2025 S.T. Yau High School Science Award (Asia)

Research Report

The Team

Registration Number: Math-007

Name of team member: Yu, Yao-Hsing (尤耀星)
School: Municipal Chien Kuo High School

Name of supervising teacher: Wang, Shao-Hsuan
Job Title: Associate professor
School: Graduate Institute of Statistics, Central University

Name of supervising teacher: Lue, Yang-Kai
Job Title: Adjunct Assistant Professor
School: Department of Applied Mathematics,Yang Ming Chiao Tung University

Title of Research Report

Geometric Analysis of the Eigenvalue Range of the Generalized Covariance Matrix

Date

14 August 2025

Geometric Analysis of the Eigenvalue Range of the Generalized Covariance Matrix

Yu, Yao-Hsing (尤耀星)

Abstract

In classical random matrix theory, the limiting spectral distribution (LSD) of a sample covariance matrix can be derived explicitly via the Stieltjes transform. However, for generalized sample covariance matrices, no closed-form expression for the LSD is available, complicating efforts to analyze their spectral behavior. In this work, we employ a combination of geometric techniques and the Stieltjes transform to derive rigorous bounds on the support of the eigenvalue distribution for generalized covariance matrices. To assess the sharpness of our theoretical estimates, we conduct numerical simulations under various parameter settings and compare the observed eigenvalue ranges with our predicted bounds. The results demonstrate that our geometric transform approach yields tight approximations to the true spectral edge. These findings offer new insights into the asymptotic behavior of the generalized covariance matrix and provide practical guidelines for applications requiring precise eigenvalue information.

Keywords: Gaussian distribution, limiting spectral distribution (LSD), random matrix,

Stieltjes transform

Acknowledgement I gratefully acknowledge the support of Taipei Municipal Chien Kuo High School and the guidance of the supervising teachers.s

1 Introduction

In modern statistics and high-dimensional data analysis, the study of large random matrices has become increasingly important. When analyzing datasets with many variables relative to the number of observations, classical statistical methods often fail, and new theoretical frameworks are needed. Random matrix theory provides powerful tools for understanding the behavior of such high-dimensional systems. A fundamental object in multivariate statistics is the *data matrix*, which organizes n observations of p variables into a structured $p \times n$ matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, where each column $\mathbf{x}_i$ corresponds to a single observation and each row represents a variable. This representation facilitates computational efficiency and provides a natural framework for multivariate analysis techniques. As the dimensions of the data matrix grow large, the empirical properties of derived quantities, such as the sample covariance matrix

$$\mathbf{S}_n = \frac{1}{n}\mathbf{X}\mathbf{X}^*,$$

become increasingly important. Here, $\mathbf{X}^*$ denotes the conjugate transpose of $\mathbf{X}$. A cornerstone result in random matrix theory is the *Marčenko-Pastur law* [7], which describes the limiting spectral distribution of sample covariance matrices. Specifically, if the entries of $\mathbf{X}$ are independent and identically distributed (i.i.d.) with mean zero and variance one, and if the ratio $p/n \to y > 0$ as $n \to \infty$, then the empirical spectral distribution of $\mathbf{S}_n$ converges almost surely to a nonrandom limit whose density is given by the Marčenko-Pastur distribution.

An important line of work has been developed concerning *generalized sample covariance matrices*. Define

$$\mathbf{B}_n = \mathbf{S}_n\mathbf{T}_n,$$

where $\mathbf{S}_n$ denotes the sample covariance matrix, and $\{\mathbf{T}_n\}_{n=1}^{\infty}$ is a sequence of nonnegative definite Hermitian matrices. This formulation encompasses many important examples across scientific fields. For instance, the standard sample covariance matrix and Fisher matrices of the form $\mathbf{F}_n$ arise naturally in statistical theory [9]. Similarly, products of i.i.d. random matrices with this structure are commonly encountered in signal processing and wireless communications. To the best of our knowledge, the exact form of the limiting spectral distribution (LSD) for generalized sample covariance matrices is not known. Therefore, instead of focusing on the explicit form of the LSD, this study investigates the support of the limiting spectral behavior of such matrices. Building upon the existing results [3, 4, 10], which provide implicit characterizations via the Stieltjes transform, we develop a geometrical approach to determine the support of the LSD.

Our main contributions are as follows:

1. We derive explicit bounds for the support of the limiting distribution when the limiting population measure H has a two-mass-point support structure of the form $H = \beta\delta_a + (1-\beta)\delta_1$.

2. We apply a geometrical technique by transforming the problem of solving the Stieltjes transform into analyzing the number of intersection points between two curves.

