J. Indones. Math. Soc.
Vol. 31, No. 01 (2025), pp. 1–22.

Limiting Spectral Distributions of Random Matrices Having Equi-Correlated Normal Structure

Atina Husnaqilati[1,2*]

[1]**Department of Mathematics, Tohoku University, Japan.**
[2]**Department of Mathematics, Universitas Gadjah Mada, Indonesia.**
husqila@gmail.com

Abstract. By rank inequalities, we show that the limiting spectral distribution of random matrices, which are Fisher matrices and Beta matrices composed of two independent samples from independent p-dimensional, centered normal populations such that all entries have unit variance and any correlation coefficient between different variables are fixed nonnegative $r_1, r_2 < 1$. Moreover, by similar method, we also present the limiting spectral distribution of Wigner matrices, Toeplitz matrices, and Hankel matrices of order p, where all entries are standard normal random variables and mutually correlated with a fixed nonnegative $r < 1$. However, the rank inequality for empirical spectral distributions is unable to show the limiting spectral distributions of Markov matrices and banded Toeplitz matrices because the perturbation matrices of those matrices have a rate rank 1.

Key words and Phrases: random matrices, limiting spectral distribution, equi-correlated normal population, rank inequalities for empirical spectral distributions, Fisher matrices.

1. INTRODUCTION

Suppose $\mathbf{M}$ is a Hermitian matrix of order p and $\lambda_1(\mathbf{M}) \geq \lambda_2(\mathbf{M}) \geq \cdots \geq \lambda_p(\mathbf{M})$ are the eigenvalues of $\mathbf{M}$. The *empirical spectral distribution* (ESD) of $\mathbf{M}$ is, by definition, a function

$$F^{\mathbf{M}}(x) = \frac{1}{p}\sum_{i=1}^{p} \mathbf{1}_{\lambda_i(\mathbf{M}) \leq x}, \qquad (x \in \mathbb{R})$$

where $\mathbf{1}$ is an indicator function. If $F^{\mathbf{M}}$ has a limit deterministic distribution function in an asymptotic framework for $p \to \infty$, it is referred to as the *limiting spectral distribution* (LSD) of $\mathbf{M}$. The LSDs of large-dimensional random matrices can be

*Corresponding author

2020 Mathematics Subject Classification: 60F15; 62H25
Received: 05-05-2022, accepted: 10-11-2024.

used to determine the majority of their bulk spectrum limiting properties. Thus, the LSDs of large dimensional random matrices have garnered considerable interest among mathematicians, physicists, and statisticians, e.g., Wigner [1, 2] for LSDs of *random Wigner matrices*, Yao et al. [3, p. 12] for the LSDs of *sample covariance matrices*, Bai et al. [4] for the LSDs of the product of two independent random matrices, Bryc et al. [5] for the LSDs of *random symmetric Toeplitz matrices*, Kargin [6] for the LSDs of *banded random Toeplitz matrices*, and Paul-Aue [7] for a review of application from random matrices to inference statistics.

Among the aforementioned works [4, 8, 9, 1, 2, 10], the assumption that the entries of the random matrices have independent random variables is usually required. However, in practical application, the independent variables assumption is a strong condition. In this paper, we consider a case in which variables are fully dependent. To allow for correlations among variables in massive models based on the sample covariance matrices and the sample correlation matrices, e.g., financial mathematics [11, 12, 13] and psychometrics [14], we concentrate on random matrices derived from an *equi-correlated normal population* (ENP). By the equi-correlation structure, we mean the population correlation matrix

$$\mathbf{C}(r) = \begin{bmatrix} 1 & r & \cdots & r \\ r & 1 & \cdots & r \\ \vdots & \vdots & & \vdots \\ r & r & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{p\times p}.$$

This matrix is clearly unbounded spectral norm because the largest eigenvalues is $1 + (p-1)r$. Hereafter, we assume $0 \le r < 1$.

The LSD of a sample covariance matrix from ENP is Marčhenko-Pastur distribution scaled by $1 - r$ [15]. This has two different proofs based on different aspects of ENP. One proof observes that the LSD of $\mathbf{C}(r)$ is that of $(1-r)\mathbf{I}$ where $\mathbf{I}$ denotes identity matrix of order p, and applies [16, Theorem 1.1] for Stieltjes transform. The other proof [15] notes the decomposition

$$N_p(0, \mathbf{C}(r)) = \sqrt{r}N_1(0,1)[1, ..., 1]^\top + \sqrt{1-r}N_p(0, \mathbf{I}), \tag{1}$$

and applies the *rank inequality* of ESD:

Proposition 1.1 ([17, Lemma 2.2, the rank inequality])**.**

$$\mathrm{K}\left(F^{\mathbf{M}_1}, F^{\mathbf{M}_2}\right) \le \frac{1}{p}\ \mathrm{rank}(\mathbf{M}_1 - \mathbf{M}_2), \qquad (\mathbf{M}_1, \mathbf{M}_2 \in \mathbb{C}^{p\times p}).$$

This rank inequality holds by the *interlacing theorem* [18, p. 242]. The first proof by [16, Theorem 1.1] and Stieltjes transform is sophisticated, then developed later by Bai-Zhou [19], Hui-Pan [20], and Bryson et al. [21], under the boundedness condition of the spectral norm of random matrices. Therefore, they are no longer applicable to ENP as the spectral norm $1 + (p-1)r$ of $\mathbf{C}(r)$ is unbounded. On the other hand, the second proof is simple because it does not use Stieltjes transform and then only depends on the rank of first term on the right-hand side of (1) (*the rank of perturbation matrix*).

The motivation of this paper is (a) to establish the LSDs of more various complicated random matrices from ENP than sample covariance matrices by extending the rank inequalities for ESDs and the decomposition (1); (b) to examine the realm of rank inequality for ESDs in order to demonstrate the LSDs of various random matrices; (c) to show the application of LSDs in real datasets. According to [22], the decomposition (1) is useful to show the characteristic of a random matrix which the first term of the decomposition (1) is the low rank perturbations of a random matrix. Moreover, by [22], the LSD of a random matrix is shown to depend explicitly on LSD of the unperturbed random matrix. Furthermore, according to [23], the decomposition (1) can be seen as a one factor-model. By this, Akama [23] showed the estimator of r in equi-correlation normal population.

The rank inequality for ESDs are convenient technique in the cases where the underlying variables are not independent and identically distributed (i.i.d.) [24, p. 503]. We apply a new rank inequality for ESDs to show the LSDs of some more complicated random matrices than the sample covariance matrices from ENP. In particular, the random matrices used are *Fisher matrices* [3, p. 25] and *Beta matrices* [25]. These matrices have key role in likelihood ratio test for verifying the equality of two covariance matrices [3, 25, p. 151].

We also consider the others large-dimensional symmetric matrices being *random Wigner matrices* [1, 2], *random Toeplitz matrices*, *random Hankel matrices*, *random Markov matrices*, and *banded random Toeplitz matrices* [6]. Those applications assume that all entries of random Wigner matrices, random Toeplitz matrices, random Hankel matrices, random Markov matrices, and banded random Toeplitz matrices are i.i.d. random variables. Moreover, the LSDs of random Toeplitz matrices, random Hankel matrices, and random Markov matrices with independent entries are included in the list of unsolved random matrix problems from [17, Section 6], and then Bryc et al. [5] provided the answer of that problem by the moment method. However, in this paper, we assume that all pairings entries from those matrices are equi-correlated and standard normal random variables, and then we find their LSDs by rank inequality for ESDs. By these, we are able to understand the realm of rank inequality for ESDs to show the LSDs of various random matrices.

Herein, we present our results. Given two independent samples $\widetilde{X}_1^{(i)}, \ldots, \widetilde{X}_{n_i}^{(i)} \overset{\text{i.i.d.}}{\sim} \mathrm{N}_p(\mathbf{0}, \mathbf{C}(r_i))$ with $0 \leq r_i < 1$ $(i = 1, 2)$. Let $\widetilde{\mathbf{X}}^{(i)} = \left[\widetilde{X}_1^{(i)}, \ldots, \widetilde{X}_{n_i}^{(i)}\right] \in \mathbb{R}^{p \times n_i}$ and $\widetilde{\mathbf{S}}_i = n_i^{-1} \widetilde{\mathbf{X}}^{(i)} \left(\widetilde{\mathbf{X}}^{(i)}\right)^\top$. By definition, a *Fisher matrix* is $\widetilde{\mathbf{F}} = \widetilde{\mathbf{S}}_1 \widetilde{\mathbf{S}}_2^{-1}$, meanwhile a *Beta matrix* of scale parameter $\alpha > 0$ is $\widetilde{\mathbf{B}} = \widetilde{\mathbf{S}}_2(\widetilde{\mathbf{S}}_2 + \alpha \widetilde{\mathbf{S}}_1)^{-1}$. In a suitable limiting regime $p/n_i \to c_i > 0$ $(i = 1, 2)$, the following results obtained by our new rank inequality and decomposition (1): The LSD of $\mathbf{F}$ ($\mathbf{B}$, resp.), which is a distribution $\mathrm{F}_{c_1,c_2}((1 - r_2)x/(1 - r_1))$ ($\mathrm{BM}_{\alpha(1-r_2)/(1-r_1),c_1,c_2}(x)$, resp.) for a fixed deterministic distribution function $\mathrm{F}_{c_1,c_2}(x)$ ($\mathrm{BM}_{\alpha,c_1,c_2}(x)$, resp.).

Likewise, for random Wigner matrices, random Toeplitz matrices, random Hankel matrices, and banded random Toeplitz matrices of order p, when all pairings of variables are equi-correlated by a fixed nonnegative $r < 1$ and all entries are standard normal random variables, the following results hold by the rank inequality (Proposition 1.1): (1) The LSDs of random Wigner matrices are semicircle law scaled by $\sqrt{1-r}$, (2) the LSDs of random Toeplitz matrices and random Hankel matrices are deterministic distribution functions scaled by $\sqrt{1-r}$. However, the LSDs of random Markov matrices and banded random Toeplitz matrices are undecided (see subsections 4.4 and 4.5).

