MESOSCOPIC PERTURBATIONS OF LARGE RANDOM MATRICES

JIAOYANG HUANG

Abstract. We consider the eigenvalues and eigenvectors of small rank perturbations of random $N \times N$ matrices. We allow the rank of perturbation $M$ increases with $N$, and the only assumption is $M = o(N)$. In both additive and multiplicative perturbation models, we prove rigidity results for the outliers of the perturbed random matrices. Based on the rigidity results we derive the empirical distribution of outliers of the perturbed random matrices. We also compute the appropriate projection of eigenvectors corresponding to the outliers of the perturbed random matrices, which are approximate eigenvectors of the perturbing matrix. Our results can be regarded as the extension of the finite rank perturbation case to the full generality up to $M = o(N)$.

1. Introduction

Enormous progress has been accomplished in the recent works on studying the influence of some perturbations on the asymptotic spectrum of large random matrices, in both the additive perturbation model,

$$\tilde{W}_N = W_N + P_N,$$

and the multiplicative perturbation model,

$$\tilde{W}_N = W_N(I_N + P_N),$$

where $W_N$ is the original random matrix, and $P_N$ is the perturbation. Most of the settings for the perturbations fall into two categories, either the perturbation is of finite rank, or the perturbation is of full rank.

In the first case when the perturbation is of finite rank, the spectrum is not much altered due to Weyl’s interlacement property of eigenvalues. However, the extreme eigenvalues of the perturbed matrix differ from that of the non-perturbed matrix if and only if the eigenvalues of the perturbation are above certain critical threshold. This phenomenon was made precise in [3], where the sharp transition (called the BBP phase transition) was first exhibited for the finite rank multiplicative perturbations of complex Gaussian Wishart matrix. In this case, it was shown that if the eigenvalues of the perturbation are above the threshold, the largest eigenvalue of the perturbed matrix deviates away from the bulk and has Gaussian fluctuation, otherwise it sticks to the bulk and fluctuates according to the Tracy-Widom law. Similar results were proved for the finite rank additive perturbations of complex Gaussian Wigner matrices in [19, 18]. Then in a series of papers [4, 2, 11, 10,
20, 21, 15, these results were generalized to Wishart and Wigner random ensembles with non-Gaussian entries. More generally, the results of BBP phase transition was extended to the case when the original matrix \( W_N \) and the perturbation \( P_N \) are orthogonally (or unitarily) independent in [8, 6, 7].

In the full rank perturbation case, when \( N \) becomes large, free probability provides us a good understanding of the global behavior of the asymptotic spectrum of the perturbed random matrices. More precisely, under some mild conditions, the empirical eigenvalue distribution of the perturbed random matrices converges to the free convolution of the limit empirical eigenvalue distributions of \( W_N \) and \( P_N \) both in expectation and almost surely. We refer to [26] and [14] for pioneering works and [1] for an introduction to free probability. For the additive perturbation of Wigner matrices (called deformed Wigner matrices), even the local statistics in the bulk of the spectrum are known. In fact, the bulk universality for deformed Wigner matrices was proved for a large class of diagonal perturbation \( P_N \) in [16, 17]. The mixer of finite rank and full rank perturbations was considered in [12, 5], where the empirical eigenvalue distribution of the perturbation converges to certain compactly supported measure and it has fixed number of fixed eigenvalues outside the support of the limit measure.

Now, how does the spectrum behave under a deterministic perturbation \( P_N \) of rank \( M \), such that the rank \( M \) may increase with \( N \), but still in the regime \( M = o(N) \)? To our knowledge, this kind of perturbations was only studied by Péché in [19, 18] for deformed complex Gaussian Wigner matrices, in which Péché obtained the limiting local statistics of the outliers of the deformed Wigner matrices. However, the method heavily depends on the explicit form of correlation function of the deformed Wigner matrices, and is hard to be extended to non-Gaussian case. In our paper, we prove rigidity result for the outliers of the perturbed random matrices in more general settings. Based on the rigidity result we derive the empirical distribution of outliers of perturbed random matrices. Following the paper [8], We also investigate the eigenvectors corresponding to the outliers of the perturbed random matrices. We prove that they are approximate eigenvectors of the perturbing matrix.

Our proofs rely on the derivation of master equation representations, like (3), of the eigenvalues and eigenvectors of the perturbed matrix, which is a standard way to study the outliers in the finite rank perturbation case (such as in [2, 8, 6, 23]). In finite rank perturbation case, this reduces the problem of understanding the outliers to the study of a finite rank matrix (the same rank as the perturbation), which is much easier to analyze. One can derive the location and the fluctuation of outliers by passing to limit. In our setting, we get an \( M \times M \) matrix, whose size increases with \( N \). Instead of passing to limit, we directly analyze such matrix. We find that such matrix is monotonic outside the spectrum of the original matrix. Based on the monotonic property and some concentration inequalities, we are able to detect the location of outliers with an exponentially high probability.
2. Additive Perturbation Case

2.1. Definition and Notation. In this section we study the eigenvalues and eigenvectors of the additively perturbed real symmetric (or Hermitian) random matrix $W_N$ by a deterministic diagonal matrix $P_N$ with small rank, 

$$\tilde{W}_N = W_N + P_N.$$ 

For the random matrix $W_N$, we consider two cases:

1. $W_N$ is orthogonally (or unitarily) independent with $P_N$. More precisely, $W_N = U^* H_N U$, where $U$ follows the Haar measure on $N \times N$ orthogonal (or unitary) group, and $H_N$ is an $N \times N$ symmetric (or Hermitian) matrix whose ordered eigenvalues we denote by $\lambda_1(H_N) \geq \lambda_2(H_N) \geq \cdots \geq \lambda_N(H_N)$. Let $\mu_{H_N}$ be the empirical eigenvalue distribution of $H_N$, i.e. $\mu_{H_N} = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(H_N)}$. Let $m_N(z)$ be the Stieltjes transform of the empirical eigenvalue distribution of $W_N$,

$$m_N(z) = \int_{\mathbb{R}} \frac{d\mu_N(x)}{z-x} = \frac{1}{N} \sum_{i=1}^N \frac{1}{z-\lambda_i(H_N)},$$

for $z \in (-\infty, \lambda_N(H_N)) \cup (\lambda_1(H_N), \infty)$.

2. $W_N = \frac{1}{\sqrt{N}} [w_{ij}]_{1 \leq i,j \leq N}$ is a Wigner ensemble. $w_{ij}$’s are i.i.d. random variables under the symmetric constraint $W_N = W_N^*$, with mean zero and variance one:

$$\mathbb{E}[w_{ij}] = 0, \quad \mathbb{E}[w_{ij}^2] = 1.$$ 

Moreover, for technical reason, we assume that $w_{ij}$’s are sub-gaussian with $\|w_{ij}\|_{\psi_2} \leq K$. We refer to [22] for the definition of the sub-gaussian norm and relative properties of sub-gaussian random variables.

The perturbation $P_N$ is an $N \times N$ real diagonal matrix having rank $M$, where $M = o(N)$. It has $M_1$ positive eigenvalues. We denote its nonzero eigenvalues by $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_{M_1} > 0 > \theta_{M_1+1} \geq \cdots \geq \theta_{M-1} \geq \theta_M$. Let $P_N = \text{diag}\{\theta_1, \theta_2, \cdots, \theta_{M-1}, \theta_M\}$ the $M \times M$ real diagonal matrix, consisting of the nonzero eigenvalues of $P_N$. Since the matrix $W_N$ is permutation invariant, without loss of generality, we can assume that

$$P_N = \begin{pmatrix} \tilde{P}_N & 0 \\ 0 & 0 \end{pmatrix}.$$ 

2.2. Orthogonally (or Unitarily) Independent Case. In this section we study the eigenvalues and eigenvectors of

$$\tilde{W}_N = U^* H_N U + P_N.$$ 

Without loss of generality we can assume that $H_N$ is a diagonal matrix. Since we concentrate only on the outliers of the perturbed matrix $\tilde{W}_N$, in the following we fixed a small universal constant $\delta > 0$, and study the eigenvalues of $\tilde{W}_N$ which are $\delta$ distance away from the bulk of $\mu_N$, i.e. on the spectral domain $(\lambda_N(H_N) - \delta, \lambda_1(H_N) + \delta)$.
Theorem 2.1. $H_N$ and $P_N$ are defined as in Section 2.1, we assume that their norms are bounded by some universal constant $B$, i.e. $\|H_N\| \leq B$, $\|P_N\| \leq B$. We denote the eigenvalues of $\tilde{W}_N = U^*H_NU + P_N$ by $\lambda_1(\tilde{W}_N) \geq \lambda_2(\tilde{W}_N) \geq \cdots \geq \lambda_N(\tilde{W}_N)$. For any $0 < \epsilon \leq \delta$ (it may depends on $M$ and $N$), and any $1 \leq i \leq M$, if the eigenvalue $\theta_i$ of the perturbation $P_N$ satisfies the following separation condition:

$$
m_N^{-1}(\frac{1}{\theta_i}) \geq \lambda_1(H_N) + 2\delta, \quad \text{if } \theta_i > 0,
$$

$$
m_N^{-1}(\frac{1}{\theta_i}) \leq \lambda_N(H_N) - 2\delta, \quad \text{if } \theta_i < 0,
$$

then it creates an outlier $\tilde{\lambda}_i$ of $\tilde{W}_N$, where $\tilde{\lambda}_i = \lambda_i(\tilde{W}_N)$ for $\theta_i > 0$, and $\tilde{\lambda}_i = \lambda_{N-M+i}(\tilde{W}_N)$ for $\theta_i < 0$, with high probability,

$$
P\left( m_N^{-1}(\frac{1}{\theta_i}) - \epsilon \leq \tilde{\lambda}_i \leq m_N^{-1}(\frac{1}{\theta_i}) + \epsilon \right) \geq 1 - C e^{-c_1Nc_2^2 + c_2M}
$$

where $c_1$, $c_2$ and $C$ are constants depending only on $\delta$ and $B$.