2 Background

Let $\mathbf{M}$ be a $p\times p$ Hermitian matrix with eigenvalues $\lambda_1,\ldots,\lambda_p$. The *empirical spectral distribution* (ESD) of $\mathbf{M}$ is defined as

$$F^{\mathbf{M}} = \frac{1}{p}\sum_{i=1}^{p}\delta_{\lambda_i},$$

where δ_{λ_i} is the Dirac delta measure at λ_i. Equivalently, the cumulative distribution function is given by

$$\mu(x) = F^{\mathbf{M}}[0,x] = \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}_{\lambda_i\le x},$$

where $\mathbf{1}_{\lambda_i\le x}$ is the indicator function. The *Stieltjes transform* of a probability measure μ on $\mathbb{R}$ is defined as

$$s_\mu(z) = \int_{\mathbb{R}}\frac{1}{x-z}\,d\mu(x), \quad z\in\mathbb{C}^+,$$

where $\mathbb{C}^+ = \{z\in\mathbb{C} : \Im(z) > 0\}$ is the upper half-plane.

2.1 Fundamental theory

We begin with two fundamental results from random matrix theory. The first extends the classical Marčenko–Pastur law to products of sample covariance matrices with deterministic sequences. The second is an inverse formula that allows us to derive the corresponding probability measure via the Stieltjes transform.

Theorem 2.1 (Generalized sample covariance matrix theorem [4, 10])**.** *Let $\mathbf{S}_n$ be the sample covariance matrix defined by $\mathbf{S}_n = \frac{1}{n}\mathbf{X}\mathbf{X}^*$, where $\mathbf{X} = (\mathbf{x}_1,\ldots,\mathbf{x}_n)$ is a $p\times n$ data matrix. Let $\{\mathbf{T}_n\}$ be a sequence of nonnegative definite Hermitian matrices of size $p\times p$. Define a generalized sample covariance matrix*

$$\mathbf{B}_n = \mathbf{S}_n\mathbf{T}_n. \tag{1}$$

Assume the following conditions hold:

(i) *The entries (x_{jk}) of the data matrix are i.i.d. with mean zero and variance 1.*

(ii) *The dimension-to-sample ratio satisfies $\frac{p}{n}\to y>0$ as $n\to\infty$.*

(iii) *The sequence $\{\mathbf{T}_n\}$ is either deterministic or independent of $\mathbf{S}_n$.*

(iv) *Almost surely, the empirical spectral distribution $H_n = F^{\mathbf{T}_n}$ of $\mathbf{T}_n$ converges weakly to a nonrandom probability measure H.*

Then almost surely, $F^{\mathbf{B}_n}$ converges weakly to a nonrandom probability measure $F_{y,H}$. Moreover, the Stieltjes transform $s(z)$ of $F_{y,H}$ satisfies the implicit equation

$$s(z) = \int\frac{1}{t(1-y-yzs(z))-z}\,dH(t), \quad z\in\mathbb{C}^+. \tag{2}$$

Remark 2.2. *The equation (2) uniquely determines the Stieltjes transform $s(z)$ in the upper half-plane, and hence uniquely determines the limiting distribution $F_{y,H}$.*

Proposition 2.3 (Inversion Formula [4, 6]). *Let μ be a probability measure on $\mathbb{R}$ with Stieltjes transform $s_\mu(z)$. Then for all continuous and compactly supported functions $\varphi : \mathbb{R} \to \mathbb{R}$,*

$$\int_{\mathbb{R}} \varphi(x)\,\mu(dx) = \lim_{v \downarrow 0} \frac{1}{\pi} \int_{\mathbb{R}} \varphi(x)\,\Im(s_\mu(x+iv))\,dx.$$

In particular, for continuity points $a < b$ of μ,

$$\mu([a,b]) = \lim_{v \downarrow 0} \frac{1}{\pi} \int_a^b \Im(s_\mu(x+iv))\,dx.$$

2.2 Applications of theory

Theorem 2.1 can be applied through several classical examples from random matrix theory and multivariate statistics.

Example 2.4 (Classical Marčenko-Pastur law). *If $\mathbf{T}_n = \mathbf{I}_p$ (the $p \times p$ identity matrix), then $\mathbf{B}_n = \mathbf{S}_n$. In this case, $H = \delta_1$ (the Dirac measure at 1), and Theorem 2.1 reduces to the classical Marčenko-Pastur law. The limiting distribution has support on $[(1-\sqrt{y})^2, (1+\sqrt{y})^2]$ when $y \le 1$, and has an additional point mass at zero when $y > 1$.*

Example 2.5 (Fisher matrix). *In multivariate analysis, the Fisher matrix is defined as $\mathbf{F}_n = \mathbf{S}_{1,n}\mathbf{S}_{2,n}^{-1}$, where $\mathbf{S}_{1,n}$ and $\mathbf{S}_{2,n}$ are independent sample covariance matrices. This can be written in the form $\mathbf{F}_n = \mathbf{S}_{1,n}\mathbf{T}_n$ with $\mathbf{T}_n = \mathbf{S}_{2,n}^{-1}$.*

Under the conditions of Theorem 2.1, and assuming that the spectral distribution of $\mathbf{T}_n$ converges appropriately, the asymptotic behavior of $\mathbf{F}_n$ can be analyzed using this framework. The limiting spectral distribution of Fisher matrices has been studied extensively by Zhang, Bai, and Hu [9].