This paper is organized as follows: Section 2 introduces a new rank inequality for the ESDs of the product of a matrix and the inverse of another matrix, to show the LSDs of Fisher matrices and Beta matrices, after the applications of Proposition 1.1 for several random matrices in [24] are described. Section 3 examines the LSDs of Fisher matrices and Beta matrices by a new rank inequality for ESDs in a population $N_p(0, \mathbf{C}(r_i))(i = 1, 2)$. Section 4 establishes the LSDs of random Wigner matrices, random Toeplitz matrices, and random Hankel matrices, assuming all pairings of variables are equi-correlated standard normal random variables by the rank inequality for ESDs (Proposition 1.1). In this section, we also show that the LSDs of random Markov matrices and banded random Toeplitz matrices are still undecided by the rank inequality for ESDs (Proposition 1.1).

2. RANK INEQUALITIES OF ESDs

Fact 2.1. *(1) For square $\mathbf{M}_1$, $\mathbf{M}_2$ of the same size,*

$$\operatorname{rank}(\mathbf{M}_1 + \mathbf{M}_2) \leq \operatorname{rank}(\mathbf{M}_1) + \operatorname{rank}(\mathbf{M}_2).$$

(2) For square $\mathbf{M}_1$, $\mathbf{M}_2$ in which $\mathbf{M}_1\mathbf{M}_2$ is defined

$$\operatorname{rank}(\mathbf{M}_1\mathbf{M}_2) \leq \min(\operatorname{rank}(\mathbf{M}_1), \operatorname{rank}(\mathbf{M}_2)).$$

(3) For matrices $\mathbf{M}_1$,

$$\operatorname{rank}({\mathbf{M}_1}^\top) = \operatorname{rank}(\mathbf{M}_1).$$

The following propositions can be found in Huber [26].

Proposition 2.2 ([26, Lemma 2.9])**.** *The Lévy distance metrizes the weak topology of the set of distribution functions.*

Proposition 2.3 ([26, p. 36])**.** *For any distribution functions F and G,*

$$\mathrm{L}(F, G) \leq \mathrm{K}(F, G)$$

where L *is the* Lévy distance *between two distribution functions F and G, see* [26, Definition 2.7]. *Also,* K *is the* Kolmogorov distance *between two distribution functions F and G is defined as*

$$\mathrm{K}(F, G) = \sup_{x \in \mathbb{R}} |F(x) - G(x)|.$$

The rank inequality for ESDs (Proposition 1.1) is applied to show almost surely the ESD of a Wigner matrix with centered i.i.d. entries weakly converges to the ESD of the same Wigner matrix from the certain truncated entries. By this and the moment method, Bai-Silverstein [24, p. 27] shows that almost surely the ESD of Wigner matrix from the certain truncated entries converges weakly to the *semicircle law*. On the other hand, we will apply the rank inequality for ESDs (Proposition 1.1) to show that almost surely the ESDs of certain random matrices with correlated pairing standard normal entries by a fixed nonnegative $r < 1$ weakly converges to the ESDs of random matrices with centered i.i.d. entries multiplied by $\sqrt{1-r}$.

We also consider the product of a random matrix and the inverse of another random matrix. Let $\mathbf{Y} = [y_{ij}]_{p\times n}$ such that y_{ij} are centered i.i.d. random variables. By the rank inequality for ESDs (Proposition 1.1), Bai-Silverstein [24, p. 70] showed that almost surely the ESD of $n^{-1}\mathbf{Y}\mathbf{Y}^\top$ converges weakly to the ESD of $n^{-1}\widetilde{\mathbf{Y}}\widetilde{\mathbf{Y}}^\top$ where $\widetilde{\mathbf{Y}}$ is the matrix $\mathbf{Y}$ with certain truncated entries. By this and the moment method, Bai-Silverstein [24, p. 71] proved that almost surely the ESD of $n^{-1}\widetilde{\mathbf{Y}}\widetilde{\mathbf{Y}^\top}\mathbf{M}$ tends to a nonrandom limit in almost surely. Here, $\mathbf{M}$ is a symmetric matrix independent of $\mathbf{Y}$, and the ESD of $\mathbf{M}$ converges to a deterministic probability distribution. In contrast, we will apply a new rank inequality for ESDs to show the LSD of the product of a random matrix and the inverse of another random matrix by extended Proposition 1.1 in the following Lemma 2.5.

Remark 2.4. It is well-known that for any square matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ of the same order, the two products $\mathbf{M}_1\mathbf{M}_2$ and $\mathbf{M}_2\mathbf{M}_1$ share a common characteristic polynomial. Therefore, for any real symmetric matrices $\mathbf{M}_1, \mathbf{M}_2$, if $\mathbf{M}_1$ is positive definite, then

$$F^{\mathbf{M}_1^{-1}\mathbf{M}_2} = F^{\mathbf{M}_2\mathbf{M}_1^{-1}} = F^{\mathbf{M}_1^{-1/2}\mathbf{M}_2\mathbf{M}_1^{-1/2}}$$

and $\mathbf{M}_1^{-1/2}$ is a symmetric matrix such that $\mathbf{M}_1^{-1} = \left(\mathbf{M}_1^{-1/2}\right)^2$.

Lemma 2.5. *For any real Hermitian matrices $\mathbf{M}_i$ $(1 \le i \le 4)$ of order p, if $\mathbf{M}_2$ is positive definite, $\mathbf{M}_3$ is positive semi-definite, and $\mathbf{M}_4$ is nonsingular, then*

$$\mathrm{K}(F^{\mathbf{M}_1\mathbf{M}_2^{-1}}, F^{\mathbf{M}_3\mathbf{M}_4^{-1}}) \le \frac{1}{p}\left(\mathrm{rank}(\mathbf{M}_1 - \mathbf{M}_3) + \mathrm{rank}(\mathbf{M}_2 - \mathbf{M}_4)\right).$$

Proof. By the triangle inequality, $\mathrm{K}(F^{\mathbf{M}_1\mathbf{M}_2^{-1}}, F^{\mathbf{M}_3\mathbf{M}_4^{-1}}) \le \mathrm{K}(F^{\mathbf{M}_1\mathbf{M}_2^{-1}}, F^{\mathbf{M}_3\mathbf{M}_2^{-1}}) +$ $\mathrm{K}(F^{\mathbf{M}_3\mathbf{M}_2^{-1}}, F^{\mathbf{M}_3\mathbf{M}_4^{-1}})$ which is, through Remark 2.4,

$$\begin{aligned}
&\mathrm{K}(F^{\mathbf{M}_2^{-1/2}\mathbf{M}_1\mathbf{M}_2^{-1/2}}, F^{\mathbf{M}_2^{-1/2}\mathbf{M}_3\mathbf{M}_2^{-1/2}}) + \mathrm{K}(F^{\mathbf{M}_3^{1/2}\mathbf{M}_2^{-1}\mathbf{M}_3^{1/2}}, F^{\mathbf{M}_3^{1/2}\mathbf{M}_4^{-1}\mathbf{M}_3^{1/2}}) \\
&\le \frac{1}{p}\,\mathrm{rank}(\mathbf{M}_2^{-1/2}(\mathbf{M}_1 - \mathbf{M}_3)\mathbf{M}_2^{-1/2}) + \frac{1}{p}\mathrm{rank}(\mathbf{M}_3^{1/2}(\mathbf{M}_2^{-1} - \mathbf{M}_4^{-1})\mathbf{M}_3^{1/2}),
\end{aligned}$$

by Proposition 1.1. For any square matrices $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3$ of the same order, $\mathrm{rank}(\mathbf{M}_2\mathbf{M}_1\mathbf{M}_3) \le \mathrm{rank}(\mathbf{M}_1)$ (the equality holds if $\mathbf{M}_2, \mathbf{M}_3$ are nonsingular). Thus, the former fraction is $\mathrm{rank}(\mathbf{M}_1 - \mathbf{M}_3)/p$ and the latter fraction is at most $\mathrm{rank}(\mathbf{M}_2^{-1} - \mathbf{M}_4^{-1})/p$. Because $\mathrm{rank}(\mathbf{M}_2^{-1} - \mathbf{M}_4^{-1}) = \mathrm{rank}(\mathbf{M}_2(\mathbf{M}_2^{-1} - \mathbf{M}_4^{-1})\mathbf{M}_4) =$ $\mathrm{rank}(\mathbf{M}_2 - \mathbf{M}_4)$, we obtain the desired consequence. □

By Lemma 2.5, we will establish the LSDs of Fisher matrices and Beta matrices from two independent ENPs in the Section 3. Moreover, we also use Proposition 1.1 in Section 4 to show the LSDs of random Wigner matrices, random Toeplitz matrices and random Hankel matrices, all being equipped with specially devised equi-correlated normal structure.