Proof. The eigenvalues of $U^*H_NU + P_N$ are the solutions of the equation

$$
det(z - H_N - UP_NU^*) = 0.
$$

For $z$ away from the bulk of $\mu_N$, on the spectral domain $(\lambda_N(H_N) - \delta, \lambda_1(H_N) + \delta]$, $z - H_N$ is invertible, we have

$$
det(z - H_N - UP_NU^*) = det(z - H_N) det(I - (z - H_N)^{-1}UP_NU^*)
$$

$$
= det(z - H_N) det(I - \tilde{U}^*(z - H_N)^{-1}\tilde{U} \tilde{P}_N),
$$

where $\tilde{U}$ is the $N \times M$ matrix, consisting of the first $M$ columns of $U$. Therefore the outliers of $U^*H_NU + P_N$ are solutions of the equation,

$$
det(\tilde{P}_N^{-1} - \tilde{U}^*(z - H_N)^{-1}\tilde{U}) = 0.
$$

In the following we concentrate on the outliers on the interval $[\lambda_1(H_N) + \delta, +\infty)$, and prove (2). Denote $D_N(z) = \tilde{P}_N^{-1} - \tilde{U}^*(z - H_N)^{-1}\tilde{U}$, which is a real symmetric (or Hermitian) matrix. Since $D_N(z)$ is analytic on the interval $[\lambda_1(H_N) + \delta, +\infty)$, we can parametrize its eigenvalues as

$$
\tau_1(z) \geq \tau_2(z) \geq \cdots \geq \tau_{M-1}(z) \geq \tau_M(z), \quad z \in [\lambda_1(H_N) + \delta, +\infty).
$$

Notice that for $z > z' \geq \lambda_1(H_N) + \delta$, we have

$$
D(z) - D(z') = \tilde{U}^*((z' - H_N)^{-1} - (z - H_N)^{-1})\tilde{U}
$$

$$
= (z - z')\tilde{U}^*(z' - H_N)^{-1}(z - H_N)^{-1}\tilde{U}
$$

$$
> 0.
$$

Therefore on the interval $[\lambda_1(H_N) + \delta, +\infty)$, $D(z)$ is an increasing function, in the sense of matrix. The following Lemma implies that $\tau_i(z)$’s are all increasing functions of $z$ on the interval $[\lambda_1(H_N) + \delta, +\infty)$, for $i = 1, 2 \cdots , M$. 

Lemma 2.2. A and B are two $N \times N$ real symmetric (or Hermitian) matrices, such that $A - B > 0$. Denote the $n$-th maximal eigenvalues of $A$ and $B$ as $\lambda_n(A)$ and $\lambda_n(B)$ respectively, then we have

$$\lambda_n(A) > \lambda_n(B), \quad n = 1, 2, 3, \ldots, N.$$ 

Proof. Take the space $W$ spanned by the eigenvectors of $B$ corresponding to its first $n$ eigenvalues. By the minimax principle,

$$\lambda_n(A) = \max_{\dim V = n} \min_{v \in V} v^* A v \geq \min_{v \in W} v^* A v.$$ 

Denote $A|_W$ and $B|_W$ the restriction of $A$ and $B$ on the space $W$, it is still true that $A|_W - B|_W > 0$. Therefore

$$\min_{v \in W} v^* A v = \min_{v \in W} v^* A|_W v > \min_{v \in W} v^* B|_W v = \lambda_n(B).$$

Remark 2.3. If we change the strictly inequality to $A - B \geq 0$, then we still have that $\lambda_n(A) \geq \lambda_n(B)$.

Determinant is the product of eigenvalues,

$$\det(\tilde{P}^{-1}_N - \tilde{U}^* (z - H_N)^{-1} \tilde{U}) = \prod_{i=1}^M \tau_i(z).$$

The determinant vanishes, if and only if some eigenvalue vanishes. Therefore $z \in [\lambda_1(H_N) + \delta, +\infty)$ corresponds to an outlier of $U^* H_N U + P_N$ if and only if $\tau_i(z) = 0$ for some $i$. Consider the point process $\{\tau_1(z), \tau_2(z), \ldots, \tau_M(z)\}$ as $z$ goes from $+\infty$ down to $\lambda_1(H_N) + \delta$.

For $z$ sufficiently large, such that $z - \lambda_1(H_N) > \theta_1$, we have

$$0 < \tilde{U}^* (z - H_N)^{-1} \tilde{U} < \frac{1}{\theta_1},$$

which is smaller than the smallest positive eigenvalue of $\tilde{P}^{-1}_N$. Therefore for $z > \lambda_1(H_N) + \theta_1$, $\tilde{P}^{-1}_N - \tilde{U}^* (z - H_N)^{-1} \tilde{U}$ has the same number of positive (and negative) eigenvalues as $\tilde{P}^{-1}_N$,

$$\tau_1(z) \geq \tau_2(z) \geq \cdots \geq \tau_{M_1}(z) > 0 > \tau_{M_1+1}(z) \geq \cdots \geq \tau_M(z).$$

As $z$ goes from $+\infty$ down to $\lambda_1(H_N) + \delta$, all the eigenvalues $\tau_i(z)$’s will decrease. At some point $z = z_1$, $\tau_{M_1}(z_1)$ becomes zero, this $z_1$ is exactly the largest outlier $\lambda_1(H_N)$. Later at some point $z = z_2$, $\tau_{M_1-1}(z_2)$ becomes zero, this $z_2$ is the second largest outlier $\lambda_2(H_N)$. We may now continue in this manner, sequentially get all the outliers of $H_N$ on the interval $[\lambda_1(H_N) + \delta, +\infty)$.

More formally, we introduce the counting function:

$$n(z) = \# \{i \in \{1, 2, \ldots, M\} : \tau_i(z) \geq 0\}.$$
Since $\tau_i(z)$’s are all strictly increasing function of $z$ on the interval $[\lambda_1(H_N) + \delta, +\infty)$, $n(z)$ is a right continuous, non-decreasing function. The $i$-th largest outlier of $\tilde{H}_N$ is given by

$$\lambda_i(\tilde{H}_N) = \inf_{z \in [\lambda_1 + \delta, +\infty)} \{ z : n(z) \geq M_1 - i + 1 \},$$

if the set on the lefthand side is not empty. The counting function $n(z)$ encodes all the information of the outliers of $\tilde{H}_N$, and it can be used to detect the location of the $i$-th largest outlier of $\tilde{H}_N$. More precisely, if we can find numbers $\lambda_i(H_N) + \delta \leq L < R$ such that

$$n(L) \leq M_1 - i, \quad n(R) \geq M_1 - i + 1,$$

then we can conclude that $\lambda_i(\tilde{H}_N) \in (L, R]$.

In the following, we prove (2), by taking in (4)

$$L = m_N^{-1}(\frac{1}{B_i}) - \epsilon, \quad R = m_N^{-1}(\frac{1}{B_i}) + \epsilon.$$ 

To do this, we use Proposition A.1, which states that roughly $\tilde{U}^*(z - H_N)^{-1}\tilde{U} \approx m_N(z)$. Take $A = (L - H_N)^{-1}$ and $\xi = \frac{4B^2}{\delta^2}$ in Proposition A.1, we have that

$$\tilde{P}^{-1}_N - \tilde{U}^*(L - H_N)^{-1}\tilde{U} \leq \tilde{P}^{-1}_N - m_N(m_N^{-1}(\frac{1}{B_i}) - \epsilon) + \frac{\epsilon}{8B^2}$$

$$\leq \tilde{P}^{-1}_N - \frac{1}{\theta_i} - \frac{\epsilon}{4B^2} + \frac{\epsilon}{8B^2}$$

$$= \tilde{P}^{-1}_N - \frac{1}{\theta_i} - \frac{\epsilon}{8B^2},$$

with probability at least $1 - C e^{-c_1 N \epsilon^2 + c_2 M}$, where $C, c_1$ and $c_2$ depends only on $\delta$ and $B$. One can find the proof for the inequality (5) in the Appendix Proposition B.1. Therefore by Lemma 2.2, we have $n(L) \leq M_1 - i$ with high probability. Similarly, if we take $A = (R - H_N)^{-1}$ and $\xi = \frac{4B^2}{\delta^2}$ in Proposition A.1, we get that $n(R) \geq M_1 - i + 1$ with probability at least $1 - C e^{-c_1 N \epsilon^2 + c_2 M}$. This finishes the proof of (2). \qed

**Remark 2.4.** Since in our setting, the rank of perturbation $P_N$ is much smaller than $N$, $M = o(N)$, (2) in fact gives us the location of outliers of the perturbed matrix $\tilde{W}_N$ with exponentially high probability. Moreover, (2) implies that the fluctuation of the outliers is at most of order $\sqrt{\frac{M}{N}}$.