Example 2.6 (Product of i.i.d. sample covariance matrices). *Consider the case where $\mathbf{T}_n$ is an independent copy of the sample covariance matrix, denoted $\mathbf{S}_{2,n}$. We define*

$$\mathbf{B}_n = \mathbf{S}_{1,n}\mathbf{S}_{2,n},$$

where $\mathbf{S}_{1,n}$ and $\mathbf{S}_{2,n}$ are independent sample covariance matrices. Such products arise de-noise technique in signal processing, especially in high dimension and low sample size data ([5, 8]). Under the framework of Theorem 2.1, assuming appropriate convergence of the spectral distributions of both matrices, one can analyze the asymptotic spectral properties of $\mathbf{B}_n$. The limiting spectral distribution is characterized by the implicit equation (2) with H being the Marčenko-Pastur distribution. The limiting spectral distribution of product matrices has been studied by [2].

3 Explicit support bounds via a geometric method

Building upon the convergence result in Theorem 2.1, we now present our main method: Geometric technique to derive explicit bounds for the support of the LSD in a special case. While

Theorem 2.1 provides the existence of the limiting distribution and its implicit characterization via the Stieltjes transform, it does not give explicit information about the support of this distribution. Our work fills this gap for an important special case.

3.1 Coarse Bounds for the Support

We consider a specific case where the limiting measure H has a simple two-point structure, allowing us to derive explicit and computable bounds for the support of the limiting distribution. Suppose the measure H is specified as

$$H = \beta\delta_a + (1-\beta)\delta_1,$$

where $a \geq 1$, $y \geq 1$, and $0 \leq \beta \leq 1$.

Theorem 3.1. *Suppose that the conditions of Theorem 2.1 hold. Then the support of $F_{y,H}$ is contained in the interval*

$$\left[\max_{t\in(0,\infty)} g(t), \min_{t\in(-1/a,0)} g(t)\right],$$

where

$$g(t) = \frac{y\beta(a-1)t + (at+1)((y-1)t-1)}{(at+1)(t^2+t)}.$$

Proof. We start with an alternative formulation involving the matrix $\underline{\mathbf{B}}_n = \frac{1}{n}\mathbf{X}^*\mathbf{T}_n\mathbf{X}$ of size $n \times n$. The matrices $\mathbf{B}_n$ and $\underline{\mathbf{B}}_n$ share the same nonzero eigenvalues, so their ESDs satisfy

$$nF^{\underline{\mathbf{B}}_n} - pF^{\mathbf{B}_n} = (n-p)\delta_0.$$

When $p/n \to y > 0$, this gives the relation

$$\underline{F}_{y,H} - yF_{y,H} = (1-y)\delta_0,$$

where $\underline{F}_{y,H}$ is the limiting distribution of $\underline{\mathbf{B}}_n$. The corresponding Stieltjes transforms $\underline{s}(z)$ and $s(z)$ are related by

$$\underline{s}(z) = -\frac{1-y}{z} + ys(z).$$

Substituting this relation into the equation for $s(z)$, we obtain the alternative formulation

$$\underline{s}(z) = -\left(z - y\int \frac{t}{1+t\underline{s}(z)}\, dH(t)\right)^{-1}.$$

Solving for z yields

$$z = -\frac{1}{\underline{s}(z)} + y\int \frac{t}{1+t\underline{s}(z)}\, dH(t). \tag{3}$$

We utilize the alternative formulation (3) to determine the support of the LSD. For the measure $H = \beta\delta_a + (1-\beta)\delta_1$, equation (3) becomes

$$z = -\frac{1}{\underline{s}} + y\int \frac{t}{1+t\underline{s}}\, dH(t) = -\frac{1}{\underline{s}} + y\left(\frac{\beta a}{1+a\underline{s}} + \frac{1-\beta}{1+\underline{s}}\right). \tag{4}$$

By rewriting (4), we obtain the following cubic polynomial of $\underline{s}$:

$$az\underline{s}^3 + \left(a(z-y+1)+z\right)\underline{s}^2 + \left(a+z-y+1-y\beta\,(a-1)\right)\underline{s}+1=0.$$

We define a real-valued function as below.

$$f(t) = azt^3 + \left(a(z-y+1)+z\right)t^2 + \left(a+z-y+1-y\beta\,(a-1)\right)t+1. \tag{5}$$

In view of Proposition 2.3, it is necessary to analyze the imaginary part of $\underline{s}(z)$ in (3.1). However, the exact support of the limiting spectral distribution cannot be readily inferred from $\underline{s}(z)$, even when an explicit expression for $\underline{s}(z)$ is available. Instead, we determine the values of z for which $\underline{s}(z)$ has a nonzero imaginary part. Our objective is to characterize the set of $z \in \mathbb{R}$ for which the equation $f(t) = 0$ admits a unique singular real solution. Conversely, any value of z for which $f(t) = 0$ has three real roots lies outside the support of the LSD. Consider

$$\begin{cases} v = g(t) & := \frac{y\beta(a-1)t+(at+1)\{(y-1)t-1\}}{(at+1)(t^2+t)}. \\ v = h(t) & := z, \quad z \geq 0. \end{cases} \tag{6}$$

Given z, $v = h(t)$ is a horizontal line. When $v = g(t)$ and $v = h(t)$ have only one point of intersection, this indicates that $f(t) = 0$ has only one real root. Therefore, we transform the solve-equation problem into a geometric problem of curve intersection. As shown in Figure 1, the left diagram has three intersection points, indicating that z is not in the support of the LSD. The right diagram has one intersection point, indicating that z may be in the support of the LSD.