3. FISHER MATRICES AND BETA MATRICES

Assumption 3.1. Let two independent matrices $\mathbf{X}^{(i)} = \left[X_1^{(i)}, \ldots, X_{n_i}^{(i)}\right] \in \mathbb{R}^{p\times n_i}$ where all entries are i.i.d. random variables with unit variance, finite fourth moments and centered random variables. Suppose that the following equations are the sample covariance matrices of $\mathbf{X}^{(i)}$.

$$\mathbf{S}_i = n_i^{-1}\mathbf{X}^{(i)}\left(\mathbf{X}^{(i)}\right)^\top.$$

Remark 3.2. *Suppose all entries of $\mathbf{X}^{(i)}$ are* i.i.d. *random variables and obey a continuous distribution. Assume without loss of generality that $p \leq n_i$. Let j-th row of $\mathbf{X}^{(i)}$ be $\mathbf{x}_j$. For the first row, $\mathbf{x}_1$ is linearly independent because it is 0 with probability 0. For the second row, because it is drawn independently from the first row, the first row is fixed. The probability of $\mathbf{x}_2$ fall in to the span of a fixed row is 0 because $\mathbf{x}_2$ has a continuous density in $\mathbb{R}^{n_i}$. For general $k \leq n$, because $p \leq n$, the first $k-1$ rows forms a subspace in $\mathbb{R}^{n_i}$ and so $\mathbf{x}_k$ falls into that subspace with probability 0 (linear subspace has Lebesgue measure 0 in $\mathbb{R}^{n_i}$). Thus the first k rows are linearly independent. Let $k = p$. As a result, $\mathbf{X}^{(i)}$ has rank p with probability 1.*

3.1. **Fisher matrix.**

The new statistical tools proposed in Yao et al. [3, p. 7] are based on linear spectral statistics (LSS) on sample covariance matrices and Fisher matrices, which play an essential role in multivariate data analysis, such as the likelihood ratio test (LRT) for assessing the equality of variances from two populations [3, p. 151] [27]. A random Fisher matrix is defined as

$$\mathbf{F} = \mathbf{S}_1\mathbf{S}_2^{-1}.$$

In multivariate analysis of variance (MANOVA), the test on the equality of means is reduced to a statistic depending on a Fisher matrix which is a functional of the "between" sum of squares and the "within" sum of squares [28, p. 328]. In multivariate linear regression, the likelihood ratio criterion for testing linear hypotheses about regression coefficients is expressed as a functional of the eigenvalues of a Fisher matrix [28, p. 294].

Since the classical limit theorems for LSS of $\mathbf{F}$ are missmatch in the large-dimensional, Yao et al. [3, p. 30] showed the limiting behavior of Fisher matrices for large-dimensional in the following proposition.

Proposition 3.3 ([3, p. 30])**.** *Assume Assumption 3.1, $p/n_1 \to c_1 \in (0,\infty)$ and $p/n_2 \to c_2 \in (0,1)$. Then, almost surely, the ESD of $\mathbf{F}$ weakly converges to a deterministic distribution function F_{c_1,c_2}.*

Moreover, Yoshida [29] showed that F_{c_1,c_2} is the free F-Distribution which defined as the distribution of the ratio of two random variables from two freely independent free Poisson random variables.

By Stieltjes transform, Yao et al. [3, p. 30] established Proposition 3.3. However, in relaxing independent entries of $\mathbf{X}^{(i)}$ $(i = 1, 2)$, the Stieltjes transform of the ESD of $\mathbf{F}$ may be complicated because the Stieltjes transform of the ESD of $\mathbf{S}_1$ from ENP has a complicated expectation from correlated pairwise rows in $\mathbf{X}^{(1)}$ [3, p. 14, 30]. Instead, we will show the LSD of $\mathbf{F}$ derived from ENPs by decomposition (1) and the new rank inequality for ESDs (Theorem 2.5) because of its simplicity.

Assumption 3.4. Suppose two independent matrices $\widetilde{\mathbf{X}}^{(i)} = [\widetilde{X}_1^{(i)}, \dots, \widetilde{X}_{n_i}^{(i)}] \in \mathbb{R}^{p\times n_i}$ such that $\widetilde{X}_1^{(i)}, \dots, \widetilde{X}_{n_i}^{(i)} \overset{\text{i.i.d.}}{\sim} \mathrm{N}_p(\mathbf{0}, \mathbf{C}(r_i))$ with $0 \le r_i < 1$ $(i = 1, 2)$. Suppose that $\widetilde{\mathbf{S}}_i = n_i^{-1}\widetilde{\mathbf{X}}^{(i)}\left(\widetilde{\mathbf{X}}^{(i)}\right)^\top$ $(i = 1, 2)$.

Remark 3.5. *Note that $\widetilde{\mathbf{X}}^{(i)}$ can be written by $\mathbf{C}(r_i)^{1/2}\mathbf{X}^{(i)}$ such that all entries of $\mathbf{X}^{(i)}$ are* i.i.d. *and obey standard normal distribution. Thus, by Remark 3.2, $\widetilde{\mathbf{X}}^{(i)}$ is full rank matrix.*

Theorem 3.6. *Assume Assumption 3.4 and $\widetilde{\mathbf{F}} = \widetilde{\mathbf{S}}_1\widetilde{\mathbf{S}}_2^\top$. Suppose $p/n_1 \to c_1 \in (0, \infty)$ and $p/n_2 \to c_2 \in (0, 1)$. Then, it holds almost surely that $F^{\widetilde{\mathbf{F}}}$ weakly converges to a distribution function*

$$x \mapsto \mathrm{F}_{c_1,c_2}\left(\frac{1-r_2}{1-r_1}x\right)$$

where F_{c_1,c_2} denotes the deterministic distribution function assured in Proposition 3.3.

Proof. By decomposition (1) and Assumption 3.1,

$$\widetilde{\mathbf{X}}^{(i)} = \sqrt{1-r_i}\mathbf{X}^{(i)} + \mathbf{N}^{(i)}, \quad (i = 1, 2). \tag{2}$$

where all entries $\mathbf{X}^{(i)}$ are independent, standard normal random variables, and $\mathbf{N}^{(i)} = \sqrt{r_i}[\eta_j^{(i)}]_{p\times n_i}$ has rank at most 1. Hence, $\widetilde{\mathbf{X}}^{(i)}(\widetilde{\mathbf{X}}^{(i)})^\top - (1-r_i)\mathbf{X}^{(i)}\left(\mathbf{X}^{(i)}\right)^\top$ is equal to

$$\mathbf{N}^{(i)}(\sqrt{1-r_i}\mathbf{X}^{(i)} + \mathbf{N}^{(i)})^\top + \sqrt{1-r_i}\mathbf{X}^{(i)}(\mathbf{N}^{(i)})^\top. \tag{3}$$

By Fact 2.1, the rank of (3) is at most

$$\begin{aligned}\operatorname{rank}(\mathbf{N}^{(i)}(\sqrt{1-r_i}\mathbf{X}^{(i)} + \mathbf{N}^{(i)})^\top) + \operatorname{rank}(\sqrt{1-r_i}\mathbf{X}^{(i)}(\mathbf{N}^{(i)})^\top) &\le 2\operatorname{rank}(\mathbf{N}^{(i)}) \\ &\le 2. \end{aligned}\tag{4}$$

Because the entries of $\mathbf{X}^{(i)}$ are independent standard normal random variables, the entries of $\mathbf{X}^{(i)}$ for $i = 1, 2$ have finite fourth moment. Moreover, since $n_2 > p$, almost surely $\widetilde{\mathbf{S}}_2^{-1}$ and $\mathbf{S}_2^{-1}$ are well-defined in the limit p.

By Lemma 2.5 and (4), $\mathrm{K}\left(F^{\widetilde{\mathbf{F}}}, F^{\frac{1-r_1}{1-r_2}\mathbf{F}}\right) \leq 4/p \to 0$ as $p \to \infty$. Note that

$$F^{\frac{1-r_1}{1-r_2}\mathbf{F}}(x) = F^{\mathbf{F}}\left(\frac{1-r_2}{1-r_1}x\right).$$

By Proposition 2.3 and Proposition 3.3, the desired consequence follows. □

To sum up, if the LSDs of Fisher matrices composed from two independent centered ENPs with unit variance and nonnegative $r_1, r_2 < 1$, the LSDs of Fisher matrices are similar to the LSDs of Fisher matrices composed of two independent centered populations with i.i.d. random variables and finite enough moment but scaled by $(1-r_2)/(1-r_1)$.

3.2. **Beta matrices.**

Several hypothesis testing apply Beta matrices in multivariate analysis when $\alpha = n_1/n_2$ with $n_1, n_2 > p$ [30, p. 332] [25]. For example, in the two-sample test for equality of variances, terms of LSS of the ESDs of Beta matrices [25] can formulate the LRT for hypothesis $H_0 : \Sigma_1 = \Sigma_2$. A random Beta matrix is defined as

$$\mathbf{B} = \mathbf{S}_1(\mathbf{S}_1 + \alpha\mathbf{S}_2)^{-1}\ (\alpha > 0).$$

In multivariate analysis of variance (MANOVA), the test on the equality of means is reduced to a statistic depending on a Beta matrix which is a functional of the "between" sum of squares and the "within" sum of squares [28, p. 331]. However, in large-dimensional, the limiting behaviors of Beta matrices are needed because the LSS of Beta matrices are not valid [25].

Remark 3.7 ([24, p. 325])**.** Assume all conditions of Proposition 3.3 are satisfied. Then, almost surely, the ESD of $\dot{\mathbf{B}} = \mathbf{S}_2(\mathbf{S}_2 + \alpha\mathbf{S}_1)^{-1}$ weakly converges to a distribution function

$$x \mapsto 1 - \mathrm{F}_{c_1,c_2}\left(\frac{1}{\alpha}\left(\frac{1}{x} - 1\right)_-\right), \quad x > 0$$

where $\mathrm{F}_{c_1,c_2}(x_-)$ is the left-limit at x.

Later it has been shown that Bai-Silverstein [24, p. 325] established the LSD of a Beta matrix is similar to LSD of a Fisher matrix. They just requires that the original variables to construct $\mathbf{S}_1$ and $\mathbf{S}_2$ are respectively i.i.d. with sufficient moments and $p/n_1 \to c_1 > 0$, $p/n_2 \to c_2 \in (0,1)$. On the other hand, for $p > \max(n_1, n_2)$ and $p < n_1 + n_2$, by Stieltjes transform, Bai et al. [25] showed the LSD of $\mathbf{B}$ which is showed in the following proposition.

Proposition 3.8 ([25, Theorem 1.1])**.** *Assume Assumption 3.1 and the following conditions:*

(1) $\alpha > 0$ *and* $p/n_1 \to c_1 > 0$.
(2) $p/n_2 \to c_2 > 0$ *and* $\frac{p}{n_1+n_2} \to \frac{c_1 c_2}{c_1+c_2} \in (0,1)$.

Then, almost surely, $F^{\mathbf{B}}$ *weakly converges to a deterministic distribution function* $\mathrm{BM}_{\alpha,c_1,c_2}$.