**Remark 2.5.** If we take $M$ as a fixed number, then Theorem 2.1 uncovers the well-known result of the finite rank perturbation of random matrices. Moreover, in this case, Theorem 2.1 gives us the correct order of fluctuations of the outliers. In fact it is proved in [6] that the outliers exhibit Gaussian fluctuations with order $N^{-\frac{1}{2}}$, and the joint distribution of normalized outliers converges to the law of the eigenvalues of independent GUE or GOE random matrices. However we do not know how to generalize their method to give the local statistics of outliers in mesoscopic case.
Remark 2.6. If there exists a large gap between some two adjacent eigenvalues of $H_N$, i.e. $\lambda_j(H_N) - \lambda_{j+1}(H_N) \gg \delta$, Theorem 2.1 can be adapted to describe the outliers on the interval $(\lambda_{j+1}(H_N), \lambda_j(H_N))$. In fact the Stieltjes transform $m_N(z)$ of the empirical eigenvalue distribution of $H_N$ is well defined on $(\lambda_{j+1}(H_N), \lambda_j(H_N))$, and is strictly decreasing. Therefore for any eigenvalue $\theta_i$ of the perturbation $P_N$, under certain separation condition analogue to (1), it will create an outlier around $m_N(\frac{1}{\theta_i})$ with high probability. This comment applies to the other theorems in this paper.

Theorem 2.7. $H_N$ and $P_N$ are defined as in Section 2.1, we assume that their norms are bounded by some universal constant $B$, i.e. $||H_N|| \leq B$, $||P_N|| \leq B$. We denote the eigenvalues of $\hat{W}_N = U^*H_NU + P_N$, i.e. $\lambda_1(\hat{W}_N) \geq \lambda_2(\hat{W}_N) \geq \cdots \geq \lambda_N(\hat{W}_N)$. For any $1 \leq i \leq M$, such that the eigenvalue $\theta_i$ of the perturbation $P_N$ satisfies the separation condition (1), with high probability, $\tilde{\lambda}_i \equiv \lambda_i(\hat{W}_N)$ for $\theta_i > 0$, and $\tilde{\lambda}_i = \lambda_{N-M+i}(\hat{W}_N)$ for $\theta_i < 0$. We denote the corresponding eigenvector by $v_i$ (if $\tilde{\lambda}_i$ is not of multiplicity one, $v_i$ can be any of its eigenvectors). For any $0 \leq \epsilon \leq \delta$, it may depends on $M$ and $N$, with probability at least $1 - C e^{-c_1 N^c + c_2 M}$, where $c_1$, $c_2$ and $C$ are constants depending only on $\delta$ and $B$, we have

1. We denote the projection of $v_i$ on the eigenspace of nonzero eigenvalues of $P_N$ by $\tilde{v}_i$, which is the projection of $v_i$ on the first $M$ coordinates. The norm of $\tilde{v}_i$ satisfies

$$||\tilde{v}_i||^2 + \frac{1}{\theta_i^2 m_N'(m_N^{-1}(\frac{1}{\theta_i}))} \leq \epsilon C_{B, \delta}.$$ (6)

2. The projection $\tilde{v}_i$ is an approximate eigenvector of $\tilde{P}_N$ with eigenvalue $\theta_i$,

$$|\tilde{P}_N \tilde{v}_i - \theta_i \tilde{v}_i| \leq \epsilon C_{B, \delta}.$$ (7)

where $C_{B, \delta}$ is a constant depending on $\delta$ and $B$.

Proof. Since $v_i$ is the eigenvector of $U^*H_NU + P_N$ corresponding to the eigenvalue $\tilde{\lambda}_i$,

$$\tilde{\lambda}_i v_i = (U^*H_NU + P_N)v_i.$$ (8)

Rearrange (8), we get

$$v_i = U^* (\tilde{\lambda}_i - H_N)^{-1} U P_N v_i.$$ (9)

Project both sides of (9) on the first $M$ coordinates, and take the norm square on both sides of (9), we obtain the following two characterizing equations for $\tilde{v}_i$,

$$\left(I_M - U^* (\tilde{\lambda}_i - H_N)^{-1} U \tilde{P}_N\right) \tilde{v}_i = 0,$$ (10)

$$\tilde{v}_i^* \tilde{P}_N \left(U^*(\tilde{\lambda}_i - H_N)^{-1} U \right) \tilde{P}_N \tilde{v}_i = 1,$$ (11)

where $\tilde{U}$ is the $N \times M$ matrix, consisting of the first $M$ columns of $U$. Approximately, we have $U^*(\lambda_i - H_N)^{-1} U \approx \frac{1}{\theta_i}$ and $U^*(\lambda_i - H_N)^{-2} \tilde{U} \approx -m_N'(m_N^{-1}(\frac{1}{\theta_i}))$, which will simplify
the algebraic relations (10) and (11). In the following, we will make these approximations more quantitative. Take \( K = \max\{\frac{8}{\delta^2}, \frac{16}{\delta^3}\} \), \( L = m_N^{-1}(\frac{1}{\delta^2}) - \frac{\epsilon}{K} \), and \( R = m_N^{-1}(\frac{1}{\delta^2}) + \frac{\epsilon}{K} \) in Proposition B.1, we obtain the following inequalities

\[
\frac{1}{\theta_i} - \frac{\epsilon}{2} \leq m_N(R) \leq \frac{1}{\theta_i} - \frac{\epsilon}{4KB^2} < \frac{1}{\theta_i} + \frac{\epsilon}{4KB^2} \leq m_N(L) \leq \frac{1}{\theta_i} + \frac{\epsilon}{2},
\]

\[-m'_N(m_N^{-1}(\frac{1}{\theta_i})) - \frac{\epsilon}{2} \leq -m'_N(R) \leq -m'_N(L) \leq -m'_N(m_N^{-1}(\frac{1}{\theta_i})) + \frac{\epsilon}{2}.\]

By taking \( \xi = \min\{\frac{\epsilon}{2}, \frac{\epsilon}{8KB^2}\} \), and \( A = (L - H_N)^{-1}, \ (R - H_N)^{-1}, \ (L - H_N)^{-2} \) and \( (R - H_N)^{-2} \) respectively in Proposition A.1, we get

\[
\frac{1}{\theta_i} - \epsilon \leq \tilde{U}^*(R - H_N)^{-1}\tilde{U} < \frac{1}{\theta_i} < \tilde{U}^*(L - H_N)^{-1}\tilde{U} \leq \frac{1}{\theta_i} + \epsilon,
\]

and

\[-m'_N(m_N^{-1}(\frac{1}{\theta_i})) - \epsilon \leq \tilde{U}^*(R - H_N)^{-2}\tilde{U} \leq \tilde{U}^*(L - H_N)^{-2}\tilde{U} \leq -m'_N(m_N^{-1}(\frac{1}{\theta_i})) + \epsilon,
\]

with exponentially high probability, i.e. at least \( 1 - Ce^{-c_1N\epsilon^2 + c_2M} \). We denote the event such that (12) and (13) hold by \( \mathcal{A} \). In the following we show that (6) and (7) hold on \( \mathcal{A} \).

The same argument as in the proof of Theorem 2.1, (12) implies that \( \tilde{\lambda}_i \in [L, R] \). Since both \( \tilde{U}^*(z - H_N)^{-1}\tilde{U} \) and \( \tilde{U}^*(z - H_N)^{-2}\tilde{U} \) are monotonic as a function of \( z \), (12) and (13) implies

\[
\frac{1}{\theta_i} - \epsilon \leq \tilde{U}^*(\tilde{\lambda}_i - H_N)^{-1}\tilde{U} \leq \frac{1}{\theta_i} + \epsilon
\]

(14)

\[-m'_N(m_N^{-1}(\frac{1}{\theta_i})) - \epsilon \leq \tilde{U}^*(\tilde{\lambda}_i - H_N)^{-2}\tilde{U} \leq -m'_N(m_N^{-1}(\frac{1}{\theta_i})) + \epsilon
\]

(15)

on the event \( \mathcal{A} \). With the quantitative estimate (14), (10) can be reduced to

\[
\left| \left( I_M - \frac{1}{\theta_i} \tilde{P}_N \right) \tilde{v}_i \right| \leq \epsilon B,
\]

(16)

which gives us (7). Similarly for (11), using the approximation \( \tilde{U}^*(\tilde{\lambda}_i - H_N)^{-2}\tilde{U} \approx -m'_N(m_N^{-1}(\frac{1}{\theta_i})) \) from (15), and \( \frac{1}{\theta_i} \tilde{P}_N \tilde{v}_i \approx \tilde{v}_i \) from (16), we get

\[
|\tilde{v}_i|^2 = \left| \frac{1}{\theta_i} \tilde{P}_N \tilde{v}_i \right|^2 + O_{B,\delta}(\epsilon) = \frac{1}{-\theta_i^2 m'_N(m_N^{-1}(\frac{1}{\theta_i}))} + O_{B,\delta}(\epsilon).
\]

(17)

This finishes the proof of (6).

\[ \square \]

2.3. **Wigner Case.** In this section we study the eigenvalues of perturbed Wigner matrices,

\[ \tilde{W}_N = W_N + P_N. \]
It is well known that the empirical eigenvalue distribution of $W_N$ converges to the famous semi-circle law with density
\[ \rho_{sc}(x) = \frac{\sqrt{4 - x^2}}{2\pi}, \quad \text{for } x \in [-2, 2]. \]

We denote its Stieltjes transform as
\[ m_{sc}(z) = \frac{z - \text{sgn}(z)\sqrt{z^2 - 4}}{2}, \quad \text{for } z \in (-\infty, -2] \cup [2, +\infty). \]

In the following we fix a small universal constant $\delta > 0$. It is known that with high probability the empirical eigenvalue distribution of Wigner matrix $W_N$ is supported on $(2 - \delta, 2 + \delta)$. Since we concentrate only on the outliers of the perturbed matrix $\tilde{W}_N$, we study the eigenvalues of $\tilde{W}_N$ on the spectral domain $(-2 - \delta, 2 + \delta)$.

