DEFORMED SINGLE RING THEOREMS

CHING-WEI HO AND PING ZHONG

Abstract. Given a sequence of deterministic matrices $A = A_N$ and a sequence of deterministic nonnegative matrices $\Sigma = \Sigma_N$ such that $A \to a$ and $\Sigma \to \sigma$ in $*$-distribution for some operators $a$ and $\sigma$ in a finite von Neumann algebra $A$. Let $U = U_N$ and $V = V_N$ be independent Haar-distributed unitary matrices. We prove that, under mild assumptions, the empirical eigenvalue distribution of $U\Sigma V + A$ converges to the Brown measure of $u\sigma + a$, where $u$ is a Haar unitary operator in $A$, and $a, u, \sigma$ are all freely independent. We show that these assumptions are satisfied when $A$ is Hermitian, unitary, or the $N \times N$ Jordan block matrix. By putting $A = 0$, our result removes a regularity assumption in the single ring theorem by Guionnet, Krishnapur and Zeitouni. We also prove a local convergence on optimal scale, extending the result of Bao, Erdős and Schnelli. In the sequel, under the assumptions on $A$ and $\Sigma$, we prove an estimate for the least singular value of $U\Sigma V + A - \lambda$, for all $\lambda \in \mathbb{C}$ except for $\lambda$ in a set with small Lebesgue measure.

Contents

1. Introduction 2
1.1. Sum of two random matrices 2
1.2. The random matrix model and the main results 3
2. Preliminaries and the limit distribution 7
2.1. Free additive convolution 7
2.2. The Brown measure and limit distribution 8
2.3. A result about the random matrix model 10
2.4. Approximate subordination functions 12
3. Least singular value estimate 13
3.1. First case: under $E_{\text{denom}} = \{\|rLv - \varepsilon L\Sigma_{(1,1)}Z\|_2 \leq K\}$ 16
3.2. Dividing $E_{\text{denom}}$ into cases 17
3.3. Proof of Theorem 3.2 20
4. The deformed local single ring theorem 21
4.1. Convergence of Cauchy transforms in the bulk 21
4.2. Proof of the local convergence 23
5. The deformed single ring theorem 27
6. Example: the Jordan block matrix 33
Acknowledgments 35
References 36

Date: October 21, 2022.

2010 Mathematics Subject Classification. 60B20, 46L54.

Key words and phrases. Non-Hermitian random matrices, single ring theorem, free probability, Brown measure.
1. Introduction

1.1. Sum of two random matrices. The study of the eigenvalues of the sum of two matrices is an old problem. The Horn’s problem [32] studies the possible eigenvalues of the sum of two deterministic Hermitian matrices with prescribed eigenvalues. Horn’s conjecture [32] was solved by Knutson and Tao [34], and Knutson, Tao and Woodward [35]. In random matrix theory, it is a fundamental question to study the limiting eigenvalue distribution of the sum of two random matrices, one of which satisfies some symmetry assumptions. In the case where the two random matrices are Hermitian, if the empirical eigenvalue distribution of the two random matrices converges to some probability measures \( \mu \) and \( \nu \) on \( \mathbb{R} \), the limiting eigenvalue distribution of this sum is much more specific than the solutions of the Horn’s problem. In terms of the bulk distribution, the empirical eigenvalue distribution of the sum of the two Hermitian random matrices converges to the free convolution \( \mu \boxplus \nu \) of \( \mu \) and \( \nu \) (see [45] and [36, Chapter 4]). In fact, the seminal work of Voiculescu [45] suggests that free probability theory is a suitable framework to study the asymptotic spectrum of the sum of large independent Hermitian or non-Hermitian random matrices.

A well-studied non-Hermitian random matrix model is the single ring random matrix model defined as follows. Consider two sequences of independent Haar-distributed unitary matrices \( U = U_N \) and \( V = V_N \) and a sequence of deterministic nonnegative matrices \( \Sigma = \Sigma_N \) such that the eigenvalue distribution of \( \Sigma \) converges weakly to a certain probability measure on \( \mathbb{R} \). Free probability theory again provides natural limit operators for \( U_N \Sigma_N V_N^* \). By Voiculescu’s asymptotic freeness result [45] (see also [17]), the random matrices \( \{U_N \Sigma_N V_N^*\}_{N=1}^\infty \) converge in *-moments to an \( R \)-diagonal operator introduced by Nica–Speicher [37]. The Brown measure [16] is an analogue of the eigenvalue distribution of operators. The Brown measure of \( R \)-diagonal operators was calculated by Haagerup–Larsen [26] and Haagerup–Schultz [28] (see also [7, 48] for alternative proofs). The Brown measure of any \( R \)-diagonal operator is supported in an annulus. In the physics literature, Feinberg–Zee [22] proved the single ring theorem which states that the limiting eigenvalue distribution of the matrix \( U \Sigma V^* \) converges to a certain rotation-invariant probability measure whose support is an annulus or a disk. Guionnet, Krishnapur and Zeitouni [25] then proved the single ring theorem rigorously under some technical assumptions; one of the assumptions is removed in a later work of Rudelson–Vershynin [40].

In this article, in addition to the matrices \( U, V, \Sigma \) introduced previously, we consider another deterministic matrix \( A = A_N \) such that given any polynomial \( P \) in two noncommuting indeterminates, the limit

\[
\lim_{N \to \infty} \text{tr}[P(A, A^*)]
\]

exists, where \( \text{tr} = (1/N) \text{Tr} \) is the normalized trace on \( N \times N \) matrices. We then prove that, under mild assumptions, the empirical eigenvalue distribution of the random matrix

\[
Y = U \Sigma V^* + A
\]

converges to the Brown measure of a sum of two free random variables, one of which is \( R \)-diagonal. We call this result a deformed single ring theorem.

