j\|^2=1+(a_N^{-1}\|\mathbf{x}_i-\mathbf{x}_j\|^2-1)$. Recall $\mathbf{M}_n=(m_{ij})=(f(a_N^{-1}\|\mathbf{x}_i-\mathbf{x}_j\|^2))_{n\times n}$. Take $x=a_N^{-1}\|\mathbf{x}_i-\mathbf{x}_j\|^2-1$ and plug into (3.29) to have

$$\left|m_{ij}-\left(f(1)+f'(1)\left(a_N^{-1}\|\mathbf{x}_i-\mathbf{x}_j\|^2-1\right)\right)\right|\leqslant C\left(a_N^{-1}\|\mathbf{x}_i-\mathbf{x}_j\|^2-1\right)^2$$

on E_n for all $i\neq j$. Since $a_N^{-1}\|\mathbf{x}_i-\mathbf{x}_j\|^2-1=(a_N^{-1}\|\mathbf{x}_i\|^2-1/2)+(a_N^{-1}\|\mathbf{x}_j\|^2-1/2)-2a_N^{-1}\mathbf{x}_i'\mathbf{x}_j$. Applying the convex inequality on function $h(x)=x^4$ we have

$$(m_{ij}-z_{ij})^2\leqslant KC^2\left(\left(\frac{\|\mathbf{x}_i\|^2}{a_N}-\frac{1}{2}\right)^4+\left(\frac{\|\mathbf{x}_j\|^2}{a_N}-\frac{1}{2}\right)^4+\left(a_N^{-1}\mathbf{x}_i'\mathbf{x}_j\right)^4\right)$$

for all $1\leqslant i<j\leqslant n$, where K is a universal constant. Recalling the definition of $\mathbf{Z}_n$ in (3.31) and noticing $tr((\mathbf{M}_n-\mathbf{Z}_n)^2)=\sum_{i\neq j}(m_{ij}-z_{ij})^2$, by the above inequality and Lemma 2.3 from [2] we have

$$L^3\big(F^{\mathbf{M}_n}, F^{\mathbf{Z}_n}\big) \leqslant \frac{1}{n} tr\big((\mathbf{M}_n - \mathbf{Z}_n)^2\big)$$
$$\leqslant (2KC^2)\sum_{i=1}^{n}\bigg(\frac{\|\mathbf{x}_i\|^2}{a_N} - \frac{1}{2}\bigg)^4 + \frac{2KC^2}{n}\sum_{1\leqslant i<j\leqslant n}\big(a_N^{-1}\mathbf{x}_i'\mathbf{x}_j\big)^4 \tag{3.32}$$
$$\to 0 \quad \text{a.s.}$$

by (ii) of Corollary 3. Denote $\Omega_2 = \{\lim_{n\to\infty}$ the term in (3.32) $= 0\}$. We then know $P(\Omega_2) = 1$. Thus, $\lim_{n\to\infty} L(F^{\mathbf{M}_n}, F^{\mathbf{Z}_n}) \to 0$ on $\Omega_1 \cap \Omega_2$. Since $P(\Omega_1 \cap \Omega_2) = 1$, we obtain

$$L\big(F^{\mathbf{M}_n}, F^{\mathbf{Z}_n}\big) \to 0 \quad \text{a.s.} \tag{3.33}$$

as $n \to \infty$. Recall z_{ij} in (3.31) with $z_{ii} = 0$ for all i. Since $f(1) + f'(1)(a_N^{-1}\|\mathbf{x}_i - \mathbf{x}_j\|^2 - 1) = f(1) - f'(1)$ for $i = j$, we have

$$\mathbf{Z}_n = f(0)\mathbf{I}_n + f(1)\mathbf{e}\mathbf{e}' + f'(1)\bigg(\frac{1}{a_N}\|\mathbf{x}_i - \mathbf{x}_j\|^2 - 1\bigg)_{n\times n} - \big(f(1) - f'(1)\big)\mathbf{I}_n$$
$$= \big(f(1) - f'(1)\big)\mathbf{e}\mathbf{e}' + \frac{f'(1)}{a_N}\big(\|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2\big)_{n\times n} + a\mathbf{I}_n + \frac{\kappa}{a_N}\big(\mathbf{x}_i'\mathbf{x}_j\big)_{n\times n} \tag{3.34}$$