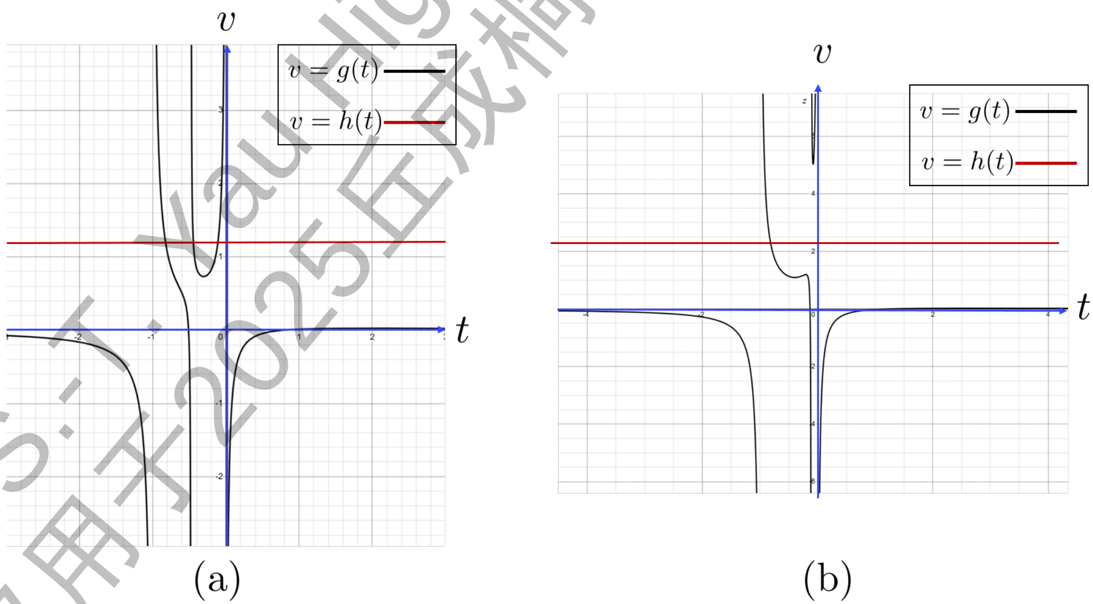

Figure 1: The graphs below depict the functions $v = g(t)$ and $v = h(t)$. The left panel corresponds to the parameter setting $(a, y, \beta) = (2.1, 1.7, 0.03)$. while the right panel uses $(a, y, \beta) = (10, 1.7, 0.03)$.

Now, we analyze the behavior of the curve $g(t)$ by directly differentiating it, as follows:

$$g'(t) = \frac{1}{t^2} - \frac{a^2 y\beta}{(at+1)^2} - \frac{y(1-\beta)}{(t+1)^2} = \frac{P_4(t)}{t^2(t+1)^2(at+1)^2}, \tag{7}$$

$$g''(t) = -\frac{2}{t^3} + \frac{2a^3 y\beta}{(at+1)^3} + \frac{2y(1-\beta)}{(t+1)^3}, \tag{8}$$

where a quartic polynomial

$$P_4(t) = (at+1)^2(t+1)^2 - a^2 y\beta\, t^2(t+1)^2 - y(1-\beta)\, t^2(at+1)^2.$$

We examine the curve of $g(t)$ across four disjoint regions of the real line: (a) $t < -1$, (b) $t \in (-1, -1/a)$, (c) $t \in (-1/a, 0)$, and (d) $t > 0$. Detailed analyses for regions (a), (c), and (d) are presented first, implies the result of Theorem 3.1. The discussion for region (b) will be presented in next section.

(a) $t < -1$: Define

$$J(t) = -\left(\frac{a^2 y\beta t^2}{(at+1)^2} + \frac{y(1-\beta)t^2}{(t+1)^2}\right).$$

We have

$$g'(t) = \frac{1}{t^2}\left(1 - \frac{a^2 y\beta t^2}{(at+1)^2} - \frac{y(1-\beta)t^2}{(t+1)^2}\right) = \frac{1}{t^2}\{1 + J(t)\}. \tag{9}$$

Note that

$$J'(t) = -2yt\left(\frac{2a^2\beta}{(at+1)^3} + \frac{(1-\beta)}{(t+1)^3}\right).$$

Since $t < -1$, we have $at+1 < 0$ and $t+1 < 0$ so $J'(t) < 0$ for $t < -1$. Thus, $J(t)$ is strictly decreasing on $(-\infty, -1)$. Moreover, $\lim_{t\to-\infty} J(t) = -y$. Therefore, $J(t) < -y$ for all $t < -1$. From (9), we have

$$g'(t) < \frac{1}{t^2}(1-y) < 0.$$

Hence, g is strictly decreasing on $(-\infty, -1)$ and admits no local extrema in this interval.