In the random matrix model described in the preceding paragraph, the deterministic matrix \( A \) and the random matrix \( U \Sigma V^* \) are asymptotically free in the following sense. Let \( (A, \tau) \) be a \( W^* \)-probability space where \( A \) is a finite von Neumann algebra equipped with a faithful, normal, tracial state \( \tau \). By [36] Theorem 9 on page 105, the matrix \( A \) and the random matrix \( U \Sigma V^* \) are asymptotically free in the sense that there exist \( a, T \in A \) that
are freely independent in the sense of Voiculescu [44] such that for any polynomial $P$ in four noncommuting indeterminates

$$\lim_{N \to \infty} \mathbb{E} \text{tr}[P(A, A^*, U\Sigma V^*, (U\Sigma V^*)^*)] = \tau[P(a, a^*, T, T^*)].$$

In other words, $Y = U\Sigma V^* + A$ converges in $*$-distribution to $y = T + a$, where $T$ is an $R$-diagonal operator. The Brown measure of $y$ is a natural candidate of the limiting eigenvalue distribution of $Y$. However, since the Brown measure is not continuous with respect to $*$-moments, knowing that $Y$ converges to $y$ in $*$-distribution does not mean the empirical eigenvalue distribution of $Y$ converges to the Brown measure of $y$. The purpose of this paper is to prove that the limiting eigenvalue distribution of $Y$ is indeed the Brown measure of $y$.

There is another non-Hermitian random matrix model closely related to the model that we consider in this paper. Ginibre [23] looked at the limiting eigenvalue distribution of the normalized random matrix with independent Gaussian entries with unit variance, now called the Ginibre ensemble. After decades of development (for example, [1]), Tao–Vu [43] showed that the limiting eigenvalue distribution of the Ginibre ensemble remains unchanged if the Gaussian entries are replaced by i.i.d. random variables with arbitrary distributions with unit variance. Śniady [42] studied the limiting eigenvalue distribution of the sum of a Ginibre ensemble and a deterministic matrix. Śniady identifies that the limiting eigenvalue distribution of the sum of an arbitrary matrix and the Ginibre ensemble is the Brown measure of the sum of two freely independent operators; one of these two operators is a Voiculescu’s circular operator, which is $R$-diagonal. While the random matrix model $Y$ being considered in this paper is different from the one Śniady considered, if the law of the $\Sigma$ in this paper converges to the quarter-circular law, the limiting eigenvalue distribution of $Y$ is the same as that of the model in Śniady’s paper. The Tao–Vu’s replacement principle [43, Corollary 1.8] shows that the empirical eigenvalue distribution of the sum of a random matrix with i.i.d. entries and a deterministic matrix converges to the same Brown measure. This Brown measure has an explicit formula computed explicitly in [13, 31, 30, 47].

In [19, 20], Erdős, Schlein and Yau investigated the local behavior of the Wigner random matrices, called a local law; they looked at the number of eigenvalues in an interval of length of order $\log N/N$ inside the bulk. Kargin [33] and Bao–Erdős–Schnelli [2] proved a local law of the sum of two large Hermitian random matrices, one of which is conjugated by a Haar-distributed unitary matrix. In the non-Hermitian framework, Bourgade, Yau and Yin [14, 15] and Yin [46] proved a local law for the normalized non-Hermitian random matrix model with i.i.d. entries. Benaych-Georges [8], later improved by Bao–Erdős–Schnelli [3] to an optimal scale, proved a local law for the random matrix $U\Sigma V^*$. They call this local law a local single ring theorem. In this paper, we also prove a deformed local single ring theorem, studying the local behavior of the eigenvalues of $Y = U\Sigma V^* + A$ in the bulk.

1.2. The random matrix model and the main results. We first introduce the random matrix model considered in this paper. Suppose that

1. $U = U_N$ and $V = V_N$ are independent Haar-distributed unitary matrices;
2. $\Sigma = \Sigma_N$ are nonnegative deterministic matrices such that the eigenvalue distribution of $\Sigma$ converges weakly to a probability measure on $\mathbb{R}$ that is not a Dirac delta measure at 0.
3. $A = A_N$ are deterministic matrices such that all the $*$-moments of $A$ converge;
(4) there is a constant $M > 0$ independent of $N$ such that
\[ \|\Sigma\|, \|A\| \leq M. \] (1.1)

In this paper, we focus on the random matrix
\[ Y = Y_N = U\Sigma V^* + A. \] (1.2)

To control the behavior of $Y$, we make an extra assumption. For an $N \times N$ matrix $L$, we denote the singular values of $L$ by
\[ s_1(L) \geq s_2(L) \geq \ldots \geq s_N(L). \]

**Assumption 1.1.** For any $\beta > 0$ and $\lambda \in \mathbb{C}$, let
\[ \gamma_1^\lambda(\beta) = \# \{ i : s_i(A - \lambda) \geq N^{-\beta} \} \quad \text{and} \quad \gamma_2(\beta) = \# \{ i : s_i(\Sigma) \geq N^{-\beta} \} \]
and denote
\[ \Gamma_N = \{ \lambda \in \mathbb{C} : \gamma_1^\lambda(\beta) + \gamma_2(\beta) < N \}. \]

We assume there exists $\alpha > 0$ such that the corresponding $\Gamma_N$ (which we suppress the dependence of $\alpha$) satisfies
\[ \lim_{N \to \infty} m(\Gamma_N) = 0, \]
where $m$ is the Lebesgue measure on the complex plane.

We emphasize that all the assumptions that we make for the random matrix $Y$ are on the deterministic matrices $\Sigma$ and $A$.

**Remark 1.2.** A large class of matrices $A$ satisfies this assumption with arbitrary nonnegative matrices $\Sigma$ whose eigenvalue distribution converges weakly to a probability measure on $\mathbb{R}$ that is not the Dirac delta measure $\delta_0$. Choose any $\alpha > 1/2$. If $A$ is normal, for any $\lambda \in \mathbb{C}$,
\[ s_N(A - \lambda) = \min_{\lambda_i} |\lambda - \lambda_i| \]
where $\lambda_i$ are the eigenvalues of $A$. This shows that $\gamma_1^\lambda(\alpha) = 0$ if $|\lambda - \lambda_i| > N^{-\alpha}$ for all $\lambda_i$. Therefore, we have $\gamma_1^\lambda(\alpha) = N$ and
\[ \Gamma_N \subseteq \bigcup_i B(\lambda_i, N^{-\alpha}) \]
where $B(a, r)$ denotes the disk centered at $a \in \mathbb{C}$ with radius $r > 0$. Therefore, Assumption 1.1 holds.