where

$$a = f(0) - f(1) + f'(1) \quad \text{and} \quad \kappa = -2f'(1), \tag{3.35}$$

$\mathbf{e} = (1, \ldots, 1)' \in \mathbb{R}^n$ and the identity $\|\mathbf{x}_i - \mathbf{x}_j\|^2 = \|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2 - 2\mathbf{x}_i'\mathbf{x}_j$ is used in the last step. Note that

$$\frac{\kappa}{a_N}\big(\mathbf{x}_i'\mathbf{x}_j\big)_{n\times n} = \frac{\kappa}{2}\bigg(\frac{Ny}{n}\bigg)^{\frac{2}{p}-1}\cdot\big(c_y n^{\frac{2}{p}-1}\mathbf{X}_n'\mathbf{X}_n\big)$$

where $\mathbf{X}_n'\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)'(\mathbf{x}_1, \ldots, \mathbf{x}_n) = (\mathbf{x}_i'\mathbf{x}_j)_{n\times n}$ and c_y is as in (3.25). Let $\mathbf{U}_n = a\mathbf{I}_n + \frac{\kappa}{a_N}(\mathbf{x}_i'\mathbf{x}_j)_{n\times n}$. By Proposition 1, with probability one,

$$F^{\mathbf{U}_n} \text{ converges weakly to the distribution of } a + bV, \tag{3.36}$$

where V has the law F_y as in (1.4) and $b = -f'(1)$. It is easy to see that the rank of $\mathbf{e}\mathbf{e}'$ is equal to 1 and

$$\operatorname{rank}\big(\|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2\big)_{n\times n} \leqslant \operatorname{rank}\big(\|\mathbf{x}_i\|^2\big)_{n\times n} + \operatorname{rank}\big(\|\mathbf{x}_j\|^2\big)_{n\times n} \leqslant 2. \tag{3.37}$$

By Lemma 2.2 from [2] and (3.34), $L(F^{\mathbf{Z}_n}, F^{\mathbf{U}_n}) \to 0$ a.s. as $n \to \infty$. This and (3.36) conclude that, with probability one, $F^{\mathbf{Z}_n}$ converges weakly to the distribution of $a + bV$, which and (3.33) imply the desired conclusion. □

Proof of Corollary 1. Notice $\mathbf{B}_n = (\|\mathbf{x}_i - \mathbf{x}_j\|^\alpha)_{n\times n}$. Write

$$N^{(\frac{2}{p}-1)\frac{\alpha}{2}}\mathbf{B}_n = \bigg(\bigg(\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{a_N}\bigg)^{\alpha/2}\bigg)\cdot N^{(\frac{2}{p}-1)\frac{\alpha}{2}}(a_N)^{\frac{\alpha}{2}}. \tag{3.38}$$

It is easily seen from (1.14) that

$$N^{(\frac{2}{p}-1)\frac{\alpha}{2}}(a_N)^{\frac{\alpha}{2}} = \bigg(2p^{\frac{2}{p}}\frac{\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})}\bigg)^{\alpha/2}. \tag{3.39}$$

Now, take $f(x) = x^{\alpha/2}$. Then, $f(0) = 0$, $f(1) = 1$ and $f'(1) = \alpha/2$. Review Theorem 3. Then $a = (\alpha/2) - 1$ and $b = -\alpha/2$. We obtain from the theorem that, with probability one, the empirical spectral distribution of $((\|\mathbf{x}_i - \mathbf{x}_j\|^2/a_N)^{\alpha/2})$ converges to the law of $c + dV$ where V has the law F_y as in (1.4). This joint with (3.38) and (3.39) gives the conclusion. □