(c) $t \in (-1/a, 0)$: Clearly, we have $a > 1$, $t+1 > 0$, $-t > 0$, and $at+1 > 0$. Then,

$$g''(t) = \frac{2}{(-t)^3} + 2y\beta\frac{(at+a)^3 - (at+1)^3}{(at+1)^3(t+1)^3} + \frac{2y}{(t+1)^3} > 0.$$

Therefore, $g'(t)$ is strictly increasing. Moreover, since

$$\lim_{t\to 0^-} g'(t) = \infty \quad \text{and} \quad \lim_{t\to(-\frac{1}{a})^+} g'(t) = -\infty,$$

there exists exactly one root of the equation $g'(t) = 0$ in the interval $(-1/a,\, 0)$, and thus $g(t)$ has a unique local minimum in this interval.

(d) $t > 0$: For the interval $(0, \infty)$, we define $I(t) = g'(t)\,(at+1)^2$. Then for $t > 0$,

$$I(t) \;=\; \frac{(at+1)^2}{t^2} - a^2 y\beta \;-\; y(1-\beta)\frac{(at+1)^2}{(t+1)^2}$$

and

$$I'(t) \;=\; -\frac{2(at+1)}{t^3} - \frac{2y(1-\beta)(a-1)(at+1)}{(t+1)^3}. \tag{10}$$

We have $I'(t) < 0$ for $t > 0$. In this case, I is strictly decreasing and $(at+1)^2$ is strictly increasing on $(0, \infty)$. Consequently, $I'(t)$ is also strictly decreasing on this interval. Since

$$\lim_{t \to 0^+} I(t) = \infty \quad \text{and} \quad \lim_{t \to \infty} I(t) = a^2(1-y) < 0, \tag{11}$$

there exists a unique $b > 0$ with $I(b) = 0$, i.e. $g'(b) = 0$. Thus g has a unique local maximum on $(0, \infty)$.

Based on (a), (c), and (d), the curves $v = g(t)$ and $v = h(t)$ may have one point. Thus,

$$\left[\max_{t\in(0,\infty)} g(t), \quad \min_{t\in(-1/a,0)} g(t)\right]$$

include the support of the LSD. The proof is done. □

3.2 Explicit Bounds for the Support

To derive the exact bounds for the support of LSD, we will complete the part (b) in Section 3.1.

(b) $t \in (-1, -1/a)$: Note that $P_4(-1) < 0$, $P_4(-1/a) < 0$, and $P_4(0) > 0$ in (7). It follows that $P_4(t)$ and $g'(t)$ share the same set of roots. From the analyses in Section 3.1, we conclude that the number of real roots of $g'(t)$ is either two or zero. Let

$$P_4(t) = c_4 t^4 + c_3 t^3 + c_2 t^2 + c_1 t + c_0, \tag{12}$$

where $c_4 = a^2(1-y)$, $c_3 = 2a^2(1-y\beta) + 2a\{1-y(1-\beta)\}$, $c_2 = a^2(1-y\beta) + 4a + \{1-y(1-\beta)\}$, $c_1 = 2a+2$, and $c_0 = 1$. Based a modified version of the Ferrari's method, we apply the Tian Heng (天珩) formulae For more information, visit https://zhuanlan.zhihu.com/p/677634589. Let

$$\begin{aligned}
D &= 3c_3^2 - 8c_4c_2,\\
E &= -c_3^3 + 4c_4c_3c_2 - 8c_4^2c_1,\\
F &= 3c_3^4 + 16c_4^2c_2^2 - 16c_4c_3^2c_2 + 16c_4^2c_3c_1 - 64c_4^3c_0,\\
A &= D^2 - 3F,\ B = DF - 9E^2,\ C = F^2 - 3DE^2,\ \Delta = B^2 - 4AC.
\end{aligned} \tag{13}$$

Assume that x_1, x_2, x_3, x_4 are the four roots of $P_4(t)$ (note that they are also the roots of $g'(t)$).

Therefore, from (a), (c), (d) in Section 3.1, we have the following cases

Case 1. When $ABC \neq 0$, $\Delta = 0$, $AB > 0$, the equation has one pair of repeated real roots and two other distinct real roots. That is,

$$\begin{cases} x_1 = x_2 = \frac{1}{4c_4}\left(-c_3 - \frac{2AE}{B}\right) \in (-1, -1/a) \\ x_3 = \frac{1}{4c_4}\left(-c_3 + \frac{2AE}{B} - \sqrt{\frac{2B}{A}}\right) \in (-1/a, 0) \\ x_4 = \frac{1}{4c_4}\left(-c_3 + \frac{2AE}{B} + \sqrt{\frac{2B}{A}}\right) \in (0, \infty) \end{cases} \tag{14}$$

It implies that $g(t)$ is monotone in $(-1, -1/a)$ so that the curves $v = g(t)$ and $v = h(t)$ has exactly one point when $v \in (g(x_4),\ g(x_3))$. Thus, it indicates that $[g(x_4),\ g(x_3)]$ is the support of the LSD.