In Section 6, we will give an example of non-normal matrices $A$ which satisfies Assumption 1.1 for arbitrary nonnegative matrices $\Sigma$ whose eigenvalue distribution converges weakly to a probability measure on $\mathbb{R}$ that is not the Dirac delta measure $\delta_0$.

We then introduce the limiting object of $Y$. Let $(A, \tau)$ be a $W^*$-probability space and $a, \sigma \in A$ such that $A$ converges to $a$ and $\Sigma$ converges to $\sigma$ in $*$-distribution. Let $T \in A$ be an $R$-diagonal freely independent from $a$ such that the law of $|T|$ is the same as that of $\sigma$; hence, if $u \in A$ is a Haar unitary operator freely independent of $\sigma$, we have $T = u \sigma$ in $*$-distribution. Set
\[ y = T + a. \] (1.3)
By [36, Theorem 9 on Page 105], \( Y \) converges in \(*\)-distribution to \( y \); that is, for any polynomial \( P \) in two noncommuting indeterminates,

\[
\lim_{N \to \infty} \mathbb{E} \text{tr}[P(Y, Y^*)] = \tau[P(y, y^*)].
\]

We are interested in the limiting eigenvalue distribution of \( Y \). Since \( y \) is the limit in \(*\)-distribution of \( Y \), a natural candidate of the limiting eigenvalue distribution of \( Y \) is the Brown measure of \( y \), which we will define shortly. Note, however, that although \( y \) is the limit in \(*\)-distribution of \( Y \), it is not automatic that the Brown measure of \( y \) is the limiting eigenvalue distribution of \( Y \); indeed, there are examples of matrices whose limiting eigenvalue distribution does not converge to the Brown measure of their limit in \(*\)-distribution (see Section 6). We now define the Brown measure introduced by Brown [16].

**Definition 1.3** ([16]). Let \( x \in A \). The function \( h(\lambda) = \tau[\log |x - \lambda|] \) defined for \( \lambda \in \mathbb{C} \) is a subharmonic function on \( \mathbb{C} \). The Brown measure \( \mu_x \) of \( x \) is defined to be

\[
\mu_x = \frac{1}{2\pi} \Delta h
\]

where the Laplacian is taken in distributional sense. The function \( h \) turns out to be the logarithmic potential of the Brown measure of \( x \) in the sense that

\[
\tau[\log |x - \lambda|] = \int_{\mathbb{C}} \log |z - \lambda| d\mu_x(z).
\]

(1.4)

For simplicity, we call \( h \) the logarithmic potential of \( x \).

The definition of the Brown measure is analogous to the eigenvalue distribution of a matrix. If we replace \( x \) in Definition 1.3 by an \( N \times N \) matrix \( X \) and the tracial state \( \tau \) by the normalized trace \( \text{tr} \) of matrices, then by Section 11.2 of [36],

\[
\text{tr}[\log |X - \lambda|] = \frac{1}{N} \sum_{k=1}^{N} \log |\lambda_k - \lambda|,
\]

(1.5)

where \( \lambda_k \) are the eigenvalues of \( X \), and \((1/2\pi) \Delta \lambda \text{tr}[\log |X - \lambda|]\) is the eigenvalue distribution of \( X \). The function defined in (1.5) is called the logarithmic potential of the eigenvalue distribution of \( X \) of the logarithmic potential of \( X \).

We introduce the main theorems of this paper. For a probability measure \( \mu \) on \( \mathbb{R} \), we denote

\[
G_\mu(z) = \int \frac{1}{z - t} d\mu(t), \quad z \in \mathbb{C}^+
\]

(1.6)

to be the Cauchy transform of \( \mu \). We also write \( \tilde{\mu} \) to be the symmetrization of \( \mu \) defined by

\[
\tilde{\mu}(B) = \frac{1}{2}[\mu(B) + \mu(-B)]
\]

for all Borel set \( B \subset \mathbb{R} \). Given a Hermitian matrix \( X \) or a Hermitian operator \( x \), we write the law of \( X \) or \( x \) as \( \mu_X \) or \( \mu_x \) respectively. There is no notational ambiguity; the law \( \mu_X \) or \( \mu_x \) are the Brown measure of \( X \) or \( x \) respectively.

The Brown measure of \( T + a \) is calculated in a recent work by Bercovici and the second author [12]. Denote

\[
\Omega(T, a) = \{ \lambda \in \mathbb{C} : \|(a - \lambda)^{-1}\|_2 ||T||_2 > 1 \text{ and } ||a - \lambda||_2 ||T^{-1}||_2 > 1 \}.
\]

(1.7)

The Brown measure of \( T + a \) is supported in the closure of \( \Omega(T, a) \); see Theorem 2.3 for a review. The main result of this paper is the following; it is proved in Theorem 5.1.
Theorem 1.4. Suppose that the random matrix $Y$ and the operator $y \in A$ are as in (1.2) and (1.3). Assume Assumption 1.1 holds. Assume further that at least one of the following is true:

1. there exists a Lebesgue measurable set $E \subset \Omega_c$ that has Lebesgue measure 0 with the following property: for any compact set $S \subset \Omega_c \cap E^c$, there exist constants $\kappa_1, \kappa_2 > 0$ such that
   \[ |G_{\tilde{\mu}, A - \lambda}(i\eta)| \leq \kappa_2 \]
   for all $\eta > N^{-\kappa_1}$ and $\lambda \in S$; or
2. there exist constants $\kappa_1, \kappa_2 > 0$ such that
   \[ |G_{\tilde{\mu}, \Sigma}(i\eta)| \leq \kappa_2 \]
   for all $\eta > N^{-\kappa_1}$.

Then the empirical eigenvalue distribution of $Y$ converges weakly to the Brown measure of $y$ in probability.