**Theorem 2.8.** Wigner matrix $W_N$ and perturbation $P_N$ are defined as in Section 2.1, we assume the norm of $P_N$ is bound by some universal constant $B$, i.e. $\|P_N\| \leq B$. We denote the eigenvalues of $\tilde{W}_N = W_N + P_N$ by $\lambda_1(\tilde{W}_N) \geq \lambda_2(\tilde{W}_N) \geq \cdots \geq \lambda_N(\tilde{W}_N)$. For any $\sqrt{\frac{M}{N}} \ll \epsilon \leq \delta$ (it may depend on $M$ and $N$), and any $1 \leq i \leq M$, if $\theta_i$ satisfies the following separation condition
\[ |\theta_i| \geq 1 + 2\delta, \quad \text{(18)} \]

then it will create an outlier $\tilde{\lambda}_i$ of $\tilde{W}_N$, where $\tilde{\lambda}_i = \lambda_i(\tilde{W}_N)$ for $\theta_i > 0$, and $\tilde{\lambda}_i = \lambda_{N-M+i}(\tilde{W}_N)$ for $\theta_i < 0$, with high probability,
\[ P \left( \theta_i + \frac{1}{\theta_i} - \epsilon \leq \tilde{\lambda}_i \leq \theta_i + \frac{1}{\theta_i} + \epsilon \right) \geq 1 - C \left( e^{-c_1N\epsilon^2 + c_2M} + \frac{1}{(N - M)c_3\ln\ln(N-M)} \right) \quad \text{(19)} \]

where $c_1$, $c_2$, $c_3$ and $C$ are constants depending only on $\delta$ and $B$.

**Proof.** The proof is quite similar to the orthogonally independent case. In the following proof, we concentrate on the outliers on the interval $[2 + \delta, +\infty)$. We decompose the Wigner matrix $W_N$ into four submatrices
\[ W_N = \begin{pmatrix} W_N^{(1)} & B^* \\ B & W_N^{(2)} \end{pmatrix}, \]

where $W_N^{(1)}$ is an $M \times M$ submatrix, $W_N^{(2)}$ is an $(N - M) \times (N - M)$ submatrix. If $z - W_N^{(2)}$ is invertible, we have
\[ \det(z - W_N - P_N) = \det(z - W_N^{(2)}) \det \left( z - W_N^{(1)} - \tilde{P}_N - B^* (z - W_N^{(2)})^{-1} B \right). \]
Since $W_N^{(2)}$ is independent of $W_N^{(1)}$ and $B$, we can go through the same proof as in Theorem 2.1 once we condition on $W_N^{(2)}$. Define the event,

$$A_N = \left\{ \|W_N^{(2)}\| \leq 2 + \frac{\delta}{2}, \sup_{z \in (-2-\delta,2+\delta)} |m_N^{(2)}(z) - m_{sc}(z)| \leq \frac{\epsilon}{2} \right\}$$

where $m_N^{(2)}$ is the Stieltjes transform of empirical eigenvalue distribution of $W_N^{(2)}$. Since $\sqrt{\frac{N}{N-M}} W_N^{(2)}$ is an $(N-M) \times (N-M)$ Wigner matrix, from the rigidity results of eigenvalues of generalized Wigner Matrices [13], for $\epsilon \gg \sqrt{\frac{M}{N}}$, we have

$$P(A_N) \geq 1 - Ce^{-c_1 N \epsilon^2} + c_2 M.$$  \ ...(20)

where the constant $c_3$ and $C$ depend only on $\delta$. For any $W_N^{(2)} \in A_N$, conditional on $W_N^{(2)}$, for any $z \in [2 + \delta, +\infty)$, $z - W_N^{(2)}$ is invertible. Therefore the outliers of $\tilde{W}_N$ are the solutions of

$$\det \left( \left( z - W_N^{(1)} - B^*(z - W_N^{(2)})^{-1} B \right) - \hat{P}_N \right) = 0.$$ 

Denote $D_N(z) = \left( z - W_N^{(1)} - B^*(z - W_N^{(2)})^{-1} B \right) - \hat{P}_N$, one can check it is increasing on the interval $[2 + \delta, +\infty)$. By Proposition A.3, we have

$$P(\|W_N^{(1)}\| \geq \frac{\epsilon}{2}) \leq Ce^{-c_1 N \epsilon^2 + c_2 M}.$$ 

Therefore $W_N^{(1)}$ is ignorable. By Proposition A.2, we have that $B^*(z - W_N^{(2)})^{-1} B \approx m_N^{(2)}(z) \approx m_{sc}(z)$. Thus the same argument as in Theorem 2.1 implies

$$P \left( \theta_i + \frac{1}{\theta_i} - \epsilon \leq \lambda_i(\tilde{W}_N) \leq \theta_i + \frac{1}{\theta_i} + \epsilon \right) \geq 1 - Ce^{-c_1 N \epsilon^2} + c_2 M$$  \ ...(21)

where $c_1$, $c_2$ and $C$ are constants depending only on $\delta$ and $B$. (20) and (21) together gives us (19).

With the rigidity result of outliers from Theorem 2.8, one can easily deduce the empirical distribution of these outliers if the empirical distribution of nonzero eigenvalues of $P_N$ goes to some limit which satisfies the separation condition (18). In fact we have the following corollary,

**Corollary 2.9.** Wigner ensemble $\tilde{W}_N$ and perturbation $P_N$ are defined as in Section 2.1, we assume the norm of $P_N$ is bound by some universal constant $B$, i.e. $\|P_N\| \leq B$. Moreover we assume that the empirical distribution of the nonzero eigenvalues of $P_N$ converges weakly to a compactly supported measure $\nu$,

$$\nu_N = \frac{1}{M} \sum_{i=1}^{M} \delta_{\theta_i} \rightarrow \nu, \quad N \rightarrow +\infty,$$ 

...
with support \( \text{supp } \mu \subset [a, b] \), and \( a > 1 \) (notice that \( \theta_i \)'s depend on \( N \)). We denote the eigenvalues of \( \tilde{W}_N = W_N + P_N \) as \( \lambda_1(\tilde{W}_N) \geq \lambda_2(\tilde{W}_N) \geq \cdots \geq \lambda_N(\tilde{W}_N) \). Then almost surely, the empirical distribution of the largest \( M \) eigenvalues of \( \tilde{W}_N \) converges weakly to the push forward measure \( \gamma_\# \mu \),

\[
\frac{1}{M} \sum_{i=1}^{M} \sigma_{\lambda_i(\tilde{W}_N)} \to \gamma_\# \nu, \quad \text{a.s.}
\]
as \( N \) goes to \( +\infty \), where \( \gamma(\theta) = \theta + \frac{1}{\theta} \).

**Proof.** Fix the constant \( \delta = \frac{1}{3}(a - 1) \), and define the event

\[
A_N = \bigcap_{i: \theta_i \geq a - \delta} \left\{ \left| \lambda_i(\tilde{W}_N) - \theta_i - \frac{1}{\theta_i} \right| \leq \epsilon_N \right\}
\]
where \( \epsilon_N \) will be chosen later. Since for these \( \theta_i \geq a - \delta \), they satisfy the separation condition (18),

\[
\theta_i \geq a - \delta \geq 1 + 2\delta.
\]

From theorem 2.8, there exist constants \( c_1, c_2, c_3 \) and \( C \) such that

\[
\mathbb{P} \left( A_N^C \right) \leq C M \left( e^{-c_1 N \epsilon^2_N} + M^{c_2} E^{c_3 \ln \ln(N-M)} \right).
\]

Therefore if we take \( \epsilon_N = \min\left\{ \delta, \left( \frac{2a}{c_1} \right)^{\frac{1}{2}} \left( \frac{M}{N} \right)^{\frac{1}{2}} \right\} \), since \( M = o(N) \), we have \( \epsilon_N \to 0 \). Moreover by Borel-Cantelli lemma, almost surely \( A_N \)'s hold. For any bounded Lipschitz test function \( f \), we need to show the following expression converges almost surely,

\[
\lim_{N \to \infty} \frac{1}{M} \sum_{i=1}^{M} f(\lambda_i(\tilde{W}_N)) = \lim_{N \to \infty} \left( \frac{1}{M} \sum_{i=1}^{M} f(\lambda_i(\tilde{W}_N)) - f(\theta_i + \frac{1}{\theta_i}) \right) + \left( \frac{1}{M} \sum_{i=1}^{M} f(\theta_i + \frac{1}{\theta_i}) \right).
\]
The first term can be decomposed into two terms. The first term corresponds to \( \theta_i \)'s outside of the support of \( \nu \). Due to weak convergence of \( \nu_N \), the portion of such \( \theta_i \)'s goes to zero,

\[
\left| \frac{1}{M} \sum_{i: \theta_i \leq a - \delta} f(\lambda_i(\tilde{W}_N)) - f(\theta_i + \frac{1}{\theta_i}) \right| \leq \frac{2\|f\|_\infty}{M} \# \{i: \theta_i \leq a - \delta \} \to 0.
\]
The second term corresponds to \( \theta_i > a - \delta \). Since almost surely \( A_N \)'s hold, which implies that simultaneously \( |\lambda_i(\tilde{W}_N) - \theta_i + \frac{1}{\theta_i}| \leq \epsilon_N \). We can control the second term by Lipschitz norm of \( f \),

\[
\left| \frac{1}{M} \sum_{i: \theta_i > a - \delta} f(\lambda_i(\tilde{W}_N)) - f(\theta_i + \frac{1}{\theta_i}) \right| \leq \epsilon_N \|f\|_\mathcal{L} \to 0, \quad \text{a.s.}
\]
Therefore almost surely we have the convergence of empirical distribution of the largest $M$ eigenvalues of $\tilde{W}_N$,

$$\lim_{N \to \infty} \frac{1}{M} \sum_{i=1}^{M} f(\lambda_i(\tilde{W}_N)) = \lim_{N \to \infty} \sum_{i=1}^{M} f(\theta_i + \frac{1}{\theta_i}) = \int f(\theta) d\gamma_{\#} \mu.$$ 