In relaxing independent entries of $\mathbf{X}^{(i)}$ $(i = 1, 2)$, similar to the reason in Fisher matrices, the Stieltjes transform of $F^{\mathbf{B}}$ may be complicated. Instead, we will show that the LSD of $\mathbf{B}$ from two centered ENPs with unit variances as $p > \max(n_1, n_2)$ and $p < n_1 + n_2$ by the new rank inequality corresponding to LSD of $\mathbf{B}$ under i.i.d. conditions (Theorem 3.8).

Theorem 3.9. *Assume Assumption 3.4 and* $\widetilde{\mathbf{B}} = \widetilde{\mathbf{S}}_1(\widetilde{\mathbf{S}}_1 + \alpha\widetilde{\mathbf{S}}_2)^{-1}$. *Suppose the conditions (1)-(2) of Proposition 3.8. Then, almost surely,* $F^{\widetilde{\mathbf{B}}}$ *weakly converges to the distribution function* BM_{s,c_1,c_2} *where* $s = \alpha(1 - r_2)/(1 - r_1)$. *Here,* BM_{s,c_1,c_2} *is the deterministic distribution function assured in Proposition 3.8.*

Proof. By decomposition (1),

$$\widetilde{\mathbf{X}}^{(i)} = \sqrt{1 - r_i}\mathbf{X}^{(i)} + \mathbf{N}^{(i)}, \quad (i = 1, 2).$$

Note that $\mathbf{S}_i$ are positive semi-definite matrices for $i = 1, 2$. Because $\frac{c_1 c_2}{c_1 + c_2} \in (0, 1)$, we have $p < n_1 + n_2$ almost surely. Thus, $\mathbf{S}_1 + \alpha\mathbf{S}_2$ is invertible almost surely in $p \to \infty$.

As with (4), we can show that $\mathrm{rank}\left(\widetilde{\mathbf{X}}^{(i)}\left(\widetilde{\mathbf{X}}^{(i)}\right)^\top - (1 - r_i)\mathbf{X}^{(i)}\mathbf{X}^{(i)\top}\right) \le 2$ $(i = 1, 2)$. Therefore, by Lemma 2.5, $\mathrm{K}\left(F^{\widetilde{\mathbf{B}}}, F^{(1-r_1)\mathbf{S}_1((1-r_1)\mathbf{S}_1 + \alpha(1-r_2)\mathbf{S}_2)^{-1}}\right)$ is at most

$$\begin{aligned}
&\frac{1}{p}\left(\mathrm{rank}(\widetilde{\mathbf{S}}_1 - (1 - r_1)\mathbf{S}_1) + \mathrm{rank}(\widetilde{\mathbf{S}}_1 + \alpha\widetilde{\mathbf{S}}_2 - \mathbf{S}_1 - \alpha(1 - r_2)\mathbf{S}_2)\right) \\
&\le \frac{1}{p}\left(\mathrm{rank}(\widetilde{\mathbf{S}}_1 - (1 - r_1)\mathbf{S}_1) + \mathrm{rank}(\widetilde{\mathbf{S}}_1 - (1 - r_1)\mathbf{S}_1) + \mathrm{rank}(\alpha\widetilde{\mathbf{S}}_2 - \alpha(1 - r_2)\mathbf{S}_2)\right) \\
&\le 6/p \to 0 \quad (p \to \infty).
\end{aligned}$$

Moreover,

$$(1 - r_1)\mathbf{S}_1((1 - r_1)\mathbf{S}_1 + \alpha(1 - r_2)\mathbf{S}_2)^{-1} = \mathbf{S}_1\left(\mathbf{S}_1 + \frac{\alpha(1 - r_2)}{(1 - r_1)}\mathbf{S}_2\right)^{-1}$$

Thus, by Proposition 2.3 and Proposition 3.8, the desired consequence follows. □

As a result, the LSDs of Beta matrices are the LSDs of Beta matrices from two independent centered populations with i.i.d. random variables and finite enough moments where the parameter is $\alpha(1-r_2)/(1-r_1)$ if the Beta matrix is composed of two independent centered ENPs with unit variances, and nonnegative $r_1, r_2 < 1$.

4. SYMMETRIC RANDOM MATRICES WITH CORRELATED PAIRING STANDARD NORMAL ENTRIES BY A FIXED COEFFICIENT

To further our exploration of the rank inequality for ESDs (Proposition 1.1), we will apply these techniques to various symmetric random matrices which all pairings of entries from those matrices are equi-correlated and standard normal random variables. Specifically, the random matrices are Wigner matrices, symmetric Toeplitz matrices, Hankel matrices, Markov matrices [17], and banded Toeplitz matrices [6].

Assumption 4.1. $\widetilde{x}_i, \widetilde{x}_{ij}$ $(i, j = 0, 1, 2, \ldots; i \le j)$ *are standard random variables mutually correlated with* r $(0 \le r < 1)$.

The following equalities are the well-known decomposition of $\widetilde{x}_i, \widetilde{x}_{ij}$ $(i, j = 0, 1, 2, \ldots)$.

$$\mathbf{Z}(\widetilde{x}_1, \widetilde{x}_2, \ldots) = \sqrt{1-r}\mathbf{Z}(x_1, x_2, \ldots) + \sqrt{r}\eta\mathbf{Z}(1, 1, \ldots) \tag{5}$$

$$\mathbf{Z}(\widetilde{x}_{11}, \widetilde{x}_{12}, \ldots) = \sqrt{1-r}\mathbf{Z}(x_{11}, x_{12}, \ldots) + \sqrt{r}\eta\mathbf{Z}(1, 1, \ldots) \tag{6}$$

where x_i, and x_{ij} $(i, j = 0, 1, 2, \ldots; i \le j)$ are i.i.d. standard normal random variables. Here, $\mathbf{Z}(.)$ is a matrix of linear combination from several random variables.

4.1. Wigner matrices.

A *Wigner matrix* of order p is a real symmetric matrix $\mathbf{W} = [x_{ij}]_{p\times p}$ such that above-diagonal entries x_{ij}, where $1 \le i \le j \le p$, are independent centered random variables such that

- the diagonal entries x_{ii} are i.i.d.; and
- the off-diagonal entries x_{ij} are i.i.d. and have unit variance.

Wigner [1, 2] found that the gaps between the lines in the spectrum of a heavy nucleus are like the gaps between the eigenvalues of an extensive $p \times p$ symmetric or Hermitian random matrix with random entries. Some of the physical applications of Wigner matrices are surveyed in [31, p. 13]. Let Sc_a be the distribution function of a *semicircle law* scaled by $a > 0$.

Proposition 4.2 ([17, Theorem 2.1])**.** *Assume that* $\mathbf{W}$ *is a Wigner matrix of order* p. *Let* $p \to \infty$. *Then, almost surely,* $F^{\mathbf{W}/\sqrt{p}}$ *converges pointwise to* Sc_2.

Theorem 4.3. *Assume* $\widetilde{\mathbf{W}} = [\widetilde{x}_{ij}]_{p\times p}$ *is a real symmetric matrix of order* p *and Assumption 4.1. Let* $p \to \infty$. *Then, almost surely,* $F^{\widetilde{\mathbf{W}}/\sqrt{p}}$ *converges pointwise to the distribution function* $\mathrm{Sc}_{2\sqrt{1-r}}$.

Proof. By applying decomposition (6), we can find *independent*, standard normal random variables x_{ij}, η $(1 \le i,\ j \le p)$ such that the rank of $\mathbf{Z}(1, 1, \ldots)$ is at most 1. By this, the Proposition 1.1 implies

$$\mathrm{K}\left(F^{\widetilde{\mathbf{W}}/\sqrt{p}}, F^{\sqrt{1-r}\mathbf{W}/\sqrt{p}}\right) \le \frac{1}{p} \to 0 \quad (p \to \infty).$$

By Proposition 4.2, $F^{\sqrt{1-r}\mathbf{W}/\sqrt{p}}$ converges pointwise to $\mathrm{Sc}_{2\sqrt{1-r}}$ almost surely. Hence, $F^{\widetilde{\mathbf{W}}/\sqrt{p}}$ does so. □

When the Wigner matrices has a perfect correlation on diagonal entries and i.i.d. random variables on off diagonal, the Wigner matrices is then symmetric Toeplitz matrices in the following subsection.

4.2. **Symmetric Toeplitz matrices.**

Let $x_0, x_1, \ldots$ be i.i.d. real random variables with mean zero and unit variance. Define a random Toeplitz matrix $\mathbf{T} = \left[x_{|i-j|}\right]_{1\le i,j\le p}$ [5], i.e., a matrix of the form

$$\mathbf{T} = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{p-2} & x_{p-1} \\ x_1 & x_0 & x_1 & & & x_{p-2} \\ x_2 & x_1 & x_0 & & \ddots & \vdots \\ \vdots & & & \vdots & & x_2 \\ x_{p-2} & & & & x_0 & x_1 \\ x_{p-1} & x_{p-2} & \cdots & x_2 & x_1 & x_0 \end{bmatrix}.$$

Hammond and Miller [32] demonstrated that almost surely the kth moment of $F^{\mathbf{T}}$ converges to the moments of a new universal distribution, independent of p. The new distribution seemed normal, and numerical simulations and heuristics corroborate a conjecture. That result was extended by Bryc, Dembo, and Jiang [5].