Proof of Corollary 2. Let a_N be as in (1.14). When $p=2$, it is easy to see that $a_N=2$. Let $\theta_{ij}\in[0,\pi]$ be the angle between vectors $\overrightarrow{\mathbf{Ox}_i}$ and $\overrightarrow{\mathbf{Ox}_j}$ for any $1\leqslant i,j\leqslant n$, where $\mathbf{O}$ is the origin. Then $d(\mathbf{x}_i,\mathbf{x}_j)=\theta_{ij}$. From the fact that $\cos\theta_{ij}=\mathbf{x}_i'\mathbf{x}_j$, we know

$$d(\mathbf{x}_i,\mathbf{x}_j)=\cos^{-1}\big(\mathbf{x}_i'\mathbf{x}_j\big)=\cos^{-1}\left(1-\frac{\|\mathbf{x}_i-\mathbf{x}_j\|^2}{2}\right).$$

Take $f(x)=\cos^{-1}(1-x)$ for $x\in[0,1]$. It is easy to check that

$$f'(x)=\frac{1}{\sqrt{2x-x^2}}\quad\text{and}\quad f''(x)=\frac{x-1}{(2x-x^2)^{3/2}}$$

for $x\in(0,1)$. Easily, $f(0)=0$, $f(1)=\pi/2$ and $f'(1)=1$. Thus,

$$a=f(0)-f(1)+f'(1)=1-\frac{\pi}{2}\quad\text{and}\quad b=-f'(1)=-1.$$

Then the conclusion follows from Theorem 3. □

4. Verifications of statements

In this section, we will verify some claims and conclusions appeared in Sections 1 and 2.

4.1. Verifications

Verification of (1.5). For $f(x)=\sqrt{x}$, it is easy to check that

$$f^{(m)}(x)=(-1)^{m-1}\frac{1\cdot 3\cdots(2m-3)}{2^m}x^{-(2m-1)/2}$$

for all $m\geqslant 2$ and $x>0$. Write $1\cdot 3\cdots(2m-3)=(2m-2)!/(2\cdot 4\cdots(2m-2))=(2m-2)!/(2^{m-1}(m-1)!)$. Then, for any $t>0$,

$$w_m(t)\geqslant\left|f^{(m)}(t)\right|=\frac{(2m-2)!}{(m-1)!}\cdot\frac{t^{-(2m-1)/2}}{2^{2m-1}}.$$

By the Stirling formula, $m!=\sqrt{2\pi m}m^me^{-m}(1+o(1))$ as $m\to\infty$. Then, for any $t>0$,

$$\begin{aligned}&\liminf_{m\to\infty}\frac{1}{m\log m}\log\omega_m(t)\\ &\quad\geqslant\liminf_{m\to\infty}\frac{(2m-2)\log(2m-2)-(m-1)\log(m-1)}{m\log m}=1.\end{aligned}\qquad\square$$

Verification of Example 1. Let $f(x)=(\sin\sqrt{x})/\sqrt{x}$ for $x\neq 0$ and $f(0)=1$. It is easy to see from the Taylor expansion that $f(x)=\sum_{m=0}^{\infty}\frac{(-1)^m}{(2m+1)!}x^m$ for all $x\in\mathbb{R}$. Thus,

$$f^{(n)}(x)=\sum_{m=n}^{\infty}\frac{(-1)^m m!}{(2m+1)!}\frac{x^{m-n}}{(m-n)!}.$$

Since $|\frac{(-1)^m m!}{(2m+1)!}|\leqslant 1$, we have $|f^{(n)}(x)|\leqslant e^{|x|}$ for all $x\in\mathbb{R}$. So $\log\omega_n(4a^2)=o(n\log n)$ as $n\to\infty$ holds for any $a>0$. Thus, Theorem 2 is true.

Now $f(0)=1$ and

$$f'(x)=\frac{\sqrt{x}\cos\sqrt{x}-\sin\sqrt{x}}{2x^{3/2}}.$$

Thus, Theorem 3 holds with

$$a = f(0) - f(1) + f'(1) = 1 + \frac{\cos 1 - 3\sin 1}{2},$$
$$b = -f'(1) = \frac{\sin 1 - \cos 1}{2}.$$

Now, assume $p = 2$. By using the formula $\Gamma(x+1) = x\Gamma(x)$ for $x > 0$ we see that $a_N = 2$. Let $g(x) = f(2x)$ for $x \geqslant 0$. Then

$$\mathbf{M}_n = \big(f\big(\|\mathbf{x}_i - \mathbf{x}_j\|^2\big)\big)_{n\times n} = \left(g\left(\frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_j\|^2\right)\right)_{n\times n}.$$

