Limiting Spectral Distributions of Random Matrices under Finite-Rank Perturbations (MIT18.338 - Project Report)

Xiaomin Li
xiaominli@g.harvard.edu

Yi Tian
yitian@mit.edu

December 17, 2022

1 Introduction

Let X_n be an $n \times n$ random Hermitian matrix and P_n be an $n \times n$ random Hermitian matrix with rank $r < n$ with deterministic nonzero eigenvalues $\theta_1 \geq \ldots \geq \lambda_s > 0 > \lambda_{s+1} \geq \ldots \lambda_r$. We are interested in the limiting spectral distributions (LSDs) of $X_n + P_n$ and $(I_n + P_n)^{1/2} X_n (I_n + P_n)^{1/2}$, which are finite-rank perturbations of random matrices.

A special case of multiplicative perturbation is well studied in the literature. Consider the matrix $Z_n \in \mathbb{C}^{p \times n}$ with i.i.d. entries, under certain conditions on the moments of its entries, we know the spectral distribution of $\frac{1}{n} Z_n Z_n^*$ converge to the famous Marchenko-Pastur law. Johnstone introduced the "spiked population model" [Joh01], where the perturbed matrix $S_n := \frac{1}{n} T_p^{1/2} Z_n Z_n^* T_p^{1/2}$, and T_p is a $p \times p$ diagonal matrix with all diagonal entries being ones except for the first r diagonal entries. In this case, matrix S_n is the result of a rank-r perturbation of the original matrix $\frac{1}{n} Z_n^* Z_n$. Previous work [Sil95, SC95, SB95, BS98, BS99, BBAP05, BY12] have studied the limit spectral distribution (LSD) of S_n under the regime $p/n \to y$ as $n \to \infty$ for constant y. The assumptions on the matrix T_p and Z_n are slightly different in the papers by different authors. For example, T_p could be diagonal/nondiagonal, random/deterministic, and entries in Z_n could be i.i.d./non-i.i.d., etc. Nonetheless, the general results indicate that if T_p has a limiting distribution, then the spikes of the LSD of T_p can induce spikes of the LSD of S_n, with the phase transition described by the derivative of a certain function. Moreover, the transformation from the i-th spike of LSD of T_p to the i-th spike of LSD of S_n can be characterized by the same function for all the spikes.

Most relevant to our problem of interest is [BGN11], which considers both additive and multiplicative perturbations under the assumption that either the original matrix or the perturbation matrix is unitarily invariant, that is, the entries has the same distribution with multiplication by a unitary matrix. Its key result is again a phase transition for both LSD and eigenvectors in both additive and multiplicative cases. Concretely, it develops a unified framework for these two cases to fully characterize the extreme (maximum and minimum) spikes of the perturbed matrix as $n \to \infty$, with the additive case characterized by the Cauchy transform and the multiplicative case by the T-transform, both having tight connections with objects in free probability.

In this project, we closely follow [BGN11]. Our first goal is to confirm its theoretical results with experiments in Julia and to understand its proof techniques. [BGN11] considers only extreme spikes, while an interesting phenomenon is that the LSD of a random matrix has multiple bulks, and spikes between the bulks may arise with perturbation. Given that spikes may correspond to signals in the real world [Joh01], the behavior of the middle spikes is also worth studying. Here, we further study this phenomenon using the technical tools from [BGN11]. Using two-bulk Wigner matrices as an example, we illustrate the different situations where there can be no spike, extreme spikes, middle spikes, or both. We provide numerical experiments to verify the theoretical results.

2 Eigenvalue phase transition

In this section, we discuss the results in [BGN11] concerning the phase transition of the LSD of the perturbed matrices. Several definitions are needed to present the main result. Let $\lambda_1(X_n) \geq \ldots \geq \lambda_n(X_n)$ be the ordered eigenvalues of X_n. Define μ_{X_n} to be its empirical spectral distribution (ESD), by

$$\mu_{X_n} := \frac{1}{n}\sum\nolimits_{i=1}^{n} \delta_{\lambda_i(X_n)}.$$

Let μ_X be the corresponding LSD as $n \to \infty$. Let a and b be the infimum and supremum of the support of μ_X, respectively.

For a compactly supported probability measure μ, the Cauchy transform is defined as

$$G_\mu(z) := \int (z-t)^{-1} \mathrm{d}\mu(t) \text{ for } z \notin \operatorname{supp}(\mu).$$

The R-transform that linearizes free sum is simply $R_\mu(z) := G_\mu^{-1}(z) - z^{-1}$. The T-transform is defined as

$$T_\mu(z) := \int t(z-t)^{-1} \mathrm{d}\mu(t) \text{ for } z \notin \operatorname{supp}(\mu).$$

The S-transform that linearizes free multiplication is simply $S_\mu(z) := (1+z)/(zT_\mu^{-1}(z)$.

For a function $f : \mathbb{R} \to \mathbb{R}$, we let $f(c^+) := \lim_{z \downarrow c} f(z)$ and $f(c^-) := \lim_{z \uparrow c} f(z)$.

The main result for the LSD of $\widetilde{X}_n := X_n + P_n$ is as follows.