A large class of matrices $A$ satisfies the condition in the above theorem. The following is proved in Theorem 5.2. By putting $A = 0$, the following theorem recovers the single ring theorem by Guionnet et al. [25] with weaker conditions, removing the regularity assumption that there exists $\kappa_1, \kappa_2 > 0$ such that
   \[ |\text{Im} G_{\Sigma}(z)| \leq \kappa_1, \quad \text{Im} z > N^{-\kappa_2}. \]
In the bulk regime, this regularity assumption has been removed; see [3, Remark 1.14]. Our result removes this regularity assumption also outside the bulk; for example, it applies to the simplest case where $A = 0$ and $\Sigma = I$. This completely solves the question in [25, Remark 2].

Theorem 1.5. Let $Y$ and $y$ be as in (1.2) and (1.3). If $A$ is Hermitian for all $N$ or if $A$ is unitary for all $N$, then the empirical eigenvalue distribution of $Y$ converges weakly to the Brown measure of $y$ in probability.

More generally, if $A$ is a normal matrix and there is a closed set $F$ with Lebesgue measure zero independent of $N$ such that all the eigenvalues of $A$ lie inside $F$ for all $N$, then the empirical eigenvalue distribution of $Y$ converges weakly to the Brown measure of $y$ in probability.

The empirical eigenvalue distribution of $Y$ is the distributional Laplacian of
   \[ (1/2\pi) \text{tr}[\log |Y - \lambda|] \]
with respect to $\lambda$. To show the weak convergence of the empirical eigenvalue distribution of $Y$ to the Brown measure of $y$, we need to show that, for any test function $f \in C_c^\infty(\mathbb{C})$,
   \[ \int \Delta f(\lambda) \text{tr}[\log |Y - \lambda|] d^2 \lambda \rightarrow \int \Delta f(\lambda) \tau[\log |y - \lambda|] d^2 \lambda \]
in probability. Note that $\text{tr}[\log |Y - \lambda|]$ is the average of the logarithm of the singular values of $Y - \lambda$. Since the logarithm is unbounded around 0, we need to estimate the least singular value of $Y - \lambda$ for all $\lambda \in \mathbb{C}$ (with limited exceptions) in order to control $\text{tr}[\log |Y - \lambda|]$. For $\lambda \in \Gamma_N$, the least singular value of $Y - \lambda$, denoted by $s_{\min}(Y - \lambda)$, is estimated in Section 3. This estimate depends on the $\alpha$ in Assumption 1.1. The following theorem follows from Theorem 3.2.
Theorem 1.6. For some positive constants $c$ and $c'$ independent of $N,$
\[ P(s_{\min}(Y - \lambda) \leq t) \leq t^c N^{c'}, \quad t > 0 \]
for all $\lambda \in \Gamma_N.$

We also prove a deformed local single ring theorem in Theorem 1.2, which concerns the local behavior of the eigenvalues of $Y$ in the bulk. The following is a simplified version of Theorem 1.2, which also states the rate of the convergence. Instead of Assumption 1.1, we make a stronger assumption that $\Gamma_N$ has Lebesgue measure zero for all $N$.

Theorem 1.7. Let $Y$ and $y$ be as in (1.2) and (1.3). Assume that $\Gamma_N$ in Assumption 1.1 has Lebesgue measure zero for all $N.$ Let $a_N \in A$ be such that $a_N$ has the same $*$-distribution as $A_N.$ Also let $T_N \in A$ be an $R$-diagonal operator such that $\mu_{\Gamma_T} = \mu_{\Sigma_N}.$ Write
\[ y_N = a_N + T_N. \]

For any compact set $K \subset D(T, a),$ $\alpha \in (0, 1/2),$ $w_0 \in K$ and any $f : \mathbb{C} \to \mathbb{R}$ be a smooth function supported in the disk centered at 0 of radius $R,$ we define a function $f_{w_0}$ depending on $N$ by
\[ f_{w_0}(w) = N^{2\alpha} f(N^{\alpha}(w - w_0)). \]

Writing $\lambda_1, \ldots, \lambda_N$ to be the eigenvalues of $Y,$ we have
\[ \frac{1}{N} \sum_{k=1}^{N} f_{w_0}(\lambda_k) - \int_{\mathbb{C}} f_{w_0}(w) \, d\mu_{y_N}(w) \to 0 \]
in probability, where $\mu_{y_N}$ to be the Brown measure of $y_N.$ This convergence is uniform in $f$ and in $w_0 \in K,$ for all large enough $N$ depending on $K,$ $R,$ $M,$ $a$ and $T.$ (The constant $M$ is defined in (1.1).)

2. Preliminaries and the limit distribution

2.1. Free additive convolution. Given a probability measure $\mu$ on $\mathbb{R},$ its Cauchy transform $G_\mu$ is defined on the complex upper half-plane $\mathbb{C}^+$ as in (1.6). Let $F_\mu : \mathbb{C}^+ \to \mathbb{C}^+$ be an analytic map defined as
\[ F_\mu(z) = \frac{1}{G_\mu(z)}, \quad z \in \mathbb{C}^+. \]

Given two probability measures $\mu_1, \mu_2$ on $\mathbb{R},$ it is known that there exist a pair of analytic functions $\omega_1, \omega_2 : \mathbb{C}^+ \to \mathbb{C}^+,$ such that, for all $z \in \mathbb{C}^+$, we have
\[ F_{\mu_1 \boxplus \mu_2}(z) = F_{\mu_1}(\omega_1(z)) = F_{\mu_2}(\omega_2(z)) = \omega_1(z) + \omega_2(z) - z. \quad (2.1) \]

The free additive convolution obeys various regularity properties. Suppose neither of $\mu_1, \mu_2$ is a point mass. Then $\mu_1 \boxplus \mu_2$ has no singular continuous part, and the absolutely continuous part of $\mu_1 \boxplus \mu_2$ is always nonzero, and its density is analytic whenever positive and finite. See [6, 9, 11] for more details.