Remark 2.10. If we assume that the extreme nonzero eigenvalues of $P_N$ converges to $a$ and $b$ respectively, the edge of the limit distribution $\nu$, then almost surely we have

$$\lambda_1(\tilde{W}_N) \to b + \frac{1}{b}, \quad \lambda_M(\tilde{W}_N) \to a + \frac{1}{a}.$$ 

3. Multiplicative Perturbation Case

3.1. Definition and Notation. In this section we study the eigenvalues and eigenvectors of the multiplicative perturbed real symmetric (or Hermitian) random matrix $W_N$ by a deterministic diagonal matrix $P_N$ with small rank,

$$\tilde{W}_N = W_N(I_N + P_N).$$

For the random matrix $W_N$, we consider two cases:

1. $W_N$ is orthogonally (or unitarily) independent with $P_N$. More precisely, $W_N = U^*H_NU$, where $U$ follows the Haar measure on $N \times N$ orthogonal (or unitary) group, and $H_N$ is a nonzero $N \times N$ positive semi-definite symmetric (or Hermitian) matrix whose ordered eigenvalues we denote by $\lambda_1(H_N) \geq \lambda_2(H_N) \geq \cdots \geq \lambda_N(H_N) \geq 0$. Let $\mu_{H_N}$ be the empirical eigenvalue distribution of $H_N$, i.e.,

$$\mu_{H_N} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i(H_N)}.$$ 

Let $T_N(z)$ be the $T$-transform of the empirical eigenvalue distribution of $H_N$,

$$T_N(z) = \int_{\mathbb{R}} \frac{x}{z - x} d\mu_{H_N}(x) = \frac{1}{N} \sum_{i=1}^{N} \frac{\lambda_i(H_N)}{z - \lambda_i(H_N)}.$$ 

for $z \in (-\infty, \lambda_N(H_N)) \cup (\lambda_1(H_N), \infty)$.

2. $W_N = X_NX_N^*$ is a Wishart ensemble, where $X_N = \frac{1}{\sqrt{p}}[x_{ij}]_{i \leq j \leq N}$ is a $N \times p$ random matrix. As $N \to +\infty$, the ratio $\frac{N}{p}$ goes to some positive constant less than one, i.e., \(\frac{N}{p} \to \phi \in (0, 1)\). $x_{ij}$’s are i.i.d. random variables with mean zero and variance one:

$$\mathbb{E}[x_{ij}] = 0, \quad \mathbb{E}[x_{ij}^2] = 1.$$ 

Moreover, for technical reason, we assume that $x_{ij}$’s are sub-gaussian with $\|x_{ij}\|_{\psi_2} \leq K$.

The perturbation $P_N$ is the same as defined in Section 2.1.
3.2. Orthogonally (or Unitarily) Independent Case. In this section we study the eigenvalues and eigenvectors of 
\[ \tilde{W}_N = U^* H_N U (I_N + P_N), \]
Without loss of generality we can assume that \( H_N \) is a diagonal matrix. Since we concentrate only on the outliers of the perturbed matrix \( \tilde{W}_N \), in the following we fix a small universal constant \( \delta > 0 \), and study the eigenvalues of \( \tilde{W}_N \) which are \( \delta \) distance away from the bulk of \( \mu_N \), i.e. on the spectral domain \( (\lambda_N(H_N) - \delta, \lambda_1(H_N) + \delta) \).

**Theorem 3.1.** \( H_N \) and \( P_N \) are defined as in Section 3.1, we assume that their norms are bounded by some universal constant \( B \), i.e. \( \|H_N\| \leq B, \|P_N\| \leq B \). We denote the eigenvalues of \( \tilde{W}_N = U^* H_N U (I_N + P_N) \) by \( \lambda_1(\tilde{W}_N) \geq \lambda_2(\tilde{W}_N) \geq \cdots \geq \lambda_N(\tilde{W}_N) \). For any \( 0 < \epsilon \leq \delta \) (it may depend on \( M \) and \( N \)), and any \( 1 \leq i \leq M \), if the eigenvalue \( \theta_i \) of the perturbation \( P_N \) satisfies the following separation condition:

\[ T_N^{-1}(\frac{1}{\theta_i}) \geq \lambda_1(H_N) + 2\delta, \quad \text{if } \theta_i > 0, \]
\[ T_N^{-1}(\frac{1}{\theta_i}) \leq \lambda_N(H_N) - 2\delta, \quad \text{if } \theta_i < 0, \]
then it creates an outlier \( \tilde{\lambda}_i \) of \( \tilde{W}_N \), where \( \tilde{\lambda}_i = \lambda_i(\tilde{W}_N) \) for \( \theta_i > 0 \), and \( \tilde{\lambda}_i = \lambda_N-M+i(\tilde{W}_N) \) for \( \theta_i < 0 \), with high probability,

\[ \mathbb{P} \left( T_N^{-1}(\frac{1}{\theta_i}) - \epsilon \leq \tilde{\lambda}_i \leq T_N^{-1}(\frac{1}{\theta_i}) + \epsilon \right) \geq 1 - C e^{-c_1 N \epsilon^2 + c_2 M} \]

where \( c_1, c_2 \) and \( C \) are constants depending only on \( \delta \) and \( B \).

**Proof.** The proof is similar to the additive perturbation case. For any \( z \in (\lambda_N(H_N) - \delta, \lambda_1(H_N) + \delta) \), \( z - H_N \) is invertible, we have

\[ \det(z - U^* H_N U (I_N + P_N)) = \det(z - H_N) \det(I_N - (z - H_N)^{-1} H_N U P_N U^*) \]
\[ = \det(z - H_N) \det(I_M - \tilde{U}^*(z - H_N)^{-1} H_N \tilde{U} \tilde{P}_N), \]
where \( \tilde{U} \) is the \( N \times M \) matrix, consisting of the first \( M \) columns of \( U \). The outliers of \( U^* H_N U (I_N + P_N) \) are solutions of the equation,

\[ \det(\tilde{P}_N^{-1} - \tilde{U}^*(z - H_N)^{-1} H_N \tilde{U}) = 0. \]

Denote \( D_N(z) = \tilde{P}_N^{-1} - \tilde{U}^*(z - H_N)^{-1} H_N \tilde{U} \). Since \( H_N \) is non-negative one can check that \( D_N(z) \) is non-decreasing respectively on the intervals \( (-\infty, \lambda_N(H_N) - \delta] \) or \( [\lambda_1(H_N) + \delta, +\infty) \). This implies that the eigenvalues of \( D_N(z) \) are non-decreasing. Since (24) has finitely many solutions, eigenvalues of \( D_N \) can not stay at zero. (In fact one can show that eigenvalues of \( D_N(z) \), as a function of \( z \), are either nonzero constant or strictly increasing.) Therefore we can go through the same proof as in Theorem 2.1. The only difference in this case is that

\[ \tilde{U}^*(z - H_N)^{-1} H_N \tilde{U} \approx T_N^{-1}(z). \]
This is the reason the T-transform appears in the statement instead of the Stieltjes transform.

\[ \square \]

**Theorem 3.2.** $H_N$ and $P_N$ are defined as in Section 3.1, we assume that their norms are bounded by some universal constant $B$, i.e. $\|H_N\| \leq B$, $\|P_N\| \leq B$. We denote the eigenvalues of $\tilde{W}_N = U^*H_NU(I_N + P_N)$ by $\lambda_1(\tilde{W}_N) \geq \lambda_2(\tilde{W}_N) \geq \cdots \geq \lambda_N(\tilde{W}_N)$. For any $1 \leq i \leq M$, such that the eigenvalue $\theta_i$ of the perturbation $P_N$ satisfies the separation condition (22), with high probability, it will create an outlier $\tilde{\lambda}_i$, where $\tilde{\lambda}_i = \lambda_i(\tilde{W}_N)$ for $\theta_i > 0$, and $\tilde{\lambda}_i = \lambda_{N-M+i}(\tilde{W}_N)$ for $\theta_i < 0$. We denote the corresponding eigenvector by $v_i$ (if $\tilde{\lambda}_i$ is not of multiplicity one, $v_i$ can be any of its eigenvectors). For any $0 < \epsilon \leq \delta$, it may depends on $M$ and $N$, with probability at least $1 - Ce^{-c_1N\epsilon^2 + c_2M}$, where $c_1$, $c_2$ and $C$ are constants depending only on $\delta$ and $B$, we have

1. We denote the projection of $v_i$ on the eigenspace of nonzero eigenvalues of $P_N$ by $\tilde{v}_i$, which is the projection of $v_i$ on the first $M$ coordinates. The norm of $\tilde{v}_i$ satisfies

\[
|\tilde{v}_i|^2 + \frac{1}{\theta_i + \theta_i^2T_N^{-1}(\frac{1}{\theta_i})T_N'(T_N^{-1}(\frac{1}{\theta_i}))} \leq \epsilon C_{B,\delta}, \tag{25}
\]

2. The projection $\tilde{v}_i$ is an approximate eigenvector of $\tilde{P}_N$ with eigenvalue $\theta_i$,

\[
|\tilde{P}_N\tilde{v}_i - \theta_i\tilde{v}_i| \leq \epsilon C_{B,\delta}. \tag{26}
\]

where $C_{B,\delta}$ is a constant depending on $\delta$ and $B$.