Theorem 1. *For each $1 \leq i \leq s$,*

$$\lambda_i(\widetilde{X}_n) \xrightarrow{\text{a.s.}} \begin{cases} G_{\mu_X}^{-1}(\theta_i^{-1}) & \text{if } \theta_i > (G_{\mu_X}(b^+))^{-1}, \\ b & \text{otherwise,} \end{cases}$$

while for each fixed $i > s$, $\lambda_i(\widehat{X}_n) \xrightarrow{\text{a.s.}} b$. Similarly, for the smallest eigenvalues, we have for each $0 \leq j < r - s$,

$$\lambda_{n-j}(\widetilde{X}_n) \xrightarrow{\text{a.s.}} \begin{cases} G_{\mu_X}^{-1}(\theta_{r-j}^{-1}) & \text{if } \theta_i < (G_{\mu_X}(a^-))^{-1}, \\ a & \text{otherwise,} \end{cases}$$

while for each fixed $j \geq r - s$, $\lambda_{n-j}(\widetilde{X}_n) \xrightarrow{\text{a.s.}} a$.

The theorem in the multiplicative case is the same as Theorem 1, except that the Cauchy transform G_{μ_X} is replaced by the T-transform T_{μ_X}.

2.1 Simulations

In this subsection, we provide numerical simulations to verify the characterizations above. Figure 1 shows the eigenvalue distribution of the Wigner matrix without perturbation, which converges to the Semicircle Law. Figures 2 and 3 plot the eigenvalues of the perturbed matrices, with different values of θ_i. We can see from the figures that when θ_i are small, no spikes are revealed, while when they are large enough, we see spikes outside of the bulk. Finally, Figure 4 verifies the correctness of the characterization of the spikes in Theorem 1.