Recall that, the symmetrization $\tilde{\mu}$ of a probability measure $\mu$ on $\mathbb{R}$ is defined by
\[ \tilde{\mu}(B) := \frac{1}{2} [\mu(B) + \mu(-B)] \]
for any Borel set $B \subset \mathbb{R}.$ Set
\[ \mu_{\sigma, \xi} := \tilde{\mu}_{\sigma} \boxplus \tilde{\mu}_{\xi}, \quad (2.2) \]
where ⊞ denotes the free additive convolution of probability measures on \( \mathbb{R} \). For any symmetric probability measure \( \mu \) on \( \mathbb{R} \), observe that, by the symmetry,

\[
G_\mu(i\eta) = \int_{\mathbb{R}} \frac{1}{i\eta - x} d\mu(x) = -i\eta \int_{\mathbb{R}} \frac{1}{\eta^2 + x^2} d\mu(x), \quad \eta > 0.
\]

Hence, \( F_\mu(i\eta) \in i(0, \infty) \). We can deduce that the subordination functions \( \omega_1, \omega_2 \) corresponding to the free convolution of two symmetric probability measures \( \tilde{\mu}_\sigma, \tilde{\mu}_\xi \) satisfy

\[
\omega_1(i\eta), \omega_2(i\eta) \in i(0, \infty)
\]

for all \( \eta > 0 \). (See [12, Proposition 3.1])

2.2. The Brown measure and limit distribution. Recall that the pair \((A, \tau)\) denotes a \( W^* \)-probability space. Given a sequence of random matrix \( \{X_N\} \), we say that \( X_N \) converges in \( * \)-distribution (or \( * \)-moments) to \( x \in (A, \tau) \) if

\[
\lim_{N \to \infty} \mathbb{E} \text{tr}(P(X_N, X_N^*)) = \tau(P(x, x^*)),
\]

for any polynomial \( P \) in two noncommuting indeterminates with complex coefficients, where \( \text{tr} \) means the normalized trace \((1/N) \text{Tr} \).

Recall that \( U = U_N \) and \( V = V_N \) are two independent Haar random unitary matrices, and \( \Sigma = \Sigma_N \) is a sequence of \( N \times N \) deterministic nonnegative definite diagonal matrices, and \( A = A_N \) is a sequence of \( N \times N \) deterministic matrices. Denote the singular values of \( \Sigma \) by \( \sigma_1, \cdots, \sigma_N \) (which are also the eigenvalues of \( \Sigma \)), and the empirical distribution of the singular values of \( \Sigma \) by

\[
\mu_\Sigma = \frac{1}{N} \sum_{i=1}^{N} \delta_{|\sigma_i|}.
\]

We assume there is a nonnegative operator \( \sigma \in A \) such that

\[
\mu_\Sigma \to \mu_\sigma.
\]

Recall that \( A \) and \( USV^* \) converge in \( * \)-moments to \( a \) and \( T \) respectively, where \( T \) is \( R \)-diagonal and \( a \) is freely independent from \( T \). The Brown measure of \( T + a \) is a natural candidate of the empirical eigenvalue distribution of \( Y_N \), and our goal is to prove that this is indeed the case.

We first establish a result concerning the subordination functions for the free convolution of \( \mu_1 \) and \( \mu_2 \), where

\[
\mu_1 = \tilde{\mu}_{|\sigma - \lambda|} \quad \text{and} \quad \mu_2 = \tilde{\mu}_{|T|} = \tilde{\mu}_\sigma.
\]

The corresponding subordination functions depend on \( \lambda \). We consider them as two-variable functions and write them as

\[
\omega_1(\lambda, \cdot) \quad \text{and} \quad \omega_2(\lambda, \cdot).
\]

The proof depends on the continuity of Denjoy–Wolff points [29, 3] (See [41, Chapter 5] for an introduction to the Denjoy–Wolff theory).

Proposition 2.1. The subordination functions \( \omega_1 \) and \( \omega_2 \) have continuous extensions on \( \mathbb{C} \times (\mathbb{C}^+ \cup \mathbb{R}) \). In particular, the continuous extensions of \( \omega_j \) satisfy

\[
\omega_j(\lambda, 0) = \lim_{\eta \to 0} \omega_j(\lambda, i\eta), \quad \lambda \in \mathbb{C}.
\]
Proof. Consider the holomorphic function
\[ \varphi_{\lambda,z}(w) = F_{\mu_2}(F_{\mu_1}(w) - w + z) - (F_{\mu_1}(w) - w + z) + z, \quad w \in \mathbb{C}^+, \]
which is a self-map on the upper half plane. Since neither \( \mu_1 \) nor \( \mu_2 \) is a single delta mass, for any \( \lambda \in \mathbb{C}, z \in \mathbb{C}^+ \cup \mathbb{R}, \omega_1(\lambda, z) \) is the Denjoy–Wolff point of \( \varphi_{\lambda,z} \).

It is clear that \( \mu_1 = \tilde{\mu}_{|a-\lambda|} \) is continuous in distribution with respect to \( \lambda \). Therefore, if \( \lambda_n \to \lambda \) in \( \mathbb{C} \) and \( z_n \to z \) in \( \mathbb{C}^+ \cup \mathbb{R} \), we have \( \varphi_{\lambda_n,z_n} \to \varphi_{\lambda,z} \) pointwise. By Theorem 1.1 of [5], \( \omega_1(\lambda_n,z_n) \to \omega_1(\lambda,z) \). This shows continuity of \( \omega_1 \). The proof of the continuity of \( \omega_2 \) is similar. \( \Box \)

Proposition 2.2. Let \( a_N, \sigma_N \in A \) be such that \( \sigma_N \geq 0 \), \( a_N \to a \) and \( \sigma_N \to \sigma \) in \(*\)-distribution.

Given \( \lambda \in \mathbb{C} \), consider the free additive convolution of \( \tilde{\mu}_{|a_N-\lambda|} \) and \( \tilde{\mu}_{\sigma_N} \) and let \( \omega_1^{(N)}(\lambda, \cdot) \) and \( \omega_2^{(N)}(\lambda, \cdot) \) be the corresponding subordination functions. Then for any compact set \( K \subset \mathbb{C} \) and \( z \in \mathbb{C}^+ \cup \mathbb{R} \),
\[ (\omega_1^{(N)}(\lambda, z), \omega_2^{(N)}(\lambda, z)) \to (\omega_1(\lambda, z), \omega_2(\lambda, z)) \]
uniformly in \( \lambda \in K \).