**Proof.** In this case we have the following two characterizing equations for $\tilde{v}_i$,

\[
\begin{align*}
(I_M - \tilde{U}^*(\tilde{\lambda}_i - H_N)^{-1}H_N\tilde{U})\tilde{v}_i &= 0, \\
\tilde{v}_i^*\tilde{P}_N(\tilde{U}^*H_N(\tilde{\lambda}_i - H_N)^{-2}H_N\tilde{U})\tilde{P}_N\tilde{v}_i &= 1.
\end{align*}
\]

Approximately, by Proposition A.1, we have $\tilde{U}^*(\tilde{\lambda}_i - H_N)^{-1}H_N\tilde{U} \approx \frac{1}{\theta_i}$ and $\tilde{U}^*H_N(\tilde{\lambda}_i - H_N)^{-2}H_N\tilde{U} \approx \frac{1}{\theta_i} - T_N^{-1}(\frac{1}{\theta_i})T_N'(T_N^{-1}(\frac{1}{\theta_i}))$. The same argument as in Theorem 2.7 leads to (25) and (26). \[ \square \]

### 3.3. Wishart Case.

In this section we study the eigenvalues of perturbed Wishart matrices,

\[
\tilde{W}_N = W_N(I_N + P_N).
\]

It is well known that the empirical eigenvalue distribution of $W_N$ converges to the famous Marchenko-Pastur law with density

\[
\rho_{mp}(x) = \frac{\sqrt{(\gamma_+ - x)(x - \gamma_-)}}{2\pi \phi x}, \quad \text{for } x \in [\gamma_-, \gamma_+],
\]
where \( \gamma_- = (1 - \sqrt{\phi})^2 \) and \( \gamma_+ = (1 + \sqrt{\phi})^2 \). We denote its T-transform as
\[
T_{mp}(z) = \frac{z - \phi - 1 - \text{sgn}(z - \gamma_+)\sqrt{(z - \gamma_-)(z - \gamma_+)}}{2\phi}, \quad \text{for } z \in (-\infty, \gamma_-) \cup (\gamma_+, +\infty).
\]

In the following we fix a small universal constant \( \delta > 0 \). It is well known that with high probability the empirical eigenvalue distribution of Wishart matrix \( W_N \) is supported on \((\gamma_- - \frac{\delta}{2}, \gamma_+ + \frac{\delta}{2})\). We concentrate on the outliers of \( \tilde{W}_N \), which are eigenvalues on the spectral domain \((\gamma_- - \delta, \gamma_+ + \delta)^c\).

**Theorem 3.3.** Wishart matrix \( W_N \) and perturbation \( P_N \) are defined as in Section 3.1, we assume that the norm of \( P_N \) is bounded by some universal constant \( B \), i.e. \( \|P_N\| \leq B \). We denote the eigenvalues of \( \tilde{W}_N = W_N(I_N + P_N) \) by \( \lambda_1(\tilde{W}_N) \geq \lambda_2(\tilde{W}_N) \geq \cdots \geq \lambda_N(\tilde{W}_N) \).

For any \( \sqrt{\frac{M}{N}} \ll \epsilon \leq \delta \) (it may depends on \( M \) and \( N \)), and any \( 1 \leq i \leq M \), if \( \theta_i \) satisfies the following separation condition
\[
|\theta_i| \geq \sqrt{\phi} + 2\delta
\]
then it will create an outlier \( \tilde{\lambda}_i \) of \( \tilde{W}_N \), where \( \tilde{\lambda}_i = \lambda_i(\tilde{W}_N) \) for \( \theta_i > 0 \), and \( \tilde{\lambda}_i = \lambda_{N-M+i}(\tilde{W}_N) \) for \( \theta_i < 0 \), with high probability,
\[
\mathbb{P}\left( \phi + 1 + \theta_i + \frac{\phi}{\theta_i} - \epsilon \leq \tilde{\lambda}_i \leq \phi + 1 + \theta_i + \frac{\phi}{\theta_i} + \epsilon \right) 
\geq 1 - C \left( e^{-c_1 N \epsilon^2 + c_2 M} + \frac{1}{(N - M)^{c_3 \ln \ln(N-M)}} \right)
\]
where \( c_1, c_2, c_3 \) and \( C \) are constants depending only on \( \delta \) and \( B \).

**Proof.** The proof is quite similar to the Wigner case (Theorem 2.8). We prove the result for the outliers on the interval \([\gamma_+ + \delta, +\infty)\). We decompose the matrix \( X_N \) into two submatrices
\[
X_N = \begin{pmatrix} X_N^{(1)} \\ X_N^{(2)} \end{pmatrix},
\]
where \( X_N^{(1)} \) is an \( M \times p \) submatrix, \( X_N^{(2)} \) is an \((N - M) \times p \) submatrix. If \( z - X_N^{(2)} X_N^{(2)*} \) is invertible, we have
\[
\det(z - W_N(I_N + P_N)) = \det(z - X_N^{(2)} X_N^{(2)*}) 
\times \det(z - X_N^{(1)} \left( I_p + X_N^{(2)*}(z - X_N^{(2)} X_N^{(2)*})^{-1} X_N^{(2)} \right) X_N^{(1)*} \left( I_M + \tilde{P}_N \right)).
\]
Since \( X_N^{(1)} \) and \( X_N^{(2)} \) are independent, we can go through the same proof as in Theorem 3.1 once we condition on \( X_N^{(2)} \). Define the event,
\[
A_N = \left\{ \gamma_- - \frac{\delta}{2} \leq \|X_N^{(2)} X_N^{(2)*}\| \leq \gamma_+ + \frac{\delta}{2}, \sup_{z \in (\gamma_- - \delta, \gamma_+ + \delta)} \left| T_N^{(2)}(z) - T_{mp}(z) \right| \leq \frac{\epsilon}{2} \right\}
\]
where \( T_N^{(2)} \) is the T-transform of \( X_N^{(2)} X_N^{(2)*} \). From the rigidity results of eigenvalues of covariance matrices [9], for \( \epsilon \gg \sqrt{\frac{M}{N}} \) we have
\[
\mathbb{P}(A_N) \geq 1 - \frac{C}{(N - M)c_3 \ln \ln(N - M)},
\]
where \( c_3 \) and \( C \) depend on \( \delta \). For any \( X_N^{(2)} \in A_N \), conditional on \( X_N^{(2)} \), for any \( z \in [\gamma_+ + \delta, +\infty) \), \( z - X_N^{(2)} X_N^{(2)*} \) is invertible. Therefore the outliers of \( \tilde{W}_N \) are the solutions of
\[
\det \left( z - X_N^{(1)} \left( I_p + X_N^{(2)*} (z - X_N^{(2)} X_N^{(2)*})^{-1} X_N^{(2)} \right) X_N^{(1)*} \left( I_M + \tilde{P}_N \right) \right) = 0.
\]
By Proposition A.2, we have
\[
X_N^{(1)} \left( I_p + X_N^{(2)*} (z - X_N^{(2)} X_N^{(2)*})^{-1} X_N^{(2)} \right) X_N^{(1)*} \approx 1 + \phi T_N^{(2)}(z) \approx 1 + \phi T_{mp}(z).
\]
which gives us (28). □

Similarly as the Wigner case, we have the following corollary on the empirical distribution of outliers of perturbed Wishart matrices.

**Corollary 3.4.** Wishart ensemble \( W_N \) and perturbation \( P_N \) are defined as in Section 3.1, we assume the norm of \( P_N \) is bound by some universal constant \( B \), i.e. \( \| P_N \| \leq B \). Moreover we assume that the empirical distribution of the nonzero eigenvalues of \( P_N \) converges weakly to a compactly supported measure \( \nu \),
\[
\hat{\nu}_N = \frac{1}{M} \sum_{i=1}^{M} \delta_{\theta_i} \rightarrow \nu.
\]
with support supp \( \mu \subset [a, b] \), and \( a > \sqrt{\phi} \) (notice here \( \theta_i \)'s depend on \( N \)). We denote the eigenvalues of \( \tilde{W}_N = W_N (I_N + P_N) \) as \( \lambda_1(\tilde{W}_N) \geq \lambda_2(\tilde{W}_N) \geq \cdots \geq \lambda_N(\tilde{W}_N) \). Almost surely, the empirical distribution of the largest \( M \) eigenvalues of \( \tilde{W}_N \) converges weakly to the push forward measure \( \gamma_{\#\mu} \),
\[
\frac{1}{M} \sum_{i=1}^{M} \sigma_{\lambda_i(\tilde{W}_N)} \rightarrow \gamma_{\#\mu}, \quad a.s.
\]
where \( \gamma(\theta) = \phi + 1 + \theta + \frac{\phi}{T} \).