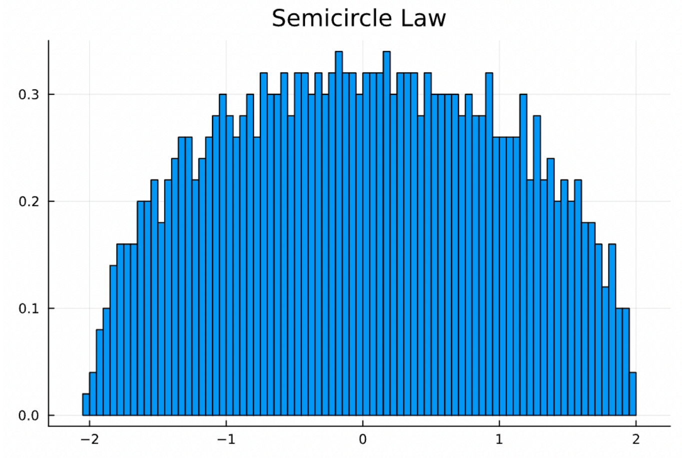


Figure 1

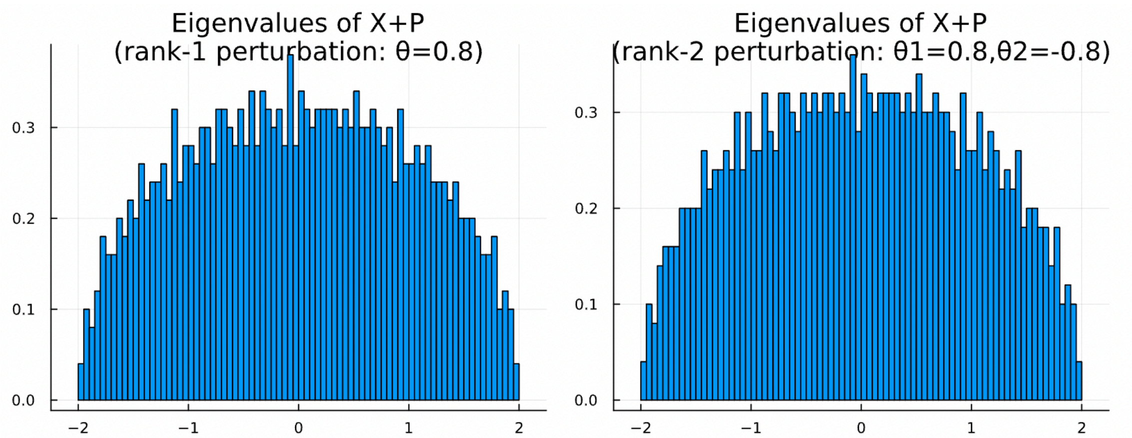


Figure 2

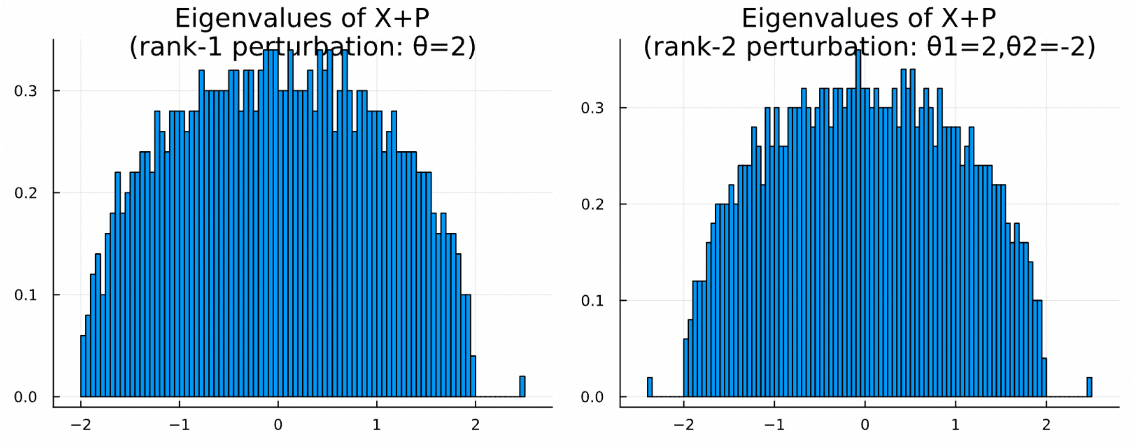


Figure 3

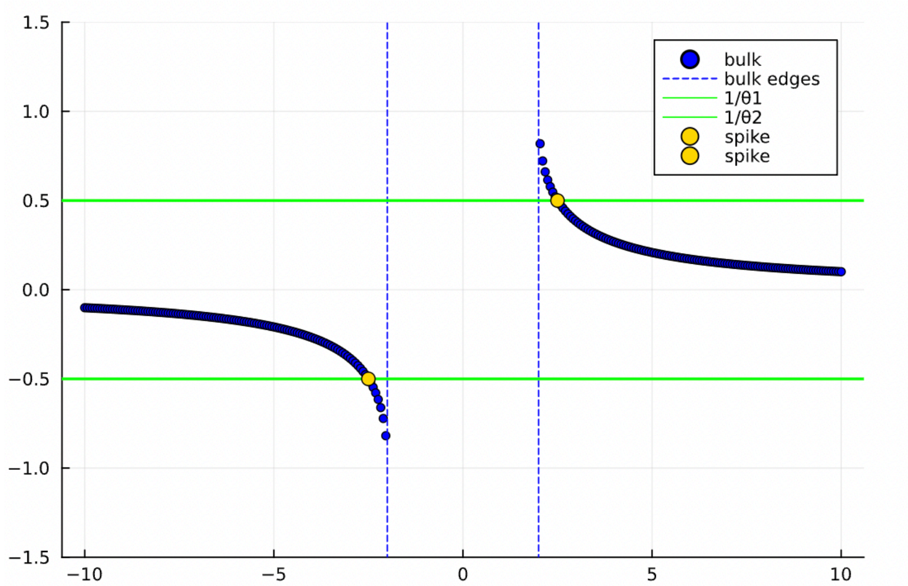


Figure 4

2.2 Outline of the proof of Theorem 1

Below we illustrate the main proof ideas for Theorem 1. The rigorous proof would rely on the concentration results in Section 4. We consider real symmetric matrices and use $A^\top$ to denote the transpose of matrix A. Since either random matrix X_n or P_n is unitarily (orthogonally) invariant,

$$\begin{aligned} X_n + P_n &= U_{X_n}\Lambda_{X_n}U_{X_n}^\top + U_{P_n}\Lambda_{P_n}U_{P_n}^\top \\ &= U_{X_n}(\Lambda_{X_n} + U_{X_n}^\top P_n U_{X_n})U_{X_n}^\top \\ &= U_{P_n}(U_{P_n}^\top X_n U_{P_n} + \Lambda_{P_n})U_{P_n}^\top. \end{aligned}$$

Without loss of generality, we can suppose $X_n = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. We shall consider rank 2 perturbation $P_n = \theta_1 u_n u_n^\top + \theta_2 v_n v_n^\top$, as they are general enough to introduce the ideas. The eigenvalues of $X_n + P_n$ are solutions to the equation

$$\det(zI - (X_n + P_n)) = 0.$$

If $zI - X_n$ is invertible, we have

$$zI - (X_n + P_n) = (zI - X_n)(I - (zI - X_n)^{-1}P_n).$$

Then,

$$\det(zI - (X_n + P_n)) = \det(zI - X_n) \cdot \det(I - (zI - X_n)^{-1}P_n).$$

This means z is an eigenvalue of $X_n + P_n$ and not an eigenvalue of X_n if and only if $\det(I - (zI - X_n)^{-1}P_n) = 0$. Moreover,

$$\begin{aligned}
&\det(I - (zI - X_n)^{-1}P_n) \\
=\ &\det(I - (zI - X_n)^{-1}(\theta_1 u_n u_n^\top + \theta_2 v_n v_n^\top)) \\
\overset{(i)}{=}\ &\det\left(I - [u_n, v_n]^\top \operatorname{diag}\left((z-\lambda_1)^{-1}, \dots, (z-\lambda_n)^{-1}\right)[\theta_1 u_n, \theta_2 v_n]\right) \\
=\ &\det\left(\begin{bmatrix} 1 - \theta_1 \sum_{i=1}^n u_i^2 (z-\lambda_i)^{-1} & \theta_2 \sum_{i=1}^n u_i v_i (z-\lambda_i)^{-1} \\ \theta_1 \sum_{i=1}^n u_i v_i (z-\lambda_i)^{-1} & 1 - \theta_2 \sum_{i=1}^n v_i^2 (z-\lambda_i)^{-1} \end{bmatrix}\right),
\end{aligned} \tag{1}$$

where (i) is due to Sylvester's determinant identity. By some concentration arguments [BGN11, Lemma A.2], in the large n limit,

$$\sum\nolimits_{i=1}^n u_i v_i (z-\lambda_i)^{-1} \overset{a.s.}{\to} 0, \quad \sum\nolimits_{i=1}^n u_i^2 (z-\lambda_i)^{-1} \overset{a.s.}{\to} n^{-1} \sum\nolimits_{i=1}^n (z-\lambda_i)^{-1}. \tag{2}$$