Proposition 4.4 ([5, Theorem 1.1])**.** *Let* $\mathbf{T}$ *be as above. Suppose* $p \to \infty$*. Then, almost surely, the ESD* $F^{\frac{1}{\sqrt{p}}\mathbf{T}}$ *converges weakly to a deterministic probability distribution* $\mathrm{F}_{\mathbf{T}}$ *that does not depend on the distribution of* x_1*, and has unbounded support.*

Theorem 4.5. *Assume* $\widetilde{\mathbf{T}} = \left[\widetilde{x}_{|i-j|}\right]_{1\le i,j\le p}$ *and Assumption 4.1. Then, almost surely, in* $p \to \infty$*, the ESD* $F^{\frac{1}{\sqrt{p}}\widetilde{\mathbf{T}}}$ *to a deterministic probability distribution* $\mathrm{F}_{\mathbf{T}}$ *scaled by* $\sqrt{1-r}$*, and has unbounded support.*

Proof. By applying decomposition (5), we can find independent, standard normal random variables η, x_k $(0 \le k \le (p-1))$ such that the rank of $\mathbf{Z}(1, 1, \ldots)$ is at most 1. By this, the Proposition 1.1 implies

$$\mathrm{K}\left(F^{\widetilde{\mathbf{T}}/\sqrt{p}},\, F^{\sqrt{1-r}\mathbf{T}/\sqrt{p}}\right) \le \frac{1}{p} \to 0 \quad (p \to \infty).$$

By Proposition 2.3 and Proposition 4.4, the desired consequence follows. □

4.3. **Hankel matrices.**

Let $x_0, x_1, \ldots$ be i.i.d. real random variables with mean zero and unit variance. Define a random Hankel matrix $\mathbf{H} = [x_{i+j-1}]_{1\le i,j\le p}$ [5], i.e., a matrix of the form

$$\mathbf{H} = \begin{bmatrix} x_1 & x_2 & \cdots & \cdots & x_{p-1} & x_p \\ x_2 & x_3 & & & x_p & x_{p+1} \\ \vdots & & & \iddots & x_{p+1} & x_{p+2} \\ x_{p-2} & x_{p-1} & \iddots & & & \vdots \\ x_{p-1} & x_p & & & x_{2p-3} & x_{2p-2} \\ x_p & x_{p+1} & \cdots & \cdots & x_{2p-2} & x_{2p-1} \end{bmatrix}.$$

Proposition 4.6 ([5, Theorem 1.2]). *Let* $\mathbf{H}$ *be as above. Suppose* $p \to \infty$. *Then, almost surely the ESD* $F^{\frac{1}{\sqrt{p}}\mathbf{H}}$ *converges weakly to a deterministic probability distribution* $\mathrm{F}_{\mathbf{H}}$ *which does not depend on the distribution of* x_1, *and has unbounded support and is unimodal.*

Theorem 4.7. *Assume* $\widetilde{\mathbf{H}} = [\widetilde{x}_{i+j-1}]_{1\le i,j\le p}$ *and Assumption 4.1. Then, almost surely,* $F^{\frac{1}{\sqrt{p}}\widetilde{\mathbf{H}}}$ *converges weakly as* $p \to \infty$ *to a deterministic probability distribution* $\mathrm{F}_{\mathbf{H}}$ *scaled by* $\sqrt{1-r}$, *and has unbounded support.*

Proof. By applying decomposition (5), we can find independent, standard normal random variables $\eta,\ x_k\ (0 \le k \le (2p-1))$ such that the rank of $\mathbf{Z}(1,1,\ldots)$ is at most 1. By this, the Proposition 1.1 implies

$$\mathrm{K}\left(F^{\widetilde{\mathbf{H}}/\sqrt{p}},\, F^{\sqrt{1-r}\mathbf{H}/\sqrt{p}}\right) \le \frac{1}{p} \to 0 \quad (p \to \infty).$$

By Proposition 2.3 and Proposition 4.6, the desired consequence follows. □

4.4. **Banded symmetric Toeplitz matrices.**

Define a symmetric Toeplitz matrix with band structure $\mathbf{T}_b = \left[x_{|i-j|}\right]_{1\le i,j\le p}$ such that $x_{|i-j|} = 0$ for $|i-j| > b$.

Proposition 4.8 ([6, Theorem 2.1]). *Suppose that* $\mathbf{T}_b$ *is a banded symmetric Toeplitz matrix with the band of width* $b = b(p)$. *Let* $k = |i-j|$. *Assume that the nonzero entries are centered* i.i.d. *random variables such that* $\mathrm{E}\, x_k^2 = 1/b$ *and* $\sup_{k,p} \mathrm{E}\left|\sqrt{b}x_k\right|^4 < \infty$. *If both* $b = b(p) \to \infty$ *and* $b/p \to 0$ *as* $p \to \infty$, *then for every* x, *the expectation of* $F^{\mathbf{T}_b}(\sqrt{2}x)$ *converges to the cumulative distribution function of standard normal distribution. Moreover,* $\mathrm{Var}(F^{\mathbf{T}_b}(x))$ *tends to 0.*

Let $\widetilde{\mathbf{T}}_b$ be a banded symmetric Toeplitz matrix with the band of width $b = b(p)$. Suppose that the entries of $\widetilde{\mathbf{T}}_b$ in the band obey a centered normal distribution with variance $1/b$ and are mutually correlated by a fixed nonnegative correlation coefficient r. By decomposition (5), we have $\sqrt{r}\eta\mathbf{Z}(1,1\ldots)$ is a $p \times p$ banded symmetric Toeplitz matrix for $b > 0$ such that all nonzero entries are $\sqrt{r}\eta$, and centered normal distribution with variance $1/b$. Note that the rank of $\sqrt{r}\eta\mathbf{Z}(1,1\ldots)$ is $p-d$ where d is the dimension of null space from matrix $\sqrt{r}\eta\mathbf{Z}(1,1\ldots)$. If $\sqrt{r}\eta\mathbf{Z}(1,1\ldots)$ is a banded symmetric Toeplitz matrix of 1 with band width b, d is at most b. By this and $b/p \to 0$,

$$\mathrm{K}\left(F^{\widetilde{\mathbf{T}}_b/\sqrt{2}}, F^{\sqrt{1-r}\mathbf{T}_b/\sqrt{2}}\right) \le \frac{p-d}{p} \to 1. \tag{7}$$

as $p \to \infty$. As a result, the rank inequality (Proposition 1.1) is inapplicable to compute the LSD of $\widetilde{\mathbf{T}}_b/\sqrt{2}$.

4.5. **Markov matrices.**

Let $\{x_{ij}\}$ for $j \geq i \geq 1$ be an infinite upper triangular array of i.i.d. random variables and define $x_{ij} = x_{ji}$ for $j > i \geq 1$. Define $\mathbf{M}$ a $p \times p$ Markov matrix given by

$$\mathbf{M} = \begin{bmatrix} -\sum_{j=2}^{p} x_{1j} & x_{12} & x_{13} & & \cdots & x_{1p} \\ x_{21} & -\sum_{j\neq 2}^{p} x_{2j} & x_{23} & & \cdots & x_{2p} \\ \vdots & & \ddots & & & \vdots \\ x_{k1} & x_{k2} & \cdots & -\sum_{j\neq k}^{p} x_{kj} & \cdots & x_{kp} \\ \vdots & \vdots & & & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & & & -\sum_{j=1}^{p-1} x_{pj} \end{bmatrix}.$$

Proposition 4.9 ([5, Theorem 1.3])**.** *Let* $\mathbf{T}$ *be as above with* $\mathrm{E}\, x_{12} = 0$ *and* $\mathrm{Var}(x_{12}) = 1$*. Suppose* $p \to \infty$*. Then, almost surely the ESD of* $F^{\mathbf{M}/\sqrt{p}}$ *converges weakly to the free convolution of the semicircle and standard normal distribution. This distribution is a nonrandom symmetric distribution with smooth bounded density, does not depend on the distribution of* x_{12} *and has unbounded support.*

Suppose that $\widetilde{\mathbf{M}}$ be a $p \times p$ Markov matrix such that $\mathrm{E}(x_{ij}x_{kl}) = r \quad ((i,j) \neq (k,l),\ 1 \leq i \leq j \leq p,$ and $1 \leq k \leq l \leq p)$. By (1), $\eta\mathbf{Z}(1, 1\ldots)$ is a $p \times p$ matrix given by

$$\begin{bmatrix} -(p-1)\eta & \cdots & \eta \\ \vdots & \ddots & \vdots \\ \eta & \cdots & -(p-1)\eta \end{bmatrix}.$$

Note that $\mathrm{P}(\eta = 0) = 0$ because η is continuous variable. Since the above matrix with the first row and the first column erased is a diagonally dominant matrix [18, p. 352], $\eta\mathbf{Z}(1, 1\ldots)$ is a full matrix of order $p-1$. Therefore, the rank of $\eta\mathbf{Z}(1, 1\ldots)$ is at least $p-1$ almost surely. Hence, by the rank inequality for ESDs (Proposition 1.1),

$$\mathrm{K}\left(F^{\widetilde{\mathbf{M}}/\sqrt{p}},\, F^{\sqrt{1-r}\mathbf{M}/\sqrt{p}}\right) \leq \mathrm{rank}\left(\frac{p-1}{p}\right) \to 1. \tag{8}$$

As a result, the rank inequality (Proposition 1.1) is inapplicable for computing the LSD of $\widetilde{\mathbf{M}}/\sqrt{p}$.

5. THE LSDs OF SIMULATION RANDOM MATRICES AND REAL DATASETS

In this section, we demonstrate the LSDs of simulation random matrices from ENP by presenting the histogram of random Fisher matrices, Beta matrices, Toeplitz matrices and Hankel matrices. Moreover, for real datasets, we show the LSDs of two independent microarray datasets by estimating r_i $(i = 1, 2)$ from Akama.

5.1. Simulation study of Fisher and Beta matrices.

Throughout the simulation of Fisher and Beta matrices, the Fisher and Beta matrices are composed of two independent samples $\mathbf{X}^{(i)}$ $(i = 1, 2)$ from independent 1000-dimensional, centered normal populations such that all entries have unit variance and any correlation coefficient between different variables are fixed non-negative $r_1, r_2 < 1$.