By Theorem 3, $\hat{\mu}(\mathbf{M}_n)$ converges weakly to $c_1 + d_1 V$ where V has distribution F_y as in (1.4), and

$$c_1 = g(0) - g(1) + g'(1) = f(0) - f(2) + 2f'(2) = 1 + \frac{\sqrt{2}\cos\sqrt{2} - 3\sin\sqrt{2}}{2\sqrt{2}},$$
$$d_1 = -g'(1) = -2f'(2) = \frac{\sin\sqrt{2} - \sqrt{2}\cos\sqrt{2}}{2\sqrt{2}}. \qquad \square$$

Verification of Example 2. Now $f(x) = (2\pi)^{-3/2}e^{-x/2}$ for $x \geqslant 0$. Recall the notation in Theorem 2. Obviously, $\omega_m(t) = (2\pi)^{-3/2}2^{-m}$ for all $t \geqslant 0$ and $m \geqslant 1$. Thus, $\log \omega_m = o(m \log m)$ and $\log \omega_m(4a^2) = o(m \log m)$ as $m \to \infty$. So Theorems 1 and 2 holds. It is easy to see that Theorem 3 holds with

$$a = f(0) - f(1) + f'(1) = (2\pi)^{-3/2}\left(1 - \frac{3}{2}e^{-1/2}\right) \quad \text{and}$$
$$b = -f'(1) = \frac{1}{2}(2\pi)^{-3/2}e^{-1/2}. \qquad \square$$

Verification of Example 3. In this example, $p = 2$. It follows from (1.14) that $a_N = 2$. Reviewing the proof of Corollary 2, we know

$$d(\mathbf{x}_i, \mathbf{x}_j) = \cos^{-1}\left(1 - \frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2}\right).$$

Let $g(x) = \cos^{-1}(1 - x)$ for $x \in [0, 1]$. From the proof of Corollary 2,

$$g'(x) = \frac{1}{\sqrt{2x - x^2}} \quad \text{and} \quad g''(x) = \frac{x - 1}{(2x - x^2)^{3/2}}$$

for $x \in (0, 1)$. Easily, $g(0) = 0$, $g(1) = \pi/2$ and $g'(1) = 1$.

(i) Set $f(x) = g(x)^\gamma$ for $x \in [0, 1]$, where $\gamma \in (0, 1]$. According to the notation $\mathbf{M}_n$ in (1.14), we have $\mathbf{M}_n = (d(\mathbf{x}_i, \mathbf{x}_j)^\gamma)_{n\times n}$. Trivially, $f'(x) = \gamma g(x)^{\gamma-1}g'(x)$. So $f(0) = g(0)^\gamma = 0$, $f(1) = g(1)^\gamma = (\pi/2)^\gamma$ and $f'(1) = \gamma(\pi/2)^{\gamma-1}$. Thus, Theorem 3 holds for $\gamma \in (0, 1]$ with

$$a = f(0) - f(1) + f'(1) = \frac{2\gamma - \pi}{2}\left(\frac{\pi}{2}\right)^{\gamma-1} < 0 \quad \text{and}$$
$$b = -f'(1) = -\gamma\left(\frac{\pi}{2}\right)^{\gamma-1}.$$

(ii) Take $f(x) = e^{-\lambda^2 g(x)^\gamma}$ for $x \in [0, 1]$, where $\gamma \in (0, 1]$. According to the notation $\mathbf{M}_n$ in (1.14), we see that $\mathbf{M}_n = (\exp(-\lambda^2 d(\mathbf{x}_i, \mathbf{x}_j)^\gamma))_{n\times n}$. Now, $f'(x) = -\gamma\lambda^2 g(x)^{\gamma-1}g'(x)e^{-\lambda^2 g(x)^\gamma}$ so that

$$f(0) = 1, \qquad f(1) = e^{-\lambda^2(\pi/2)^\gamma} \quad \text{and} \quad f'(1) = -\gamma\lambda^2\left(\frac{\pi}{2}\right)^{\gamma-1}e^{-\lambda^2(\pi/2)^\gamma}.$$