Setting $\det(I - (zI - X_n)^{-1}P_n)$ to be zero gives

$$1 - \theta_1 n^{-1} \sum\nolimits_{i=1}^n (z-\lambda_i)^{-1} = 0, \text{ or } 1 - \theta_2 n^{-1} \sum\nolimits_{i=1}^n (z-\lambda_i)^{-1} = 0.$$

By definition,

$$G_{\mu_X}(z) = \lim_{n\to\infty} n^{-1} \sum\nolimits_{i=1}^n (z-\lambda_i)^{-1}.$$

Hence, new limiting eigenvalues z of $X_n + P_n$ satisfy

$$G_{\mu_X}(z) = \theta_1^{-1}, \text{ or } G_{\mu_X}(z) = \theta_2^{-1}. \tag{3}$$

3 Eigenvector phase transition

[BGN11] also characterizes the eigenvectors associated with the leading and tailing spike eigenvalues. We will state the results informally in the case when P_n has rank $r = 1$ and provide the core ideas of the proof. The rigorous proof also relies on the concentration results in Section 4.

Theorem 2 (Norm of Eigenvector Projection). *Let $P_n = \theta u u^\top$. Let $\widetilde{\lambda}$ be the largest (smallest) eigenvalue of $\widetilde{X}_n$ if $\theta > 0$ ($\theta < 0$) and $\widetilde{u}$ be the eigenvector of $\widetilde{X}_n$ associated with $\widetilde{\lambda}$. When $\frac{1}{\theta} \in (G_{\mu_X}(a^-), G_{\mu_X}(b^+))$,*

$$\langle \widetilde{u}, u \rangle^2 \xrightarrow{a.s.} \frac{-1}{\theta^2 G'_{\mu_X}(\rho)}, \tag{4}$$

where $\rho \overset{\text{def}}{=} G_{\mu_X}^{-1}(\frac{1}{\theta})$ is the limit of $\widetilde{\lambda}$ as $n \to \infty$. When $\frac{1}{\theta} \notin (G_{\mu_X}(a^-), G_{\mu_X}(b^+))$, we assume G_{μ_X} has infinite derivative at a and b, then

$$\langle \widetilde{u}, u \rangle^2 \xrightarrow{a.s.} 0. \tag{5}$$

The main result in the multiplicative case is also similar, with the limit (4) given by $-(\theta^2 \rho T'_\mu(\rho) + \theta)^{-1}$, where $\rho = T_{\mu_X}^{-1}(\theta^{-1})$.

3.1 Outline of the proof of Theorem 2

Similar to the discussion of eigenvalues, we can decompose

$$X_n + P_n = U_{X_n}(\Lambda_{X_n} + U_{X_n}^\top U_{P_n}\Lambda_{P_n}U_{P_n}^\top U_{X_n})U_{X_n}^\top.$$

Hence for any eigenvector $\widetilde{u}$ of $\Lambda_{X_n} + U_{X_n}^\top U_{P_n}\Lambda_{P_n}U_{P_n}^\top U_{X_n}$, $U_{X_n}\widetilde{u}$ is an eigenvector of $\widetilde{X}_n$. Therefore, we could assume $X_n = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ without loss of generality.

By the definiton of $\widetilde{u}$, we have $\widetilde{X}_n\widetilde{u} = z\widetilde{u}$, which implies that $(\widetilde{\lambda}I - X_n)\widetilde{u} = P_n\widetilde{u} = \theta uu^\top\widetilde{u} = (\theta u^\top\widetilde{u})u$. Therefore, $\widetilde{u}$ is proportional to $(\widetilde{\lambda}I - X_n)^{-1}u$. By convention, let the eigenvectors have unit-norm. Then

$$\widetilde{u} = \frac{(\widetilde{\lambda}I - X_n)^{-1}u}{\|(\widetilde{\lambda}I - X_n)^{-1}u\|} = \frac{(\widetilde{\lambda}I - X_n)^{-1}u}{\sqrt{u^\top(\widetilde{\lambda}I - X_n)^{-2}u}} \tag{6}$$

and

$$\langle\widetilde{u}, u\rangle^2 = \frac{(u^\top(\widetilde{\lambda}I - X_n)^{-1}u)^2}{u^\top(\widetilde{\lambda}I - X_n)^{-2}u} = \frac{\left(\sum_{i=1}^n \frac{u_i^2}{(\widetilde{\lambda}-\lambda_i)}\right)^2}{\sum_{i=1}^n \frac{u_i^2}{(\widetilde{\lambda}-\lambda_i)^2}} \tag{7}$$

$$= \frac{(G_{\mu_n}(\widetilde{\lambda}))^2}{-G'_{\mu_n}(\widetilde{\lambda})} \approx \frac{(1/\theta)^2}{-G'_{\mu_X}(\rho)} = \frac{-1}{\theta^2 G'_{\mu_X}(\rho)} \tag{8}$$

It suffices to only discuss the case when $\theta > 0$. When $\frac{1}{\theta} < G_{\mu_X}(b^+)$, we know $\widetilde{\lambda}$ is a spike of $\widetilde{X}_n$ and $\widetilde{\lambda} \to \rho = G_{\mu_X}(\frac{1}{\theta}) > 0$. Then $\langle\widetilde{u}, u\rangle^2 \to \frac{-1}{\theta^2 G'_{\mu_X}(\rho)} > 0$. When $\frac{1}{\theta} \geq G_{\mu_X}(b^+)$, $\widetilde{\lambda}$ is hidden in the bulk and $\widetilde{\lambda} \to \rho = G_{\mu_X}(\frac{1}{\theta}) = b$. With further assumption that G_{μ_X} has infinite derivative at b, we have $\langle\widetilde{u}, u\rangle^2 \to \frac{-1}{\theta^2 G'_{\mu_X}(\rho)} \to 0$.