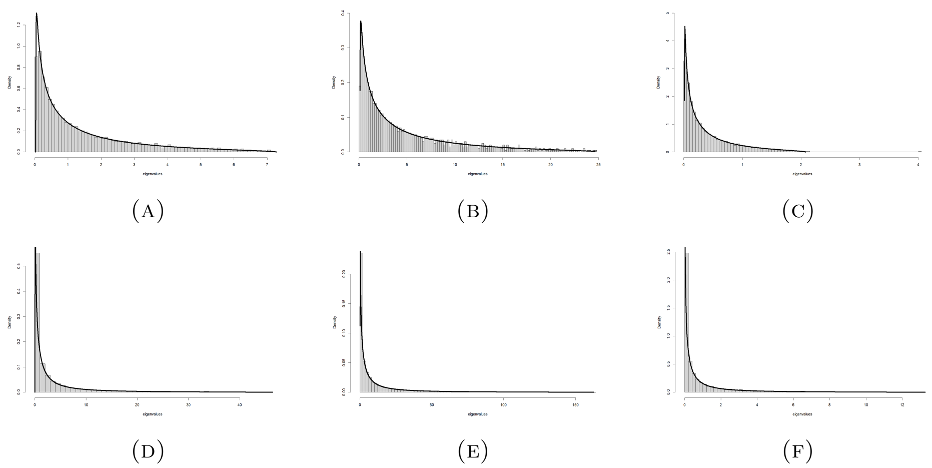
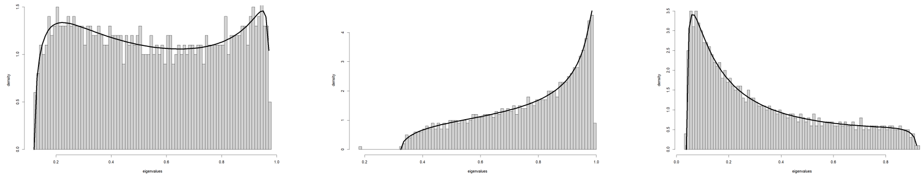


FIGURE 1. Histogram of numerical eigenvalues of 1000×1000 Fisher matrices. From (A)-(F), $(p/n_1, p/n_2, r_1, r_2)$ is equal to (0.7,0.3,0,0), (0.7,0.3,0.3,0.8), (0.7,0.3,0.8,0.3), (1.3,0.7,0,0), (1.3,0.3,0.3,0.8), (0.7,0.3,0.8,0.3). The tick curve is the density function of $\mathbf{F}_{p/n_1,p/n_2}$ scaled by $(1-r_2)/(1-r_1)$.

The Figure 1 shows that the density functions of $\mathbf{F}_{p/n_1,p/n_2}$ for $0 \le r < 1$ fit to histogram of Fisher matrices which is composed from two independent population where all elements are standard normal distribution mutually correlated with $0 \le r < 1$. Since $(1-r_1)/(1-r_2) < 1$ for $(r_1, r_2) = (0.8, 0.3)$, as Figure 2 (B), the histogram has the largest eigenvalue outside of thick curve from $(1-r_2)/(1-r_1)\mathbf{F}'((1-r_2)/(1-r_1)x)$ for all $x \in [\frac{(1-r_1)(1-(c_1+c_2-c_1c_2)^{1/2})^2}{(1-r_2)(1-c_2)^2}, \frac{(1-r_1)(1+(c_1+c_2-c_1c_2)^{1/2})^2}{(1-r_2)(1-c_2)^2}]$.

The Figure 2 shows that the density functions of $\mathrm{BM}_{s,p/n_1,p/n_2}$ for $0 \le r < 1$ fit to histogram of Fisher matrices which is composed from two independent population where all elements are standard normal distribution mutually correlated with $0 \le r < 1$. However, as Figure 2 (B), the histogram has the extreme values outside of thick curve from $\mathrm{BM}'_{s,p/n_1,p/n2}$. This extreme value will be identified as future work for our research about LSDs of Beta matrices from ENP.

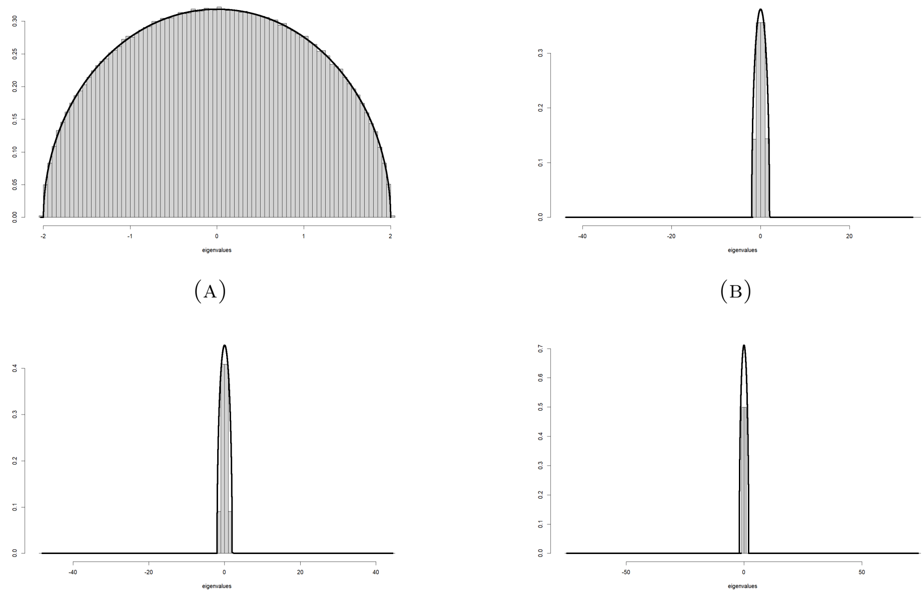


(A) (B) (C)

FIGURE 2. Histogram of numerical eigenvalues of 1000×1000 Beta matrices with $\alpha = 1$. From (A)-(C), $(p/n_1, p/n_2, r_1, r_2)$ is equal to (0.7,0.3,0,0), (0.7,0.3,0.3,0.8), (0.7,0.3,0.8,0.3). The tick curve is the density function of $\mathrm{BM}_{s,p/n_1,p/n_2}$ where $s = (1 - r_2)/(1 - r_1)$.

5.2. **Simulation study of symmetric matrices.**

In this simulation, we construct the symmetric matrices following decomposition (5) and (6) such that x_i, and x_{ij} $(i, j = 0, 1, 2, \ldots, p\ ; i \leq j)$ are i.i.d. standard normal random variables. We set $p = 500$ and $r = 0, 0.3, 0.5, 0.8$.

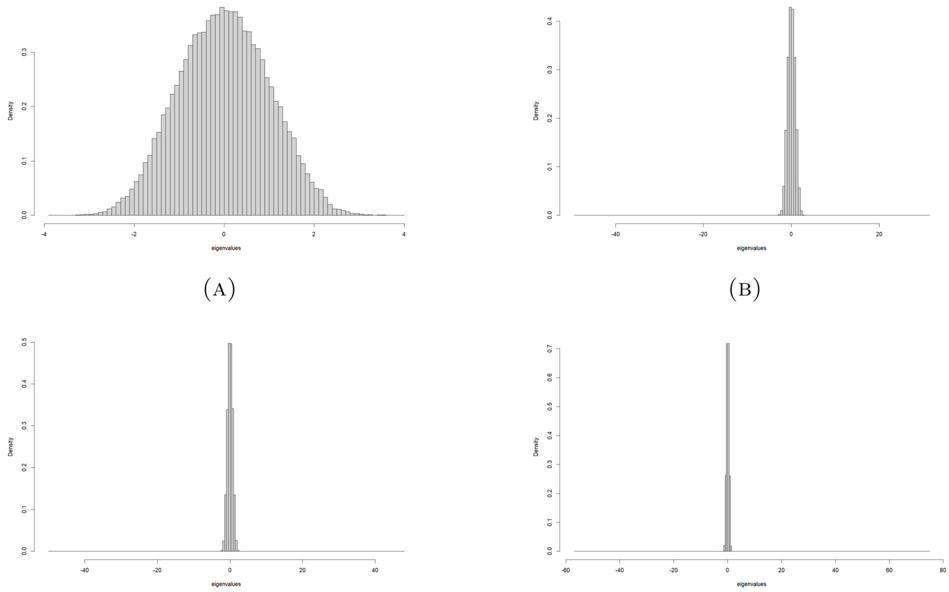


(A) (B)

(C) (D)

FIGURE 3. Histogram of numerical eigenvalues Wigner matrices from 100 matrices. From (A)-(D), r is equal to 0,0.3,0.5,0.8; and the tick curve is $\mathrm{Sc}_{2\sqrt{1-r}}$.

The Figure 3 shows that the density functions of $\mathrm{Sc}_{2\sqrt{1-r}}$ for $0 \le r < 1$ fit to histogram of Wigner matrices which all elements are standard normal distribution mutually correlated with $0 \le r < 1$. Moreover, as Figure 2, (B)-(D) present the histogram and thick curve of semi-circle distribution scaled by $\sqrt{1-r}$.



(A) (B) (C) (D)

FIGURE 4. Histogram of numerical eigenvalues symmetric Toeplitz matrices from 100 matrices. From (A)-(D), r is equal to 0,0.3,0.5,0.8.

The Figure 4 shows that the histogram of eigenvalues of symmetric Toeplitz matrices with all entries are standard normal and duplicate correlated entries r. Moreover, as Figure 4, (B)-(D) present the histogram is scaled by $\sqrt{1-r}$. Figure 5 presents that the histogram of eigenvalues of symmetric Toeplitz matrices with all entries are standard normal and duplicate correlated entries r. Moreover, as Figure 5, (B)-(D) present the histogram is scaled by $\sqrt{1-r}$.