Hence, Theorem 3 holds with

$$a = f(0) - f(1) + f'(1) = 1 - e^{-\lambda^2(\pi/2)^\gamma} - \gamma\lambda^2\left(\frac{\pi}{2}\right)^{\gamma-1} e^{-\lambda^2(\pi/2)^\gamma};$$
$$b = -f'(1) = \gamma\lambda^2\left(\frac{\pi}{2}\right)^{\gamma-1} e^{-\lambda^2(\pi/2)^\gamma}.$$

Observe that $e^x > 1 + x \geqslant 1 + tx$ for all $x > 0$ and $t \leqslant 1$. Then

$$a = e^{-\lambda^2(\pi/2)^\gamma}\left[e^{\lambda^2(\pi/2)^\gamma} - 1 - \frac{2\gamma}{\pi}\cdot\lambda^2\left(\frac{\pi}{2}\right)^{\gamma}\right] > 0. \tag{4.1}$$

(iii) Now, given $0 < \gamma \leqslant 2$, let $f(x) = e^{-\lambda^2(2x)^{\gamma/2}}$ for $x \geqslant 0$. Then, by the definition of $\mathbf{M}_n$ in (1.14), we get that $\mathbf{M}_n = (\exp(-\lambda^2\|\mathbf{x}_i - \mathbf{x}_j\|^\gamma))_{n\times n}$. Note that $f'(x) = -\gamma\lambda^2(2x)^{\gamma/2-1}e^{-\lambda^2(2x)^{\gamma/2}}$ for $x \geqslant 0$. Thus,

$$f(0) = 1, \qquad f(1) = e^{-\lambda^2 2^{\gamma/2}} \quad \text{and} \quad f'(1) = -\gamma\lambda^2 2^{\gamma/2-1}e^{-\lambda^2 2^{\gamma/2}}.$$

Then Theorem 3 holds with

$$a = f(0) - f(1) + f'(1) = 1 - e^{-\lambda^2 2^{\gamma/2}} - \frac{\gamma}{2}\left(\lambda^2 2^{\gamma/2}\right)e^{-\lambda^2 2^{\gamma/2}};$$
$$b = -f'(1) = \frac{\gamma}{2}\left(\lambda^2 2^{\gamma/2}\right)e^{-\lambda^2 2^{\gamma/2}} > 0.$$

By the same argument as in (4.1), we know $a > 0$ for all $0 < \gamma \leqslant 2$. □

Verification of Example 4. Review $\mathbf{M}_n = (f(\|\mathbf{x}_i - \mathbf{x}_j\|^2))_{n\times n}$ where $f(x) = x^m + \alpha_1 x^{m-1} + \cdots + \alpha_m$ for $m \geqslant 1$. Let $\mathbf{B}_{n,k} = (\|\mathbf{x}_i - \mathbf{x}_j\|^{2k})_{n\times n}$ with $1 \leqslant k \leqslant m$ and $\mathbf{B}_{n,0}$ be the matrix whose entries are all equal to 1. Then

$$\begin{aligned}\mathbf{U}_1 &:= N^{(\frac{2}{p}-1)m}\left(f\left(\|\mathbf{x}_i - \mathbf{x}_j\|^2\right)\right)_{n\times n}\\ &= \underbrace{N^{(\frac{2}{p}-1)m}\left(\|\mathbf{x}_i - \mathbf{x}_j\|^{2m}\right)_{n\times n}}_{\mathbf{U}_2} + \underbrace{\sum_{k=0}^{m-1} N^{(\frac{2}{p}-1)m}(\alpha_{m-k}\mathbf{B}_{n,k})}_{\mathbf{U}_3}.\end{aligned} \tag{4.2}$$

From Corollary 1, with probability one, $F^{\mathbf{U}_2}$ converges weakly to the distribution of $c + dV$ where V has the distribution F_y as in (1.4),

$$c = (m-1)\left(2p^{\frac{2}{p}}\frac{\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})}\right)^m \quad \text{and} \quad d = -m\left(2p^{\frac{2}{p}}\frac{\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})}\right)^m.$$