4 Concentration Arguments

The paper [BGN11] relies on concentration results to make all the arguments rigorous in the proof outlines above for both spike eigenvalues and the corresponding eigenvectors. In this section, we briefly include the main ideas of the rigorous proof of Theorem 1.

Step 1: Prove the extreme eigenvalues of $\widetilde{X}_n$ which does not tend to a limit in $\mathbb{R}$ $[a, b]$ tends to either a or b. That is, for any fixed $1 \leq i \leq s$, we have $\lambda_i(\widetilde{X}_n) \xrightarrow{a.s.} b$ and for any $1 \leq i \leq r - s$, $\lambda_{n-i}(\widetilde{X}_n) \xrightarrow{a.s.} a$. This step could be simply done by using Weyl's interlacing inequalities to bound the i-th (fixed i) eigenvalue of $\widetilde{X}_n$.

Step 2: Consider the perturbation matrix P_n with eigendecomposition $U_n\Theta_nU_n^\top$, where $\Theta = \operatorname{diag}(\theta_1, \ldots, \theta_r)$. Define

$$M_n(z) := I_r - U_n^\top(zI_n - X_n)^{-1}U_n\Theta_n. \tag{9}$$

The rank-2 special case of $M_n(z)$ is in (1). For $G_{\mu_X}(z)$ of the variable $z \in \mathbb{C} \setminus [a, b]$, define

$$M_{G_{\mu_X}}(z) := \operatorname{diag}(1 - \theta_1 G_{\mu_X}(z), \ldots, 1 - \theta_r G_{\mu_X}(z)). \tag{10}$$

The hardest part in the rigorous proof is to show that $M_n(z) \approx M_{G_{\mu_X}}(z)$ and the z's such that $M_n(z)$ is singular converge to the z's such that $M_{G_{\mu_X}}(z)$ is singular, which supports our approximation in (2). To prove this, the paper [BGN11] uses

Cauchy's argument principle to show that for any interval (c, d) outside of (a, b), the number of z's in (c, d) such that $M_n(z) = 0$ converges to number of $\widetilde{z}$'s in (c, d), where $\widetilde{z}$'s are the solutions to $M_{G_{\mu_X}}(\widetilde{z}) = 0$.

Step 3: Note that the $\widetilde{z}$'s such that $M_{G_{\mu_X}}(\widetilde{z}) = 0$, which are exactly the solutions to $G_{\mu_X}(\widetilde{z}) = \theta_i^{-1}$. Inverting G_{μ_X} gives the formula in (3).

Step 4: The last step is to deal with the case when θ_i have multiplicity. A perturbation argument (perturb to separate the repeated eigenvalues θ_i) makes it reduced to the simple-eigenvalue case.

5 Spikes between multiple bulks

[BGN11] considers the extreme eigenvalues. That is, suppose the smallest and largest eigenvalues of X_n converge to a and b, how perturbation P_n moves the eigenvalues beyond $[a, b]$. Here we study how perturbation creates new spikes between a and b. Suppose the support of μ_X is $[a_1, b_1]$ and $[a_2, b_2]$ where $b_1 < a_2$. The question is, under what condition we can observe spikes in (b_1, a_2) in the limit distribution of $X_n + P_n$, and where the spikes are. Below, we apply the techniques introduced in [BGN11] to random matrices whose eigenvalue distributions have two bulks. Although we consider only the two-bulk case in this section, the analysis generalizes to the case of multiple bulks.

For simplicity, we consider additive rank-2 perturbation $P_{2n} = \theta_1 u_{2n} u_{2n}^\top + \theta_2 v_{2n} v_{2n}^\top$; the analysis generalizes to finite-rank perturbation and the multiplicative case. Following the derivation in Section 2, we know that z is a limit eigenvalue of $X_n + P_n$ and not a limit eigenvalue of X_n if and only if

$$G_{\mu_X}(z) = \theta_1^{-1}, \text{ or } G_{\mu_X}(z) = \theta_2^{-1}. \tag{11}$$

So the problem reduces to finding G_{μ_X}.