5.3. **Real dataset.**

For the scaling parameter $1-r_i$ of Fisher matrices of Theorem 3.6, Akama [23] proposed $1 - \lambda_1(\mathbf{S}_i)/p$ $(i = 1, 2)$ in:

Theorem 5.1 (Akama [23])**.** *Let* $\mathbf{S}_i$ $(i = 1, 2)$ *be a sample covariance matrix formed from i-th population* $\mathrm{N}_p(\mathbf{0},\, \mathbf{C}(r_i))$ *for a deterministic constant* $r_i \in [0,\, 1)$. *Suppose*

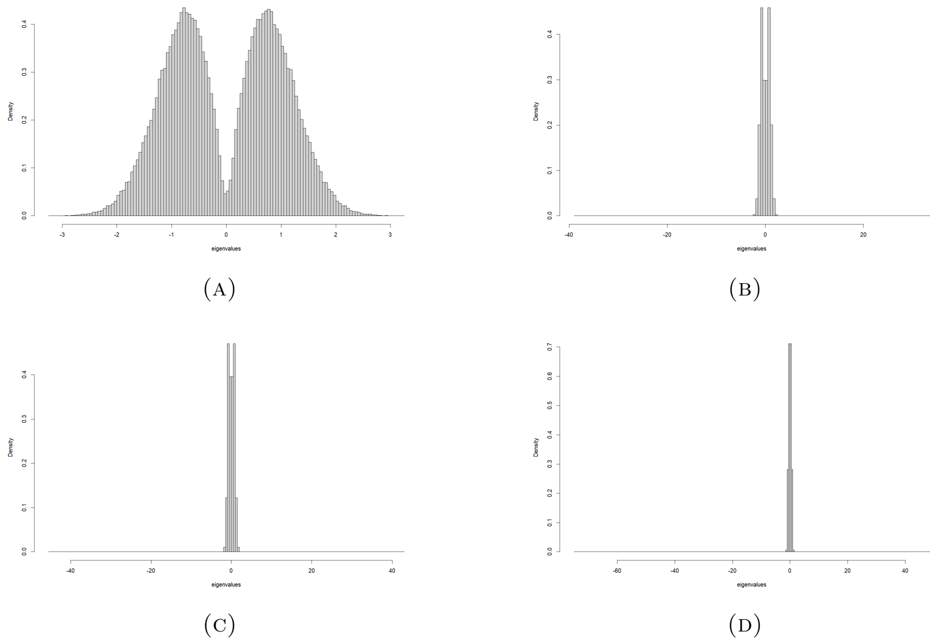


Figure 5. Histogram of numerical eigenvalues Hankel matrices from 100 matrices. From (A)-(D), r is equal to 0,0.3,0.5,0.8.

$p, n \to \infty$ and $p/n \to c \in (0, \infty)$. Then, almost surely,

$$\frac{\lambda_1(\mathbf{S}_i)}{p} \to r.$$

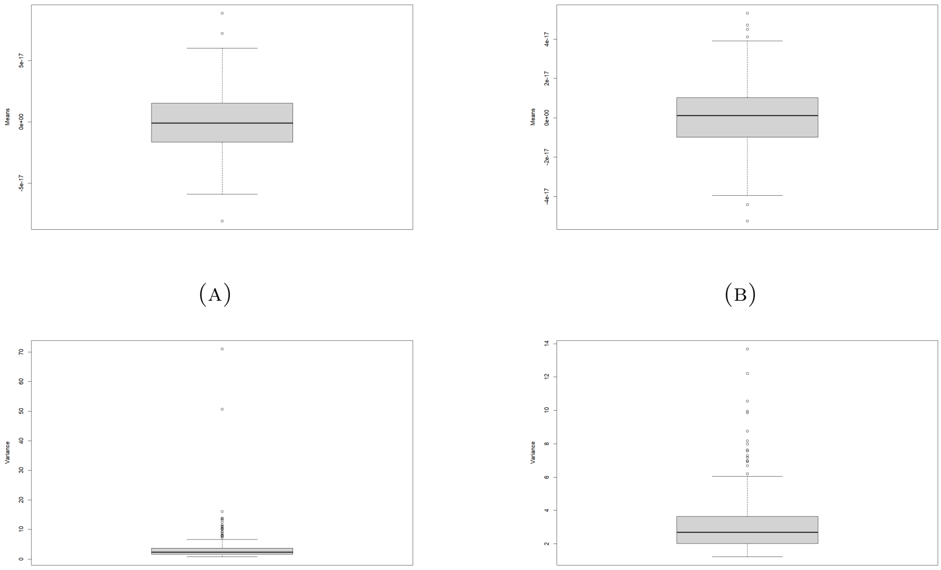
With this, we show Theorem 3.6 by real datasets such as the two independent returns of S&P500 stocks of two specific periods and two independent class of a microarray dataset from [33].

5.3.1. *Finance dataset.*

We consider the two datasets of returns of p S&P500 stocks for n trading days. Table 1 is the list of p/n, $\lambda_1(\mathbf{S})/p$ and $p = 212$ S&P500 stocks of two periods.

Table 1. The returns of S&P500 datasets.

No	Period	p/n	$\lambda_1(\mathbf{S})/p$	p
1	1993-01-04-1995-12-29	.280	.110	210
2	2012-08-01-2022-08-01	.083	.400	210



(A) (B)

(C) (D)

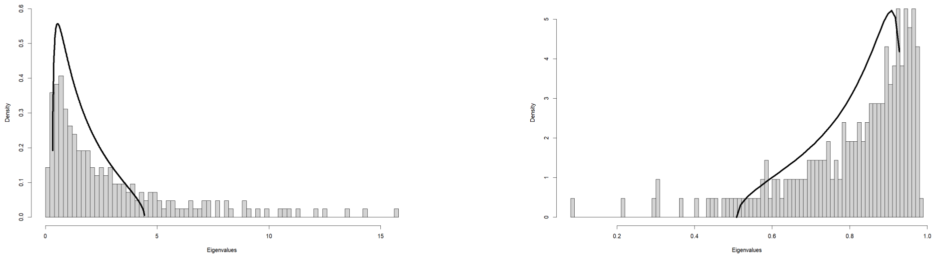
FIGURE 6. The box plots of S&P500 stocks in Table 1. (A)-(B) are the means and variance from returns of the first dataset. (C)-(D) are the means and variance from returns of the second dataset.

As Figure 6, the box plots (A)-(B) show that the means is close to zero. However, the box plots (C)-(D) have very large variance and there are many outlier.

Let $\widetilde{\mathbf{S}}_1$ and $\widetilde{\mathbf{S}}_2$ are the sample covariance matrix of the first and second dataset in Table 1. As Figure 7, the thick curve fits to histogram (A) of Fisher matrix but some bins of histogram outside the thick curve. Otherwise, the thick curve unfits to histogram (B) of Fisher matrix and many bins of histogram outside the thick curve. Following 6 (C), the variance of the first dataset are very far from 1 which is different with our assumption in Theorem 3.6 and Theorem 3.6. By this, the histogram of Fisher and Beta matrix may be disparate from our distribution function of F_{c_1,c_2} and BM_{s,c_1,c_2}. However, this result will be our future research to more generalize the assumption for variance from each independent datasets.

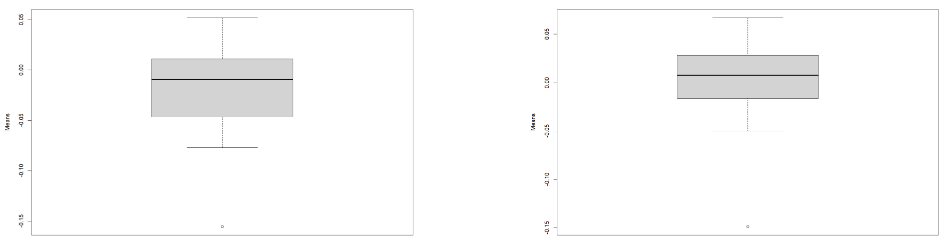
5.3.2. *Microarray dataset.*

We used microarray data sets of breast cancer with restricted variables $p = 50$ genes. The data sets consist of two classes: π_1 :cancer (57 samples) and π_2 :normal breast (111 samples). See Gravier et al. [34] for the details. The data sets are available at [33]. Table 2 is the list of p/n, $\lambda_1(\mathbf{S})/p$ and $p = 50$ genes from two classes.

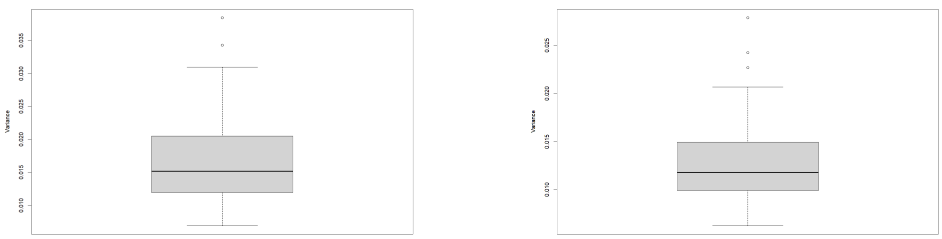


(A) (B)

Figure 7. (A) The histogram of Fisher matrix $\widetilde{\mathbf{F}}$ composed by two independent returns of S&P500 datasets from Table 1. The thick curve is density function of $\mathrm{F}_{0.28,0.4}((1-\lambda_1(\mathbf{S}_2))/(1-\lambda_1(\mathbf{S}_2))x)$ for all $x \in \mathbb{R}$. (B) The histogram of Beta matrix $\widetilde{\mathbf{B}}$ composed by two independent returns of S&P500 datasets from Table 1. The thick curve is density function of $\mathrm{BM}_{0.2,0.28,0.4}(x)$ for all $x \in \mathbb{R}$.



(A) (B)



(C) (D)

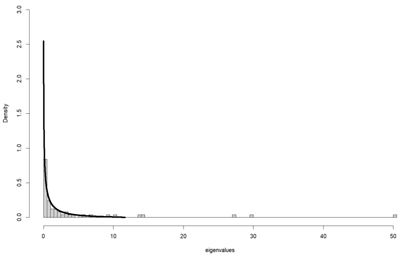
Figure 8. The box plots of the microarray dataset in Table 2. (A)-(B) are the means and variance from the microarray dataset of the first class dataset. (C)-(D) are the means and variance from the microarray dataset of the second class dataset.

Table 2. The microarray dataset of breast cancer patients.