**Appendix A. Concentration of Measure**

**Proposition A.1.** \( A \) is a deterministic \( N \times N \) matrix. \( \tilde{U} \) is the first \( M \) columns of an \( N \times N \) orthogonal (or unitary) matrix following Haar measure. For any \( \xi > 0 \), we have
\[
\mathbb{P} \left( \left\| \tilde{U}^* A \tilde{U} - \frac{\text{Tr} A}{N} I_M \right\| \leq \xi \right) \geq 1 - Ce^{-cN\xi^2 + M \ln 7},
\]
where \( c \) and \( C \) depend on the norm of \( A \).
Proof. We will prove the orthogonal case, the unitary case is exactly the same. Denote \( \tilde{A} = A - \frac{1}{N} \mathbb{A} / N \), then \( \text{Tr} \tilde{A} = 0 \). (31) is equivalent to show that \( \| \tilde{U}^* \tilde{A} \| \leq \xi \) with high probability.

The proof follows from the standard epsilon net argument. We refer to [24] and [25] for a detailed discussion on epsilon net argument. For any \( v \in S^{M-1} = \{x \in \mathbb{R}^M : |x| = 1\} \), \( \tilde{U} v \) is uniformly distributed on the sphere \( S^{N-1} \). From the application of a well-known concentration of measure result, we have

\[
\mathbb{P} \left( \left| v^* \tilde{U}^* \tilde{A} \tilde{U} v \right| \geq \frac{\xi}{3} \right) = \mathbb{P} \left( \left| v^* \tilde{U}^* A \tilde{U} v - \mathbb{E} \left[ v^* \tilde{U}^* A \tilde{U} v \right] \right| \geq \frac{\xi}{3} \right) \leq C e^{-cN \xi^2}. \tag{32}
\]

for any \( \xi > 0 \), where \( c \) and \( C \) depends on the norm of \( A \).

Take \( \Sigma \) be a maximal \( \frac{1}{4} \)-net of the sphere \( S^{M-1} \), i.e. a set of points in \( S^{M-1} \) that are separated from each other by a distance of at least \( \frac{1}{3} \), and which is maximal with respect to set inclusion. The volume argument gives us that \( \Sigma \) has cardinality

\[
|\Sigma| \leq (1 + \frac{1}{2} \times \frac{1}{3})^M = 7^M.
\]

For any \( v \in S^{M-1} \), there exists some \( w \in \Sigma \) such that \( |v - w| \leq \frac{1}{3} \). Then we have,

\[
\left| v^* \tilde{U}^* \tilde{A} \tilde{U} v \right| \leq \left| w^* \tilde{U}^* \tilde{A} \tilde{U} w \right| + \left| (v - w)^* \tilde{U}^* \tilde{A} \tilde{U} (v - w) \right| \\
\leq \left| w^* \tilde{U}^* \tilde{A} \tilde{U} w \right| + \frac{2}{3} \| \tilde{U}^* \tilde{A} \tilde{U} \|.
\]

Therefore, we can replace the sphere \( S^{M-1} \) by its \( \frac{1}{2} - net \), and get

\[
\| \tilde{U}^* \tilde{A} \tilde{U} \| = \sup_{v \in S^M} \left| v^* \tilde{U}^* \tilde{A} \tilde{U} v \right| \leq 3 \sup_{v \in \Sigma} \left| v^* \tilde{U}^* \tilde{A} \tilde{U} v \right|. \tag{33}
\]

Combining (32) and (33), we have

\[
\mathbb{P}(\| v^* \tilde{U}^* \tilde{A} \tilde{U} v \| \geq \xi) \leq \mathbb{P} \left( \bigcup_{v \in \Sigma} \left| v^* \tilde{U}^* \tilde{A} \tilde{U} v \right| \geq \frac{\xi}{3} \right) \\
\leq \sum_{v \in \Sigma} \max_{v \in \Sigma} \left\{ \mathbb{P} \left( \left| v^* \tilde{U}^* \tilde{A} \tilde{U} v \right| \geq \frac{\xi}{3} \right) \right\} \\
\leq C e^{-cN \xi^2 + M \ln 7}.
\]

This finishes the proof. \( \square \)

**Proposition A.2.** \( B = \frac{1}{\sqrt{N}} [b_{ij}]_{1 \leq i \leq N}^{1 \leq j \leq M} \) is an \( N \times M \) random matrix. \( b_{ij} \)'s are i.i.d. random variables such that

\[ \mathbb{E}[b_{ij}] = 0, \quad \mathbb{E}[b_{ij}^2] = 1. \]
Moreover, for technical reason, we assume that $b_{ij}$'s are sub-gaussian with $\|b_{ij}\|_\psi^2 \leq K$. $A$ is a deterministic $N \times N$ matrix, then for $\xi > 0$, we have
\[
P \left( \left\| B^*AB - \frac{\text{Tr} A}{N} I_M \right\| \leq \xi \right) \geq 1 - 2e^{-cN\xi^2 + M\ln 7}
\]
where $c$ depends on $K$ and the norm $\|A\|$.

**Proof.** We can go through the same argument as in Proposition A.1, if we can prove analogue concentration of measure inequality as (32) in the i.i.d. setting. In fact we have, for any $v = (v_1, v_2, \cdots, v_M)^t \in S^{M-1}$,
\[
P \left( |v^*B^*A Bv - E[v^*B^*A Bv]| \geq \frac{\xi}{3} \right) \leq 2e^{-cN\xi^2}.
\]
(34)

Denote $u = Bv$, then $i$-th entry $u_i$ of $u$ is also sub-gaussian
\[
E[e^{tu_i}] = E\left[e^{t \sum_{j=1}^M b_{ij} v_j}\right] = \prod_{j=1}^M E[e^{t b_{ij} v_j}] \leq \prod_{j=1}^M \left[E[e^{t^2 v_j^2 K^2}]\right] = e^{t^2 K^2},
\]
with $\|u_i\|_\psi^2 \leq K$. Then (34) follows from Hanson-Wright inequality. We refer to [22] for a modern proof of Hanson-Wright inequality.

\[\square\]

**Proposition A.3.** $W_N = [w_{ij}]_{1 \leq i, j \leq N}$ is an $N \times N$ random matrix. $w_{ij}$'s are i.i.d. random variables such that
\[
E[w_{ij}] = 0, \quad E[w_{ij}^2] = 1.
\]
Moreover, for technical reason we assume that $w_{ij}$'s are sub-gaussian with $\|w_{ij}\|_\psi^2 \leq K$. Then for any $\xi > 0$, we have the following bound on the norm of $W_N$,
\[
P \left( \|W_N\| \geq \xi \right) \leq 14e^{-c\xi^2 + N\ln 7}
\]
where $c$ depends on $K$.

**Proof.** This kind of bound can also be proved by the epsilon-net argument. For any $v = (v_1, v_2, \cdots, v_N)^t \in S^N$,
\[
P \left( |v^*W_Nv| \geq \xi \right) = \mathbb{P} \left( \left| \sum_{i \leq j} (1 + \delta_{ij}) v_i v_j w_{ij} \right| \geq \xi \right) \leq 2e^{-c\xi^2}.
\]
(35)

Since here $\left| \sum_{i \leq j} (1 + \delta_{ij}) v_i v_j w_{ij} \right|_\psi^2 \leq \sqrt{2K}$. Therefore by the same epsilon-net argument as in Proposition A.1, we get
\[
P \left( \|W_N\| \geq \xi \right) \leq \mathbb{P} \left( \bigcup_{v \in \Sigma} |v^*W_Nv| \geq \frac{\xi}{3} \right) \leq \sum_{v \in \Sigma} \max \left\{ \mathbb{P} \left( |v^*W_Nv| \geq \xi \right) \right\} \leq 2e^{-c\xi^2 + \ln 7N}
\]
\[\square\]
Appendix B. Bounds on Stieltjes Transform and T-Transform

Proposition B.1. $H_N$ and $P_N$ are defined as in Section (2.1), such that their norms are bounded by $B$. $m_N$ is the Stieltjes transform of the empirical measure of $H_N$. For any $|\xi| \leq \delta$, we have

\[
\frac{|\xi|}{4B^2} \leq \left| m_N \left( m_N^{-1} \left( \frac{1}{\theta_i} \right) + \xi \right) - \frac{1}{\theta_i} \right| \leq \frac{4|\xi|}{\delta^2} \tag{36}
\]

\[
\frac{|\xi|}{8B^3} \leq \left| m_N' \left( m_N^{-1} \left( \frac{1}{\theta_i} \right) + \xi \right) - m_N' \left( m_N^{-1} \left( \frac{1}{\theta_i} \right) \right) \right| \leq \frac{8|\xi|}{\delta^3} \tag{37}
\]

where $\theta_i$ are eigenvalues of $P_N$ such that $m_N^{-1} \left( \frac{1}{\theta_i} \right) \geq \lambda_1(H_N) + 2\delta$, or $m_N^{-1} \left( \frac{1}{\theta_i} \right) \leq \lambda_N(H_N) - 2\delta$.