Suppose the limit eigenvalue distribution μ_X has two bulks. That is, let

$$\mu_Y = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(X_{2n})}, \quad \mu_Z = \lim_{n\to\infty} \frac{1}{n} \sum_{i=n+1}^{2n} \delta_{\lambda_i(X_{2n})}.$$

Then, $\operatorname{supp}(\mu_Z) = [a_1, b_1]$ and $\operatorname{supp}(\mu_Y) = [a_2, b_2]$, where $b_1 < a_2$. Without loss of generality, we consider diagonal $X_{2n} = \operatorname{diag}(Y_n, Z_n)$ where $Y_n = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and $Z_n = \operatorname{diag}(\lambda_{n+1}, \ldots, \lambda_{2n})$. By definition, the Cauchy transforms of μ_{Y_n} and μ_{Z_n} are given by

$$G_{\mu_{Y_n}}(z) = n^{-1} \sum\nolimits_{i=1}^{n} (z - \lambda_i)^{-1}, \quad G_{\mu_{Z_n}}(z) = n^{-1} \sum\nolimits_{i=n+1}^{2n} (z - \lambda_i)^{-1},$$

and G_{μ_Y}, G_{μ_Z} are their limits. Hence,

$$G_{\mu_{X_n}}(z) = (2n)^{-1} \sum\nolimits_{i=1}^{2n} (z - \lambda_i)^{-1} = \frac{1}{2}(G_{\mu_{Y_n}}(z) + G_{\mu_{Z_n}}(z)),$$

and $G_{\mu_X} = (G_{\mu_Y} + G_{\mu_Z})/2$.

5.1 Example: Wigner matrices with two bulks

Suppose μ_Y and μ_Z are two semicircle distributions. Then

$$G_{\mu_Z}(z) = \frac{8z - 4(a_1 + b_1) - 2\operatorname{sign}(2z - (a_1 + b_1))\sqrt{(2z - (a_1 + b_1))^2 - (b_1 - a_1)^2}}{(b_1 - a_1)^2},$$

where $z \le a_1$ or $z \ge b_1$, and

$$G_{\mu_Z}(z) = \frac{8z - 4(a_2+b_2) - 2\,\mathrm{sign}(2z-(a_2+b_2))\sqrt{(2z-(a_2+b_2))^2-(b_2-a_2)^2}}{(b_2-a_2)^2},$$

where $z \le a_2$ or $z \ge b_2$. Hence,

$$\begin{aligned} G_{\mu_X}(z) = &\frac{4z - 2(a_1+b_1) - \mathrm{sign}(2z-(a_1+b_1))\sqrt{(2z-(a_1+b_1))^2-(b_1-a_1)^2}}{(b_1-a_1)^2} \\ &+ \frac{4z - 2(a_2+b_2) - \mathrm{sign}(2z-(a_2+b_2))\sqrt{(2z-(a_2+b_2))^2-(b_2-a_2)^2}}{(b_2-a_2)^2}, \end{aligned}$$

where $z \le a_1$, $b_1 \le z \le a_2$, or $z \ge b_2$.

If $a_1 = -2, b_1 = 2, a_2 = 6, b_2 = 10$, we have

$$G(-2) = -G(10) = \sqrt{6} - 3 \approx -0.5505, \quad G(2) = -G(6) = \sqrt{2} - 1 \approx 0.4142.$$

That is, if $|\theta_i|^{-1} \ge 0.5505$, there is no new spike brought by it; if $0.4142 \le |\theta_i|^{-1} < 0.5505$, there is a new spike outside $[a_1, b_2]$ but none inside $[b_1, a_2]$; if $|\theta_i|^{-1} < 0.4142$, there is both a new spike outside $[a_1, b_2]$ and one inside $[b_1, a_2]$. Note that if the two semicircles have different sizes, it is also possible to only have a new middle spike.

5.2 Simulations

Below we show the simulations on different perturbations on the matrix X_n with two Wigner bulks.