Case 2. When $\Delta < 0, D > 0, F > 0$, the equation has four distinct real roots. Let $x_{i,j} = u \pm w$ denote $x_i = u + w$ and $x_j = u - w$ for simplicity. We have

$$\begin{cases} x_{1,2} = \frac{1}{4c_4}\left\{-c_3 + \mathrm{sgn}(E)\sqrt{y_1} \pm (\sqrt{y_2} + \sqrt{y_3})\right\} \\ x_{3,4} = \frac{1}{4c_4}\left\{-c_3 - \mathrm{sgn}(E)\sqrt{y_1} \pm (\sqrt{y_2} - \sqrt{y_3})\right\}, \end{cases} \tag{15}$$

where

$$y_1 = \frac{1}{3}\left(D - 2\sqrt{A}\cos\left(\frac{\theta}{3}\right)\right) \quad \text{and} \quad y_{2,3} = \frac{1}{3}\left(D - 2\sqrt{A}\cos\left(\frac{\theta}{3} \pm \frac{2\pi}{3}\right)\right)$$

with

$$\theta = \arccos\left(\frac{3B - 2AD}{2A\sqrt{A}}\right).$$

Let $x_{(1)} \leq x_{(2)} \leq x_{(3)} \leq x_{(4)}$ denote the ordered (increasing) values of x_1, x_2, x_3, x_4. From (a), (c), (d), $x_{(1)}, x_{(2)} \in (-1, -1/a)$, $x_{(3)} \in (-1/a, 0)$, and $x_{(4)} \in (0, \infty)$. Thus, the curves $v = g(t)$ and $v = h(t)$ has exactly one point when $v \in (g(x_{(4)}), g(x_{(1)}))$ or $v \in (g(x_{(2)}), g(x_{(3)}))$. Thus, it indicates that the set $[g(x_{(4)}), g(x_{(1)})] \bigcup [g(x_{(2)}), g(x_{(3)})]$ is the support of the LSD, as described in Figure 1 (b).

Case 3. When $\Delta > 0$, the equation has two distinct real roots. We have

$$x_{1,2} = \frac{1}{4c_3}\left(-c_3 + \mathrm{sgn}(E)\sqrt{\frac{D + u_1^{1/3} + u_2^{1/3}}{3}} \pm \sqrt{\frac{2D - (u_1^{1/3} + u_2^{1/3}) + 2\sqrt{u_3}}{3}}\right), \tag{16}$$

where

$$u_{1,2} = AD + 3\left(\frac{-B \pm \sqrt{\Delta}}{2}\right), \quad u_3 = D^2 - D(u_1^{1/3} + u_2^{1/3}) + (u_1^{1/3} + u_2^{1/3})^2 - 3A.$$

Let $x_{(1)} \leq x_{(2)}$ denote the ordered (increasing) values of x_1, x_2. In this case, the curves

$v = g(t)$ and $v = h(t)$ has exactly one point when $v \in (g(x_{(2)}),\ g(x_{(1)}))$. Thus, it indicates that $[g(x_{(2)}),\ g(x_{(1)})]$ is the support of the LSD, as described in Figure 1 (a).

4 Numerical study

In this section, we present a numerical study to evaluate the performance of the proposed method under various parameter settings. The code is available : https://github.com/yu314-coder and https://huggingface.co/spaces/euler314.

Figure 2 illustrates the setting $y = 10, a = 6$ with varying $\beta \in (0, 1)$. First, we apply the eigen-decomposition technique to compute the bounds of eigenvalues of the generalized sample covariance matrix, denoted as simulated bounds. We plot two curves corresponding to the estimated bounds from Theorem 3.1 and the simulated bounds with varing $\beta \in (0, 1)$, respectively. The close alignment of the two curves indicates that the estimated bounds perform well.