Proof. If the conclusion does not hold, there exist \( \varepsilon_0 > 0 \) and a sequence \( \lambda_N \in K \) such that
\[ |\omega_1^{(N)}(\lambda_N, z) - \omega_1(\lambda_N, z)| \geq \varepsilon_0 \] (2.3)
or
\[ |\omega_2^{(N)}(\lambda_N, z) - \omega_2(\lambda_N, z)| \geq \varepsilon_0. \]
Without loss of generality, we assume the former holds. By extracting a subsequence of \( (\lambda_N) \) if necessary, we assume \( \lambda_N \to \lambda \) for some \( \lambda \in K \).

Write \( \mu_{1,N} = \tilde{\mu}_{|a_N-\lambda_N|} \) and \( \mu_{2,N} = \tilde{\mu}_{\sigma_N} \). Consider the holomorphic function
\[ \varphi_N(w) = F_{\mu_{2,N}}(F_{\mu_{1,N}}(w) - w + z) - (F_{\mu_{1,N}}(w) - w + z) + z, \quad w \in \mathbb{C}^+ \]
which is a self-map on the upper half plane. Define \( \varphi_{\lambda,z} \) as in the proof of Proposition 2.1. Then \( \varphi_N \to \varphi_{\lambda,z} \) pointwise. But then by Theorem 1.1 of [5], \( \omega_1^{(N)}(\lambda_N, z) \to \omega_1(\lambda, z) \), contradicting (2.3). \( \Box \)

Now we discuss the support of the Brown measure of \( T + a \). Recall that
\[ \Omega(T, a) = \{ \lambda \in \mathbb{C} : ||(a - \lambda)^{-1}||_2||T||_2 > 1 \quad \text{and} \quad ||a - \lambda||_2||T^{-1}||_2 > 1 \}. \] (2.4)
We set
\[ S(T, a) = \{ \lambda \in \mathbb{C} : \mu_1(\{0\}) + \mu_2(\{0\}) \geq 1 \}. \] (2.5)
Since neither \( \mu_1 \) nor \( \mu_2 \) is a single delta mass at zero, it follows that \( \lambda \in S(T,a) \) implies that 0 is an eigenvalue of \( |T| \) or \( |a - \lambda| \). We see that \( \lambda \) satisfies the defining conditions for \( \Omega(T, a) \). Hence, \( S(T, a) \subset \Omega(T, a) \).

Theorem 2.3. [12] For any \( \lambda \in \mathbb{C} \) and \( \eta > 0 \), \( \omega_j(\lambda, i\eta) \) is a purely imaginary number \( (j = 1, 2) \). The set \( S(T,a) \) consists of finitely many elements and we have
\[ \Omega(T,a) \setminus S(T,a) = \{ \lambda \in \mathbb{C} : |\omega_1(\lambda, 0)| \in (0, \infty) \} \]
\[ = \{ \lambda \in \mathbb{C} : |\omega_2(\lambda, 0)| \in (0, \infty) \} \]
\[ = \{ \lambda \in \mathbb{C} : 0 < f_{\mu_1 \# \mu_2}(0) < \infty \}, \] (2.6)
where \( f_{\mu_1 \boxplus \mu_2} \) denotes the density function of the absolutely continuous part of \( \mu_1 \boxplus \mu_2 \).

Moreover,

\[
S(T,a) = \{ \lambda \in \mathbb{C} : \omega_1(\lambda,0) = \omega_2(\lambda,0) = 0 \},
\]

and

\[
\mathbb{C}\setminus\Omega(T,a) = \{ \lambda \in \mathbb{C} : \text{exactly one of } \omega_1(\lambda,0), \omega_2(\lambda,0) \text{ is infinity} \}.
\]

The Brown measure of \( T + a \) is supported in the closure of \( \Omega(T,a) \) and is absolutely continuous in \( D(T,a) \), where

\[
D(T,a) = \Omega(T,a) \setminus S(T,a).
\]

**Definition 2.4.** For any \( \varepsilon > 0 \), we denote

\[
D^{(\varepsilon)}(T,a) = \{ \lambda \in \mathbb{C} : |\omega_1(\lambda,0)|, |\omega_2(\lambda,0)| \in (\varepsilon, 1/\varepsilon) \}.
\]

We then have

\[
D(T,a) = \bigcup_{n=1}^{\infty} D^{(1/n)}(T,a).
\]

**Proposition 2.5.** Let \( K \subset \Omega^c \) be a compact set. Consider the free additive convolution of \( \mu_1 \) and \( \mu_2 \). Then there exists a constant \( C > 0 \) such that

\[
G_{\mu_1 \boxplus \mu_2}(i\eta) \leq C
\]

for all \( \eta \geq 0 \) and \( \lambda \in K \).

**Proof.** By Proposition 2.1, \( (\lambda, \eta) \mapsto G_{\mu_1 \boxplus \mu_2}(i\eta) \) is continuous in \( \lambda \in K \) and \( \eta \geq 0 \). By Theorem 2.3, exactly one of \( \omega_1(\lambda,0), \omega_2(\lambda,0) \) is infinity when \( \lambda \in \mathbb{C}\setminus\Omega(T,a) \), thus \( G_{\mu_1 \boxplus \mu_2}(0) = 0 \). This shows that \( G_{\mu_1 \boxplus \mu_2}(i\eta) \) is bounded for \( (\lambda, \eta) \in K \times (0,1] \). The proposition then follows from the observation that \( G_{\mu_1 \boxplus \mu_2}(i\eta) \leq 1 \) if \( \eta \geq 1 \). \( \square \)

### 2.3. A result about the random matrix model.

For any \( \lambda \in \mathbb{C} \), we have to study the limit of the singular value distribution of the random matrix \( Y - \lambda I \). Since \( U \) and \( V \) are independent Haar-distributed unitary matrices, both \( Y - \lambda I = U\Sigma V^* + A - \lambda I \) and \( U\Sigma V^* + |A - \lambda I| \) have the same singular values. To this end, we consider the \( N \times N \) random matrix of the form

\[
X = X_N = U\Sigma V^* + \Xi
\]

where \( U = U_N \) and \( V = V_N \) are two independent Haar random unitary matrices, \( \Sigma = \Sigma_N \) and \( \Xi \) are sequences of deterministic matrices. The general result about \( X \) can be applied to \( Y - \lambda I \) by choosing \( \Xi = |A - \lambda I| \).