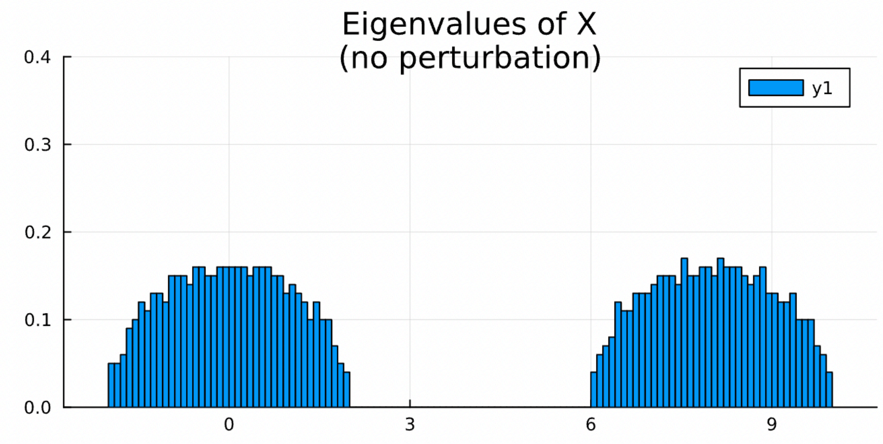


Figure 5

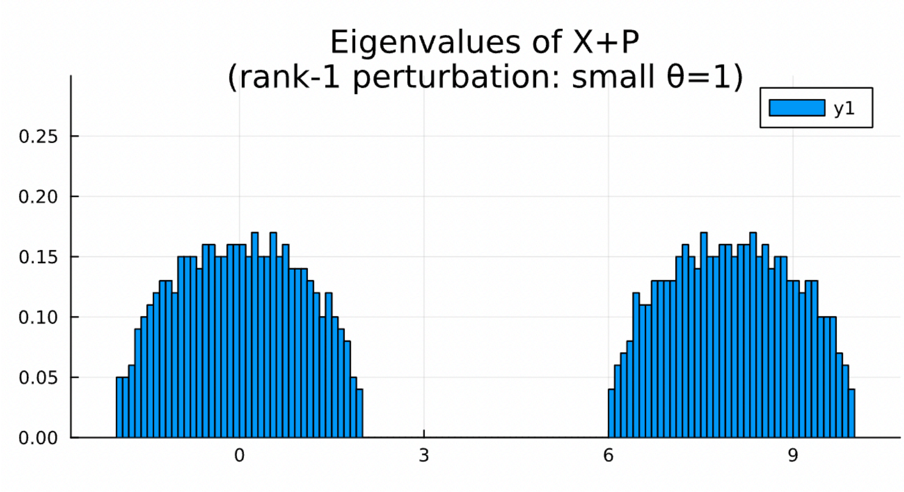


Figure 6

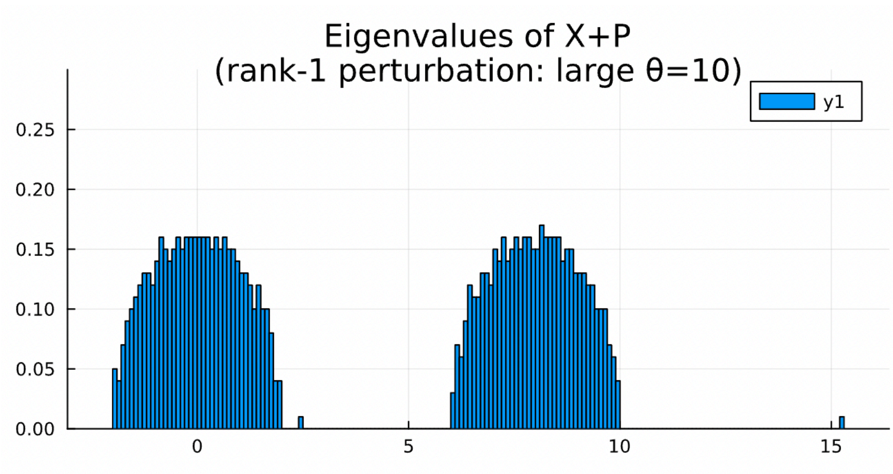


Figure 7

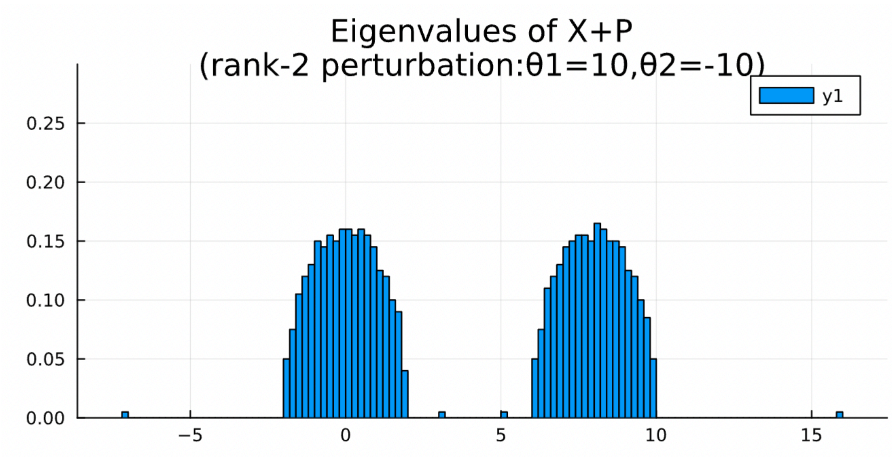


Figure 8

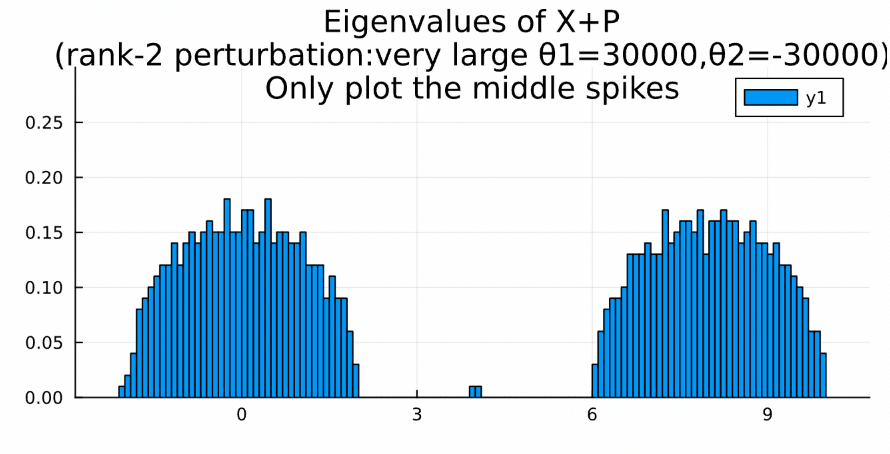
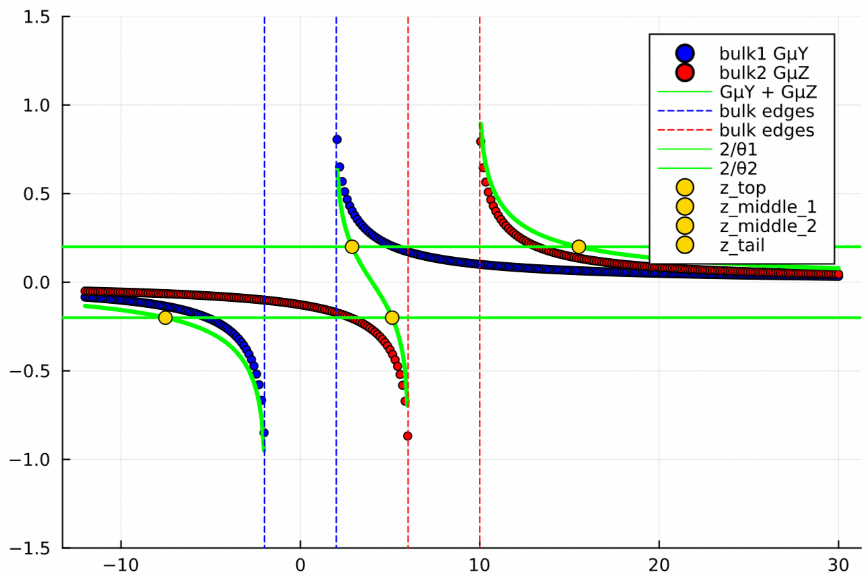