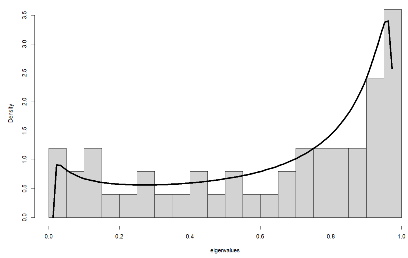
No	Class	p/n	$\lambda_1(\mathbf{S})/p$	p
1	π_1 (cancer)	.877	.58	50
2	π_2 (normal)	.450	.56	50

As Figure 8, the box plots (A)-(B) show that the means is close to zero. Otherwise, the box plots (C)-(D) have very small variance. However, the outlier just two variables for box plot (C) and three variables for box plot (C).

Let $\widetilde{\mathbf{S}}_1$ and $\widetilde{\mathbf{S}}_2$ are the sample covariance matrix of the first and second class in Table 2. As Figure 9, the thick curve fits to histogram (A) of Fisher matrix but



(A)



(B)

Figure 9. (A) The histogram of Fisher matrix $\widetilde{\mathbf{F}}$ composed by two independent class of the microarray dataset from Table 2. The thick curve is density function of $\mathrm{F}_{0.87,0.45}((1-\lambda_1(\widetilde{\mathbf{S}}_2))/(1-\lambda_1(\widetilde{\mathbf{S}}_2))x)$ for all $x \in \mathbb{R}$. (B) The histogram of Beta matrix $\widetilde{\mathbf{B}}$ composed by two independent class of the microarray dataset from Table 2. The thick curve is density function of $\mathrm{BM}_{0.31,0.28,0.4}(x)$ for all $x \in \mathbb{R}$.

some bins of histogram outside the thick curve. Similarly, the thick curve fits to histogram (B) of Beta matrix. Following 8, the variance of the first dataset are close to 1 and the the mean is very small which is different with our assumption in Theorem 3.6 and Theorem 3.6. However, the outlier is not as much the box plots in Figure 6 (C)-(D). By this, the histogram of Fisher and Beta matrix may be suitable from our distribution function of F_{c_1,c_2} and BM_{s,c_1,c_2}.

REFERENCES

[1] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *Ann. Math.*, vol. 62, no. 3, pp. 548–564, 1955. `https://www.jstor.org/stable/1970079`.

[2] E. P. Wigner, “On the distribution of the roots of certain symmetric matrices,” *Ann. Math.*, vol. 67, no. 2, pp. 325–327, 1958. https://www.jstor.org/stable/1970008.

[3] J. Yao, S. Zheng, and Z. D. Bai, *Sample covariance matrices and high-dimensional data analysis*. New York: Cambridge University Press, 2015. https://doi.org/10.1017/CBO9781107588080.

[4] Z. D. Bai, Y. Q. Yin, and P. R. Krishnaiah, “On the limiting empirical distribution function of the eigenvalues of a multivariate F matrix,” *Theory Probab. Its Appl.*, vol. 32, no. 3, pp. 490–500, 1988. https://doi.org/10.1137/1132067.

[5] W. Bryc, A. Dembo, and T. Jiang, “Spectral measure of large random Hankel, Markov and Toeplitz matrices,” *Ann. Probab.*, vol. 34, no. 1, pp. 1–38, 2006. https://doi.org/10.1214/009117905000000495.

[6] V. Kargin, “Spectrum of random Toeplitz matrices with band structure,” *Electron. Commun. Probab.*, vol. 14, pp. 412–423, 2009. https://doi.org/10.1214/ECP.v14-1492.

[7] D. Paul and A. Aue, “Random matrix theory in statistics: A review,” *J. Stat. Plan. Inference*, vol. 150, pp. 1–29, 2014. https://doi.org/10.1016/j.jspi.2013.09.005.

[8] V. A. Marčenko and L. Pastur, “Distribution of eigenvalues for some sets of random matrices,” *Math USSR Sb.*, vol. 1, pp. 457–483, 1967. https://doi.org/10.1070/SM1967v001n04ABEH001994.

[9] J. W. Silverstein, “The limiting eigenvalue distribution of a multivariate F matrix,” *SIAM J. Math. Anal.*, vol. 16, no. 3, pp. 641–646, 1985. https://doi.org/10.1137/0516047.

[10] Y. Q. Yin, “Limiting spectral distribution for a class of random matrices,” *J. Multivar. Anal.*, vol. 20, no. 1, pp. 50–68, 1986. https://doi.org/10.1016/0047-259X(86)90019-9.

[11] E. R. Berndt and N. E. Savin, “Estimation and hypothesis testing in singular equation systems with autoregressive disturbances,” *Econometrica*, pp. 937–957, 1975. https://doi.org/10.2307/1911336.

[12] E. J. Elton and M. J. Gruber, “Estimating the dependence structure of share prices–implications for portfolio selection,” *J. Finance.*, vol. 28, no. 5, pp. 1203–1232, 1973. https://doi.org/10.2307/2978758.

[13] R. Engle and B. Kelly, “Dynamic equicorrelation,” *J Bus Econ Stat.*, vol. 30, no. 2, pp. 212–228, 2012. https://doi.org/10.1080/07350015.2011.652048.

[14] H. F. Kaiser, “A measure of the average intercorrelation,” *Educ. Psychol. Meas.*, vol. 28, no. 2, pp. 245–247, 1968. https://doi.org/10.1177/001316446802800203.

[15] Y. Akama and A. Husnaqilati, “A dichotomous behavior of guttman-kaiser criterion from equi-correlated normal population,” *J. Indones. Math. Soc.*, vol. 28, no. 3, p. 272–303, 2022. https://jims-a.org/index.php/jimsa/article/view/1158.

[16] J. W. Silverstein, “Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices,” *J. Multivar. Anal.*, vol. 55, no. 2, pp. 331–339, 1995. https://doi.org/10.1006/jmva.1995.1083.

[17] Z. D. Bai, “Methodologies in spectral analysis of large dimensional random matrices, a review,” *Advances in statistics*, pp. 174–240, 1999. http://www.jstor.org/stable/24306605.

[18] R. A. Horn and C. R. Johnson, *Matrix analysis*. New York: Cambridge university press, 2nd ed., 2012. https://www.cambridge.org/jp/academic/subjects/mathematics/algebra/matrix-analysis-2nd-edition?format=PB&isbn=9780521548236.

[19] Z. Bai and W. Zhou, “Large sample covariance matrices without independence structures in columns,” *Stat. Sin.*, pp. 425–442, 2008. https://www.jstor.org/stable/24308489.

[20] J. Hui and G. Pan, “Limiting spectral distribution for large sample covariance matrices with m-dependent elements,” *Commun. Stat. - Theory Methods*, vol. 39, no. 6, pp. 935–941, 2010. https://doi.org/doi.org/10.1080/03610920902807929.

[21] J. Bryson, R. Vershynin, and H. Zhao, “Marchenko–Pastur law with relaxed independence conditions,” *Random Matrices Theory Appl*, p. 2150040, 2021. https://doi.org/10.1142/S2010326321500404.

[22] F. Benaych-Georges and R. R. Nadakuditi, "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices," *Adv. Math.*, vol. 227, no. 1, pp. 494–521, 2011. https://doi.org/10.1016/j.aim.2011.02.007.

[23] Y. Akama, "Correlation matrix of equi-correlated normal population: Fluctuation of the largest eigenvalue, scaling of the bulk eigenvalues, and stock market," *Int. J. Theor. Appl. Finance*, vol. 26, no. 02n03, p. 2350006, 2023. https://doi.org/10.1142/S0219024923500061.

[24] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, vol. 20 of *Springer Series in Statistics*. New York: Springer, 2010. https://link.springer.com/book/10.1007/978-1-4419-0661-8.

[25] Z. Bai, J. Hu, G. Pan, and W. Zhou, "Convergence of the empirical spectral distribution function of Beta matrices," *Bernoulli*, vol. 21, no. 3, pp. 1538–1574, 2015. https://doi.org/10.3150/14-BEJ613.

[26] P. J. Huber and E. M. Ronchetti, *Robust Statistics*. Hoboken, NJ: John Wiley & Sons, 2011. https://doi.org/10.1002/9780470434697.

[27] Z. Bai, D. Jiang, J. Yao, and S. Zheng, "Corrections to LRT on large-dimensional covariance matrix by RMT," *Ann. Stat.*, vol. 37, no. 6B, pp. 3822–3840, 2009. https://doi.org/10.1214/09-AOS694.

[28] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. Hoboken: Wiley, 3rd ed., 2003.

[29] H. Yoshida, "Remarks on a free analogue of the beta prime distribution," *J. Theor. Probab.*, vol. 33, no. 3, pp. 1363–1400, 2020. https://doi.org/10.1007/s10959-019-00924-x.

[30] R. J. Muirhead, *Aspects of multivariate statistical theory*. Hoboken: John Wiley & Sons, 2009. https://doi.org/10.1002/9780470316559.

[31] M. L. Mehta, *Random matrices*. Amsterdam: Elsevier, 3rd ed., 2004. https://www.elsevier.com/books/random-matrices/lal-mehta/978-0-12-088409-4.

[32] C. Hammond and S. J. Miller, "Distribution of eigenvalues for the ensemble of real symmetric Toeplitz matrices," *J. Theor. Probab.*, vol. 18, no. 3, pp. 537–566, 2005. https://doi.org/10.1007/s10959-005-3518-5.

[33] J. Ramey, *Datamicroarray*. R package version 1.14.4., 2013.

[34] E. Gravier, G. Pierron, A. Vincent-Salomon, N. Gruel, V. Raynal, A. Savignoni, Y. De Rycke, J.-Y. Pierga, C. Lucchesi, F. Reyal, A. Fourquet, S. Roman-Roman, F. Radvanyi, X. Sastre-Garau, B. Asselain, and O. Delattre, "A prognostic dna signature for t1t2 node-negative breast cancer patients," *Genes Chromosomes Cancer*, vol. 49, no. 12, pp. 1125–1134, 2010. https://doi.org/10.1002/gcc.20820.