Second, by Lemma 2.3 from [2] we have

$$L^3\left(F^{\mathbf{U}_1}, F^{\mathbf{U}_2}\right) \leqslant \frac{1}{n}tr\left(\mathbf{U}_3^2\right). \tag{4.3}$$

Recall the Frobenius norm $\|\mathbf{E}\|_F = (tr(\mathbf{E}^2))^{1/2} = (\sum_{1\leqslant i,j\leqslant n} e_{ij}^2)^{1/2}$ for any symmetric matrix $\mathbf{E} = (e_{ij})_{n\times n}$. Observe that $p > 4m/(2m-1) > 2$ for all $m \geqslant 1$. Then, $\|\mathbf{v}\| = \|\mathbf{v}\|_2 \leqslant \|\mathbf{v}\|_p \leqslant 1$ for all $\mathbf{v} \in B_{N,p}$. It follows that $\|\mathbf{x} - \mathbf{y}\| \leqslant 2$ for all $\mathbf{x}, \mathbf{y} \in B_{N,p}$. This says

$$K := \sup_{\mathbf{x},\mathbf{y}\in B_{N,p}} \max_{0\leqslant k\leqslant m-1}\left(1 + \alpha_k + \|\mathbf{x} - \mathbf{y}\|^{2k}\right) < \infty.$$

Thus, $tr((\alpha_{m-k}\mathbf{B}_{n,k})^2) \leqslant n^2K^4$ for each $0 \leqslant k \leqslant m-1$. By the triangular inequality, $(tr(\mathbf{U}_3^2))^{1/2} = \|\mathbf{U}_3\|_F \leqslant (mK^2)N^{(\frac{2}{p}-1)m}n$. This and (4.3) imply

$$L^3\big(F^{\mathbf{U}_1}, F^{\mathbf{U}_2}\big) \leqslant \big(mK^2\big)^2 N^{2(\frac{2}{p}-1)m} n \to 0$$

as $n \to \infty$ since $n/N \to y \in (0,\infty)$ and $p > 4m/(2m-1)$ for all $m \geqslant 1$. By the convergence of $F^{\mathbf{U}_2}$, we know that, with probability one, $F^{\mathbf{U}_1}$, and hence $\hat{\mu}(N^{(\frac{2}{p}-1)m}\mathbf{M}_n)$, converges weakly to the distribution of $c + dV$. $\square$