$H_N$ and $P_N$ are defined as in Section (3.1), such that their norms are bounded by $B$. $T_N$ is the Stieltjes transform of the empirical measure of $H_N$. For any $|\xi| \leq \delta$, we have

\[
\frac{|\xi|}{4B^2} \leq \left| T_N \left( T_N^{-1} \left( \frac{1}{\theta_i} \right) + \xi \right) - \frac{1}{\theta_i} \right| \leq \frac{4B|\xi|}{\delta^2} \tag{38}
\]

\[
\frac{|\xi|}{8B^3} \leq \left| T_N' \left( T_N^{-1} \left( \frac{1}{\theta_i} \right) + \xi \right) - T_N' \left( T_N^{-1} \left( \frac{1}{\theta_i} \right) \right) \right| \leq \frac{8B|\xi|}{\delta^3} \tag{39}
\]

where $\theta_i$ are eigenvalues of $P_N$ such that $T_N^{-1} \left( \frac{1}{\theta_i} \right) \geq \lambda_1(H_N) + 2\delta$, or $T_N^{-1} \left( \frac{1}{\theta_i} \right) \leq \lambda_N(H_N) - 2\delta$.

Proof. We will only prove (36) for positive eigenvalues of $P_N$, the other inequalities can be proved in the same way. On the interval $(\lambda_1(H_N) + \delta, +\infty)$, $m_N(z)$ is positive and strictly decreasing. By Taylor expansion, there exists some $\gamma \in [0, 1]$ such that

\[
\left| m_N \left( m_N^{-1} \left( \frac{1}{\theta_i} \right) + \xi \right) - \frac{1}{\theta_i} \right| = |\xi| m_N' \left( m_N^{-1} \left( \frac{1}{\theta_i} \right) + \gamma \xi \right)
\]

\[
= \frac{|\xi|}{N} \sum_{k=1}^{N} \frac{1}{\left( m_N^{-1} \left( \frac{1}{\theta_i} \right) + \gamma \xi - \lambda_k(H_N) \right)^2}
\]

\[
\geq \frac{|\xi|}{4N} \sum_{k=1}^{N} \frac{1}{\left( m_N^{-1} \left( \frac{1}{\theta_i} \right) - \lambda_k(H_N) \right)^2} \tag{40}
\]

\[
\geq |\xi| \left( \frac{1}{2N} \sum_{k=1}^{N} \frac{1}{m_N^{-1} \left( \frac{1}{\theta_i} \right) - \lambda_k(H_N)} \right)^2 \tag{41}
\]

\[
= \frac{|\xi|}{(4\delta_i)^2} \geq \frac{|\xi|}{4B^2}
\]
For the inequality (40), we use the fact that \( m_N^{-1}(\frac{1}{\theta_i}) - \lambda_k(H_N) \geq \delta + |\xi| \). (41) is from AM-GM inequality. For the upper bound,

\[
\left| m_N(m_N^{-1}(\frac{1}{\theta_i}) + \xi) - \frac{1}{\theta_i} \right| \leq \frac{4|\xi|}{N} \sum^{N}_{k=1} \frac{1}{m_N^{-1}(\frac{1}{\theta_i}) - \lambda_k(H_N)}^2 \leq \frac{4|\xi|}{\delta^2}.
\]

For the T-transform, although we can not directly lower bound expressions like (41) by AM-GM inequality. We still have the following bound from Cauchy inequality

\[
\left( \frac{1}{N} \sum^{N}_{k=1} \lambda_k(H_N) \right) \left( \frac{1}{N} \sum^{N}_{k=1} \frac{\lambda_k(H_N)}{T_N^{-1}(\frac{1}{\theta_i}) - \lambda_k(H_N)} \right)^t \geq \left( \frac{1}{N} \sum^{N}_{k=1} \frac{\lambda_k(H_N)}{T_N^{-1}(\frac{1}{\theta_i}) - \lambda_k(H_N)} \right)^2 = \frac{1}{(\theta_i)^2} \geq \frac{1}{B^2}.
\]

Rearrange the above inequality we get

\[
\frac{1}{N} \sum^{N}_{k=1} \frac{\lambda_k(H_N)}{T_N^{-1}(\frac{1}{\theta_i}) - \lambda_k(H_N)} \geq \frac{1}{B^{t+1}}, \quad t = 2, 3.
\]

Acknowledgement

The author expresses grateful thanks to Prof. Alice Guionnet, for her introducing to this problem and helpful discussions throughout the research process.

References

[1] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. An introduction to random matrices. Vol. 118. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2010, pp. xiv+492.

[2] Zhidong Bai and Jian-feng Yao. “Central limit theorems for eigenvalues in a spiked population model”. In: Ann. Inst. Henri Poincaré Probab. Stat. 44.3 (2008), pp. 447–474.

[3] Jinho Baik, Gérard Ben Arous, and Sandrine Péché. “Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices”. In: Ann. Probab. 33.5 (2005), pp. 1643–1697.

[4] Jinho Baik and Jack W. Silverstein. “Eigenvalues of large sample covariance matrices of spiked population models”. In: J. Multivariate Anal. 97.6 (2006), pp. 1382–1408.

[5] Serban Teodor Belinschi et al. “Outliers in the spectrum of large deformed unitarily invariant models”. In: (2012). url: http://arxiv.org/abs/1207.5443.
REFERENCES

[6] Florent Benaych-Georges, Alice Guionnet, and Mylène Maïda. “Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices”. In: Electron. J. Probab. 16 (2011), no. 60, 1621–1662.

[7] Florent Benaych-Georges, Alice Guionnet, and Mylène Maïda. “Large deviations of the extreme eigenvalues of random deformations of matrices”. In: Probab. Theory Related Fields 154.3-4 (2012), pp. 703–751.

[8] Florent Benaych-Georges and Raj Rao Nadakuditi. “The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices”. In: Adv. Math. 227.1 (2011), pp. 494–521.

[9] Alex Bloemendal et al. “Isotropic local laws for sample covariance and generalized Wigner matrices”. In: Electron. J. Probab. 19 (2014), no. 33, 53.

[10] Mireille Capitaine, Catherine Donati-Martin, and Delphine Féral. “Central limit theorems for eigenvalues of deformations of Wigner matrices”. In: Ann. Inst. Henri Poincaré Probab. Stat. 48.1 (2012), pp. 107–133.

[11] Mireille Capitaine, Catherine Donati-Martin, and Delphine Féral. “The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations”. In: Ann. Probab. 37.1 (2009), pp. 1–47.

[12] Mireille Capitaine et al. “Free convolution with a semicircular distribution and eigenvalues of spiked deformations of Wigner matrices”. In: Electron. J. Probab. 16 (2011), no. 64, 1750–1792.

[13] László Erdős, Horng-Tzer Yau, and Jun Yin. “Rigidity of eigenvalues of generalized Wigner matrices”. In: Adv. Math. 229.3 (2012), pp. 1435–1515.

[14] Fumio Hiai and Dénes Petz. “Asymptotic freeness almost everywhere for random matrices”. In: Acta Sci. Math. (Szeged) 66.3-4 (2000), pp. 809–834.

[15] Antti Knowles and Jun Yin. “The outliers of a deformed Wigner matrix”. In: Ann. Probab. 42.5 (2014), pp. 1980–2031.

[16] Ji Oon Lee and Kevin Schnelli. “Local deformed semicircle law and complete delocalization for Wigner matrices with random potential”. In: J. Math. Phys. 54.10 (2013), pp. 103504, 62.

[17] Ji Oon Lee et al. “Bulk Universality for Deformed Wigner Matrices”. In: (2014). URL: http://arxiv.org/abs/1405.6634.

[18] Sandrine Péché. “Erratum: “The largest eigenvalue of small rank perturbations of Hermitian random matrices” [Probab. Theory Related Fields 134 (2006), no. 1, 127–173; MR2221787]”. In: Probab. Theory Related Fields 134.1 (2006), p. 174.

[19] Sandrine Péché. “The largest eigenvalue of small rank perturbations of Hermitian random matrices”. In: Probab. Theory Related Fields 134.1 (2006), pp. 127–173.

[20] Alessandro Pizzo, David Renfrew, and Alexander Soshnikov. “On finite rank deformations of Wigner matrices”. In: Ann. Inst. Henri Poincaré Probab. Stat. 49.1 (2013), pp. 64–94.

[21] David Renfrew and Alexander Soshnikov. “On finite rank deformations of Wigner matrices II: Delocalized perturbations”. In: Random Matrices Theory Appl. 2.1 (2013), pp. 1250015, 36.
[22] Mark Rudelson and Roman Vershynin. “Hanson-Wright inequality and sub-Gaussian concentration”. In: Electron. Commun. Probab. 18 (2013), no. 82, 9.

[23] Terence Tao. “Outliers in the spectrum of iid matrices with bounded rank perturbations”. In: Probab. Theory Related Fields 155.1-2 (2013), pp. 231–263.

[24] Terence Tao. Topics in random matrix theory. Vol. 132. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2012, pp. x+282. ISBN: 978-0-8218-7430-1.

[25] Roman Vershynin. “Introduction to the non-asymptotic analysis of random matrices”. In: Compressed sensing. Cambridge Univ. Press, Cambridge, 2012, pp. 210–268.

[26] Dan Voiculescu. “Limit laws for random matrices and free products”. In: Invent. Math. 104.1 (1991), pp. 201–220.

Department of Mathematics, Harvard University

E-mail address: jiaoyang@math.harvard.edu