Without loss of generality, we may assume that \( \Xi \) is a diagonal matrix

\[
\Xi = \text{diag}(\xi_1, \cdots, \xi_N),
\]

where \( \xi_i \in \mathbb{C} \) for all \( i = 1, \cdots, N \). For \( \lambda \) in some compact set \( K \subset \mathbb{C} \), there exists some constant \( C_K \) independent of \( N \) such that

\[
||\Sigma||, ||\Xi|| \leq C_K.
\]

Denote the empirical distribution of the singular values of \( \Xi \) by

\[
\mu_\Xi = \frac{1}{N} \sum_{i=1}^{N} \delta_{|\xi_i|},
\]

and assume that there is a \( \xi \in \mathcal{A} \) such that

\[
\mu_\Xi \to \mu_\xi.
\]
Throughout this paper, we always assume that $\mu_\sigma$ and $\mu_\xi$ are not a single point mass at zero.

A general approach for a non-Hermitian random matrix is to consider its Hermitian reduction. We consider Hermitian random matrices $H$ defined by

$$H = \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} 0 & \Sigma \\ \Sigma^* & 0 \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & V^* \end{bmatrix} + \begin{bmatrix} 0 & \Xi \\ \Xi^* & 0 \end{bmatrix}.$$  \hfill (2.9)

The eigenvalues of the matrix $\begin{bmatrix} 0 & \Sigma \\ \Sigma^* & 0 \end{bmatrix}$ are exactly $\{|\sigma_i|, -|\sigma_i| : i = 1, \cdots, N\}$, where $|\sigma_i|$ are singular values of $\Sigma$. Similarly, the eigenvalues of the matrix $\begin{bmatrix} 0 & \Xi \\ \Xi^* & 0 \end{bmatrix}$ are exactly $\{\xi_i, -\xi_i : i = 1, \cdots, N\}$, where $\xi_i$ are singular values of $\Xi$.

By the asymptotic freeness result \[39, 45\], the random matrices $(U\Sigma V^* + \Xi)(U\Sigma V^* + \Xi)^*$ converges in $*$-moment to $(T + \xi)(T + \xi)^*$, where $T$ is an $R$-diagonal operator and $\xi$ is $*$-free from $T$. Hence, $\mu_{U\Sigma V^* + \Xi} \to \mu_{T + \xi}$ weakly. Let $u$ be a Haar unitary operator that is $*$-free from $\{T, \xi\}$, then $[T + u\xi]$ and $T + \xi$ have the same $*$-distribution. It is known that $u\xi$ is $R$-diagonal. Hence, $T + u\xi$ is a sum of two $R$-diagonal operators. By \[27\ Proposition 3.5\] (see also \[38\]), we have

$$\tilde{\mu}_{T + u\xi} = \tilde{\mu}_\sigma \boxplus \tilde{\mu}_\xi = \mu_{\sigma, \xi}.$$  

We conclude that the empirical eigenvalue distribution of $H$ converges weakly to $\mu_{\sigma, \xi}$. In \[3\], a qualitative version of this convergence was obtained. Given an interval $I \subset \mathbb{R}$ and $0 \leq a \leq b$, we denote

$$S_I(a, b) = \{z = x + i\eta : x \in I, a \leq \eta \leq b\}. \hfill (2.10)$$

Following \[3\], we use the following definition taken from \[18\].

**Definition 2.6** (Stochastic domination). Let $\mathcal{X} = \mathcal{X}^{(N)}$, $\mathcal{Y} = \mathcal{Y}^{(N)}$ be two sequence of nonnegative random variables. We say that $\mathcal{Y}$ stochastically dominates $\mathcal{X}$ if, for all (small) $\varepsilon > 0$ and (large) $D > 0$,

$$\mathbb{P}(\mathcal{X}^{(N)} > N^\varepsilon \mathcal{Y}^{(N)}) \leq N^{-D}, \hfill (2.11)$$

for sufficiently large $N \geq N_0(\varepsilon, D)$, and we write $\mathcal{X} \prec \mathcal{Y}$. When $\mathcal{X}^{(N)}$ and $\mathcal{Y}^{(N)}$ depend on a parameter $w \in W$. We say $\mathcal{X}(w) \prec \mathcal{Y}(w)$ uniformly in $w$ if $N_0(\varepsilon, D)$ can be chosen independent from $w$.

**Definition 2.7.** For two Borel probability measures $\nu_1, \nu_2$ on $\mathbb{R}$ and neither of $\nu_1, \nu_2$ is a point mass, denote by $f_{\nu_1 \boxplus \nu_2}$ the density function of $\nu_1 \boxplus \nu_2$. The bulk of $\nu_1 \boxplus \nu_2$ is defined as

$$\mathcal{B}_{\nu_1 \boxplus \nu_2} = \{x \in \mathbb{R} : 0 < f_{\nu_1 \boxplus \nu_2}(x) < \infty, \nu_1 \boxplus \nu_2\{\{x\}\} = 0\}. \hfill (2.12)$$

**Theorem 2.8.** Let $\mu_\sigma, \mu_\xi$ be two compactly supported probability measures on $[0, \infty)$ such that neither $\mu_\sigma$ nor $\mu_\xi$ is a single point mass and at least one of them is supported at more than two points. Fix some $L > 0$ and let $I$ be any compact subinterval of the bulk $\mathcal{B}_{\tilde{\mu}_\sigma \boxplus \tilde{\mu}_\xi}$. Then there exists a constant $b_0 > 0$ and $N_0 \in \mathbb{N}$, depending on $\mu_\sigma, \mu_\xi, I$ and the constant $C$ in \[2.8\], such that whenever

$$\sup_{N \geq N_0} \left(d_L(\mu_\Sigma, \mu_\sigma) + d_L(\mu_\Xi, \mu_\xi)\right) \leq 2b,$$

for some $b \leq b_0$, then

$$|G_H(z) - G_{\tilde{\mu}_\sigma \boxplus \tilde{\mu}_\xi}(z)| < \frac{1}{N\eta(1 + \eta)} \hfill (2.13)$$
holds uniformly on $S_I(0, N^L)$, for $N$ sufficiently large depending on $\mu_\sigma, \mu_\xi, I, L$ and the constant $C$ given in (2.8).