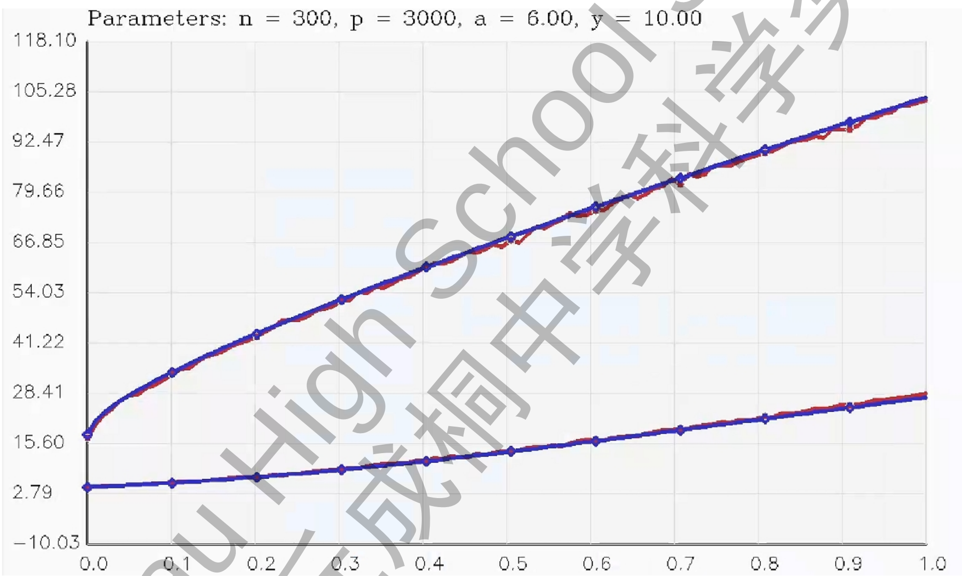

Figure 2: The graphs above compare the estimated bounds (blue line) from Theorem 3.1 with simulated bounds (red line), under the setting $y = 10, a = 6$ with varing $\beta \in (0, 1)$.

From Cases 2 and 3 in Section 3.2, the criterion Δ in (13) can be used to distinguish between two situations: (1) two disjoint intervals, and (2) a single interval. In each case, the corresponding estimated bounds are provided. To further demonstrate the effectiveness of our proposed method in Section 3.2, we present two settings:

Setting 1. The spectral distribution for $\beta = 0.05, a = 10, n = 1000, p = 1700$ (i.e. $y = 1.7$) reveal two disjoint intervals.

Setting 2. The spectral distribution for $\beta = 0.5, a = 10, n = 1000, p = 1700$ (i.e. $y = 1.7$) form a single interval.

Under **Setting 1**, the criterion $\Delta \approx -8.38 \times 10^{25} < 0$ is satisfied, implying the support of the LSD is formed by two disjoint intervals. The *Coarse Method* provides bounds as established in Theorem 3.1 and discussed in Section 3.1. The *Explicit Method* offers explicit estimates for the bounds $\{g(x_{(j)})\}$s, introduced in Section 3.2. The *Simulated Method* refers to directly computing eigenvalues via eigen-decomposition of a sample covariance matrix. Table 1 shows that the estimates of both the coarse and explicit methods are close to those obtained from the Simulated Method. Figure 3 showed that the estimated bounds in *Explicit Method* are very close to the simulated bounded. For **Setting 2**, the criterion $\Delta \approx 2.14 \times 10^{26} > 0$ is satisfied, implying the support of the LSD is an interval. Table 2 shows that the estimates from both the Coarse and Explicit Methods are close to those obtained from the Simulated Method. Figure 4 showed that the estimated bounds in *Explicit Method* are very close to the simulated bounded.

Method	Interval 1		Interval 2	
Simulated Method	[0.1019	4.9566]	[7.0041	18.4447]
Coarse Method	[0.0979	]	[	18.4314]
Explicit Method	[0.0979	4.8998]	[6.9009	18.4314]

Table 1: Comparison of three methods for estimating the bounds of the support of the spectral distribution under **Setting 1**. The *Coarse Method* provides coarse estimated bounds as described in Section 3.1, while the *Explicit Method* presents explicit estimated bounds introduced in Section 3.2.

Method	Interval	
Simulated Method	[0.2107	37.3610]
Coarse Method	[0.2075	37.8357]
Explicit Method	[0.2075	37.8357]

Table 2: Comparison of three methods for estimating the bounds of the support of the spectral distribution under **Setting 2**. The *Coarse Method* provides coarse estimated bounds as described in Section 3.1, while the *Explicit Method* presents explicit estimated bounds introduced in Section 3.2.

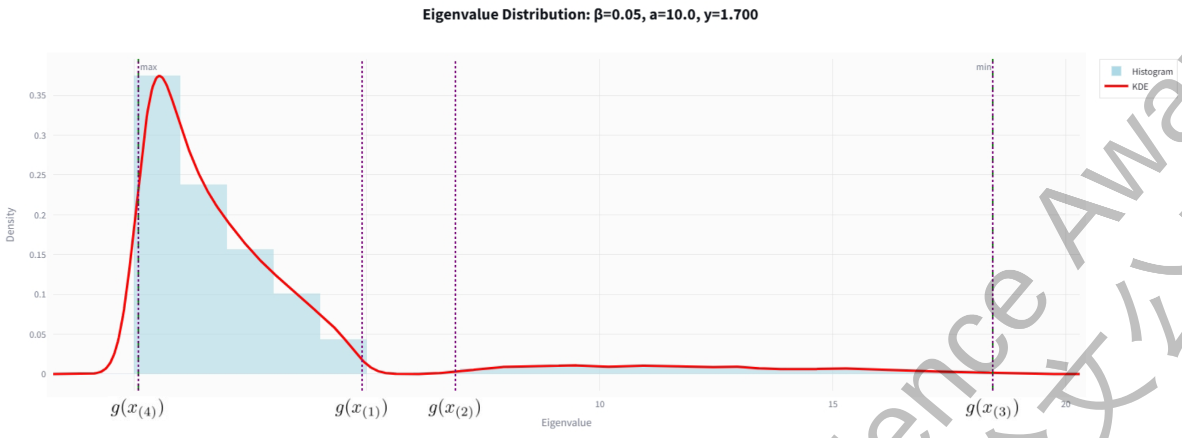

Figure 3: The histogram and kernel density estimate (KDE) of the spectral distribution for $\beta = 0.05, a = 10, y = 1.7$ reveal two disjoint intervals. In this case, the criterion $\Delta \approx -8.38 \times 10^{25} < 0$ is satisfied, and $g(x_{(1)}), \ldots, g(x_{(4)})$ represent the four estimated bounds corresponding to Case 2 in Section 3.2.