References

[1] G. Aubrun, Random points in the unit ball of l_p^n, Positivity 10 (2006) 755–759.
[2] Z.D. Bai, Methodologies in spectral analysis of large dimensional random matrices, a review, Statist. Sinica 9 (1999) 611–677.
[3] Z.D. Bai, J.W. Silverstein, Spectral Analysis of Large Dimensional Random Matrices, second edition, Springer, 2009.
[4] F. Barthe, F. Gamboa, L. Lozada-Chang, A. Rouault, Generalized Dirichlet distributions on the ball and moments, ALEA Lat. Am. J. Probab. Math. Stat. 7 (2010) 319–340.
[5] E. Bogomolny, O. Bohigas, C. Schmidt, Distance matrices and isometric embeddings, J. Math. Phys. Anal. Geom. 4 (1) (2008) 7–23.
[6] E. Bogomolny, O. Bohigas, C. Schmidt, Spectral properties of distance matrices, J. Phys. A 36 (2003) 3595–3616.
[7] C. Bordenave, Eigenvalues of Euclidean random matrices, Random Structures Algorithms 33 (4) (2008) 515–532.
[8] W. Bryc, A. Dembo, T. Jiang, Spectral measure of large random Hankel, Markov and Toeplitz matrices, Ann. Probab. 34 (1) (2006) 1–38.
[9] T. Cai, T. Jiang, Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices, Ann. Statist. 39 (3) (2011) 1496–1525.
[10] A. Cavagna, I. Giardina, G. Parisi, An investigation of the hidden structure of states in a mean-field spin-glass model, J. Phys. A 30 (20) (1997) 7021–7038.
[11] X. Cheng, A. Singer, The spectrum of random inner-product kernel matrices, arXiv:1202.3155, available at: http://arxiv.org/pdf/1202.3155v2.pdf, 2012.
[12] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, second edition, Springer, 1998.
[13] P. Diaconis, S. Goel, S. Holmes, Horseshoes in multidimensional scaling and local kernel methods, Ann. Appl. Stat. 2 (3) (2008) 777–807.
[14] Y. Do, V. Vu, The spectrum of random kernel matrices, arXiv:1206.3763, available at: http://arxiv.org/abs/1206.3763, 2012.
[15] R. Durrett, Probability: Theory and Examples, second edition, The Duxbury Press, 1995.
[16] N.E. El Karoui, The spectrum of kernel random matrices, Ann. Statist. 38 (1) (2010) 1–50.
[17] J. Felsenstein, Inferring Phylogenies, second edition, Sinauer Associates, Sunderland, MA, 2003.
[18] T.W. Gamelin, Complex Analysis, first edition, Springer, 2001.
[19] R.J. Gardner, The Brunn–Minkowski inequality, Bull. Amer. Math. Soc. 39 (3) (2002) 355–405.
[20] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1994.
[21] V. Koltchinskii, E. Giné, Random matrix approximation of spectra of integral operators, Bernoulli 6 (1) (2000) 113–167.
[22] M. Mézard, G. Parisi, A. Zee, Spectra of euclidean random matrices, Nuclear Phys. B 559 (1999) 689–701.
[23] D.M. Mount, Bioinformatics: Sequence and Genome Analysis, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2004.
[24] A. Naor, The surface measure and cone measure on the sphere of l_p^n, Trans. Amer. Math. Soc. 359 (2007) 1045–1079.
[25] A. Naor, D. Romik, Projecting the surface measure of the sphere of l_p^n, Ann. Inst. H. Poincaré Probab. Stat. 39 (2) (2003) 241–261.
[26] A. Pajor, L. Patur, On the limiting empirical measure of eigenvalues of the sum of rank one matrices with log-concave distribution, Studia Math. 195 (2009) 11–29.
[27] G. Parisi, Euclidean random matrices: solved and open problems, in: Applications of Random Matrices in Physics, in: NATO Sci. Ser., vol. 221, 2006, pp. 219–260.
[28] M. Penrose, Random Geometric Graphs, Oxford Stud. Probab., Oxford University Press, Oxford, 2003.
[29] V.M. Preciado, A. Jadbabaie, Spectral analysis of virus spreading in random geometric networks, in: IEEE Conference on Decision and Control, 2009.
[30] I. Rajapaksea, M. Groudineb, M. Mesbahi, Dynamics and control of state-dependent networks for probing genomic organization, Proc. Natl. Acad. Sci. USA 108 (42) (2011) 17257–17262.
[31] L. Reid, X. Sun, Distance matrices and ridge functions interpolation, Canad. J. Math. 45 (6) (1993) 1313–1323.
[32] I.J. Schoenberg, Metric spaces and completely monotone functions, Ann. of Math. 39 (1938) 811–841.
[33] I.J. Schoenberg, Metric spaces and positive definite functions, Trans. Amer. Math. Soc. 44 (1938) 522–536.
[34] I.J. Schoenberg, On certain metric spaces arising from Euclidean spaces by a change of metric and their imbedding in Hilbert space, Ann. of Math. 38 (1937) 787–793.
[35] F. Sinz, M. Bethge, L_p-nested symmetric distributions, J. Mach. Learn. Res. 11 (2010) 3409–3451.

[36] S.E. Skipetrov, A. Goetschy, Eigenvalue distributions of large Euclidean random matrices for waves in random media, J. Phys. A 44 (2011) 065102.
[37] A. Song, A.K. Gupta, L_p-norm uniform distribution, Proc. Amer. Math. Soc. 125 (1997) 595–601.
[38] A.M. Vershik, Random metric spaces and universality, Russian Math. Surveys 59 (2004) 259–295.
[39] T.M. Wun, R.F. Loring, Phonons in liquids: A random walk approach, J. Chem. Phys 97 (11) (1992) 8568–8575.