Figure 9

Figure 5 shows the eigenvalues of X_n without perturbation. In Figure 6, we add rank-1 perturbation with a small eigenvalue: $\theta = 1$. Since it is below the critical threshold, the spikes are not revealed. After increased the θ to be 10, we see the two top spikes in Figure 7, each from one of the two bulks.

In Figure 8, we plot the eigenvalues $X_n + P_n$ with a rank-2 perturbation matrix P_n. Since $\theta_1 > 0$ and $\theta_2 < 0$. each bulk has two spikes on each side of the bulk. When we set $|\theta_1|, |\theta_2|$ to be extremely large in Figure 9, the two spikes between the spikes are extremely close to each other. This is easily explained by our theoretical analysis at the beginning of this section. When $|\theta_1|, |\theta_2|$ both go to infinity, $\frac{1}{\theta_1} \approx \frac{1}{\theta_2}$. Therefore, the two solutions in (11) will collide.



6 Concluding remarks

In this project, we studied the limiting eigenvalues of the random matrices under finite-rank perturbations. We focus on studying [BGN11], which provides a general framework to study such problems without assuming the distribution of the original random matrix. After numerically verifying its claims and studying its techniques, we used the technical tools in [BGN11] to explore the problem where the spectrum of the original matrix consists of more than one bulk. We illustrated the analysis through the running example of the Wigner matrices, where the LSD of the original matrix has two semicircle bulks. The new spikes can again be precisely characterized by equations involving the Cauchy transform of LSD of the original matrix, which we verify through numercial simulations.

Below, we remark on the generalization of the analysis and possible future work.

- **Multiple bulks, finite-rank perturbation, and general ensembles.** In our example in Section 5, we only analyzed the spikes between two bulks with rank-2 perturbation. The analysis clearly generalizes to the case where the number of bulks and the rank of the perturbation matrix exceed two. Moreover, we used the example of Wigner matrices to verify our results. It would be nice to have case studies on other ensembles.

- **Multiplicative perturbation.** We mostly considered the additive perturbation $\widetilde{X}_n = X_n + P_n$. [BGN11] also characterizes the spike eigenvalue and eigenvectors for the multiplicative perturbation $\widetilde{X}_n = X_n(I + P_n)$. The analysis is similar but instead of using the Cauchy transform, we need to use the T-transform (see Theorem 2.6 in [BGN11]) and solve for $T_{\mu_X}(z) = \theta_i^{-1}$ to obtain the solutions for spike eigenvalues.

- **Complex random matrices.** In this project, we only considered real symmetric matrices for simplicity, but the arguments are general enough to be applied in the complex case. More details can be found in the paper [BGN11].

- **Shifting bulks in the complex plane.** Suggested by Prof. Alan Edelman, one could try shifting a bulk with imaginary numbers. The geometrical picture of the new spikes in this case can be very interesting. It should be noted that in this case the original matrix X_n is no longer Hermitian.

References

[BBAP05] Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *The Annals of Probability*, 33(5):1643–1697, 2005.

[BGN11] Florent Benaych-Georges and Raj Rao Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 227(1):494–521, 2011.

[BS98] Zhi-Dong Bai and Jack W Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *The Annals of Probability*, 26(1):316–345, 1998.

[BS99] ZD Bai and Jack W Silverstein. Exact separation of eigenvalues of large dimensional sample covariance matrices. *Annals of probability*, pages 1536–1555, 1999.

[BY12] Zhidong Bai and Jianfeng Yao. On sample eigenvalues in a generalized spiked population model. *Journal of Multivariate Analysis*, 106:167–177, 2012.

[Joh01] Iain M Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *The Annals of statistics*, 29(2):295–327, 2001.

[SB95] Jack W Silverstein and ZD Bai. On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate analysis*, 54(2):175–192, 1995.

[SC95] Jack W Silverstein and Sang-Il Choi. Analysis of the limiting spectral distribution of large dimensional random matrices. *Journal of Multivariate Analysis*, 54(2):295–309, 1995.

[Sil95] Jack W Silverstein. Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *Journal of Multivariate Analysis*, 55(2):331–339, 1995.