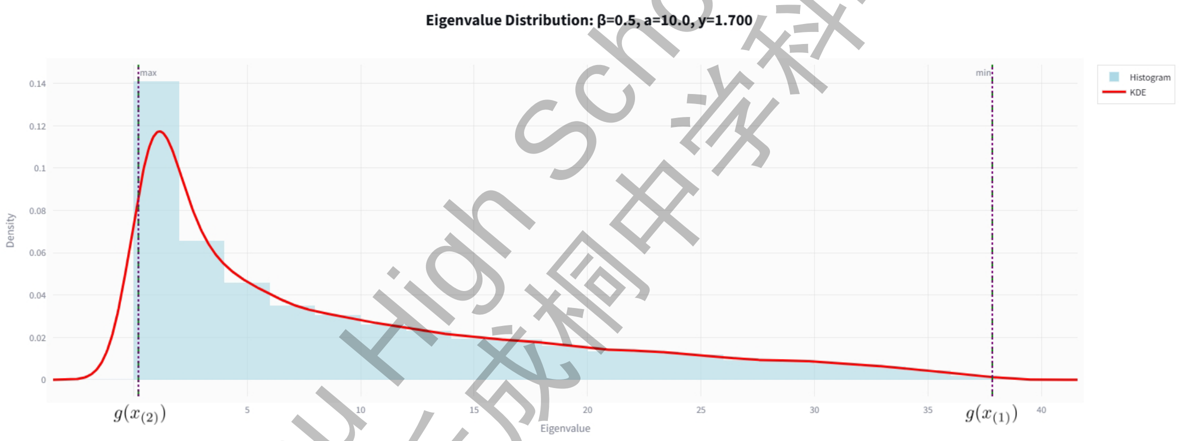

Figure 4: The histogram and kernel density estimate (KDE) of the spectral distribution for $\beta = 0.5, a = 10, y = 1.7$ reveal two disjoint intervals. In this case, the criterion $\Delta \approx 2.14 \times 10^{26} > 0$ is satisfied, and $g(x_{(1)})$ and $g(x_{(2)})$ represent the two estimated bounds corresponding to Case 3 in Section 3.2.

5 Conclusion

Based on existing convergence results for the limiting spectral distributions of a generalized sample covariance matrices, we have derived explicit bounds for the support of the LSD in the important special case of a two-mass-point limiting measure. Our main theorem provides concrete upper and lower bounds that can be computed through optimization problems involving elementary functions. By geometric properties of quadratic functions, precise spectral bounds of the LSD can be obtained. The explicit nature of our bounds makes them particularly valuable for computational applications and numerical verification. Future research directions include extending the bounds to support structures for more general limiting measure H, investigating the sharpness of the derived bounds, and developing efficient numerical algorithms for computing the extremal values in the optimization problems that define the support boundaries.

References

[1] D. Voiculescu. Limit laws for random matrices and free products. *Inventiones Mathematicae*, 104(1):201–220, 1991.

[2] H. Hung and C. C. Yeh, S. Y. Huang On the asymptotic properties of product-PCA under the high-dimensional setting. *arXiv*, 2024.

[3] J. W. Silverstein. Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *Journal of Multivariate Analysis*, 55(2):331–339, 1995.

[4] J. Yao and S. Zheng and Z. Bai Sample covariance matrices and high-dimensional data analysis. *Cambridge UP, New York*, 2015.

[5] K. Yata and M. Aoshima. Effective PCA for high-dimension, low-sample-size data with singular value decomposition of cross data matrix. *Journal of multivariate analysis*, 101(9), 2060-2077.

[6] N. I. Akhiezer. *The Classical Moment Problem and Some Related Questions in Analysis*. Oliver and Boyd, Edinburgh, 1965.

[7] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457–483, 1967.

[8] S. H. Wang, and S. Y. Huang Perturbation theory for cross data matrix-based PCA. *Journal of Multivariate Analysis*, 190, 104960.

[9] X. Zhang and Z. Bai, and J. Hu, Limiting spectral distribution of high-dimensional non-central Fisher matrices and its analysis. *Sci. China Math*, 66: 393-408, 2023.

[10] Z. Bai and J. W. Silverstein. CLT for linear spectral statistics of large-dimensional sample covariance matrices. *The Annals of Probability*, 32(1A):553–605, 2004.