Moreover, there exists a constant $\eta_M \geq 1$, independent of $N$, such that (2.13) holds uniformly on $S_I(\eta_M, N^L)$, for any compact interval $I \subset \mathbb{R}$, for $N$ sufficiently large depending on $\mu_\sigma, \mu_\xi, I, L$ and the constant $C$ in (2.8).

2.4. Approximate subordination functions. Free additive convolution can be studied using the subordination functions. When we work with a sum of two random matrices that are asymptotically free, there is also a pair of approximate subordination functions [33]. To study the singular value distribution of the sum, we use the results in [8]. In the following, we introduce the approximate subordination functions in [8] and state a direct consequence of [8, Theorem 1.5].

Let $\Sigma$ and $\Xi$ be deterministic $N \times N$ matrices such that there is a constant $M$ independent of $N$ such that

$$\|\Sigma\|, \|\Xi\| \leq M. \quad (2.14)$$

Let $U$, $V$ be independent $N \times N$ Haar-distributed unitary matrices and $\tilde{\Sigma} = U \Sigma V^*$. Define

$$A = \begin{pmatrix} 0 & \Xi \\ \Xi^* & 0 \end{pmatrix}; \quad B = \begin{pmatrix} 0 & \tilde{\Sigma} \\ \tilde{\Sigma}^* & 0 \end{pmatrix}; \quad H = \begin{pmatrix} 0 & \Xi + \tilde{\Sigma}^* \\ (\Xi + \tilde{\Sigma})^* & 0 \end{pmatrix}. \quad (2.15)$$

Following [8, Eq.(67)], we define a complex-valued function $\omega_A$ by

$$\omega_A(z) = z + \frac{1}{\mathbb{E} G_H(z)} \left( \frac{1}{2N} \mathbb{E} \text{Tr} \left[ (z - H)^{-1} B \right] \right)$$

and similarly define $\omega_B$ by changing $B$ to $A$.

**Theorem 2.9.** For some functions $r_A(z)$ and $r_B(z)$ depending on $z \in \mathbb{C}^+$, we have

$$\mathbb{E} G_H(z) = \mathbb{E} G_A(\omega_A(z)) + r_A(z),$$

$$\mathbb{E} G_H(z) = \mathbb{E} G_B(\omega_B(z)) + r_B(z).$$

Write $\eta = \text{Im}(z)$. There exists a constant $C > 0$ that only depends on $M$ in (2.14) such that

1. if $N \eta^5 \geq C$, then $\text{Im} \omega_A(z), \text{Im} \omega_B(z) \geq \eta - \frac{C}{N \eta^7}$; and
2. if $N \eta^8 \geq C$, then $|r_A(z)|, |r_B(z)| \leq \frac{C}{N \eta^9}$.

**Proof.** Let $R_A$ and $R_B$ as in (24) and (25) in [8]. Let $r_A(z) = \frac{1}{2N} \text{Tr}[R_A(z)]$ and $r_B(z) = \frac{1}{2N} \text{Tr}[R_B(z)]$. Meanwhile, the function $S_B$ in [8, Eq.(67)] is related to $\omega_A$ by

$$\omega_A(z) = z + S_B(z)$$

and similarly $\omega_B(z) = z + S_A(z)$. The conclusion of the theorem follows directly from [8, Theorem 1.5]. \hfill \Box

We only need the values of the functions $\omega_A(z)$ and $\omega_B(z)$ for purely imaginary $z$.

**Lemma 2.10.** For any $\eta > 0$, $\omega_A(i\eta)$ and $\omega_B(i\eta)$ are purely imaginary.

**Proof.** Since $\omega_A$ is analytic, we only need to show that $\omega_A(i\eta)$ is purely imaginary for large enough $\eta$. That $\omega_A(i\eta)$ is also purely imaginary for all $\eta$ then follows from Schwarz reflection principle.
Write $H = \Xi + \tilde{\Sigma}$. Since $H$ has a symmetric distribution, $\mathbb{E}[G_H(i\eta)]$ is purely imaginary. By the formula (2.13) of $\omega_A$, it suffices to show that when $z = i\eta$ for all large enough $\eta$,

$$\frac{1}{2N} \mathbb{E} \text{Tr} \left[ (z - H)^{-1} B \right].$$

(2.16) is real.

We expand (2.16) into a power series

$$\sum_{n=0}^{\infty} \frac{1}{z^{n+1}} \mathbb{E} \frac{1}{2N} \text{Tr} \left[ \begin{pmatrix} 0 & H \\ H^* & 0 \end{pmatrix}^n \begin{pmatrix} 0 & \tilde{\Sigma} \\ \tilde{\Sigma}^* & 0 \end{pmatrix} \right].$$

(2.17)

It is straightforward to see (for example, by mathematical induction) that, if $n = 2m$,

$$\begin{pmatrix} 0 & H \\ H^* & 0 \end{pmatrix}^n = \begin{pmatrix} (HH^*)^m & 0 \\ 0 & (H^*H)^m \end{pmatrix}$$

and if $n = 2m + 1$,

$$\begin{pmatrix} 0 & H \\ H^* & 0 \end{pmatrix}^n = \begin{pmatrix} 0 & (HH^*H)^mH^* \\ (H^*H)^mH^* & 0 \end{pmatrix}.$$ 

Thus,

$$\frac{1}{2N} \text{Tr} \left[ \begin{pmatrix} 0 & H \\ H^* & 0 \end{pmatrix}^n \begin{pmatrix} 0 & \tilde{\Sigma} \\ \tilde{\Sigma}^* & 0 \end{pmatrix} \right] = \begin{cases} \text{Re} \text{tr}[(HH^*H)^mH\tilde{\Sigma}^*] & \text{if } n = 2m + 1 \\ 0 & \text{if } n = 2m \end{cases}.$$ 

This shows only the odd terms in (2.17) survive; in this case the $n$-th term is real if $z$ is purely imaginary. □

### 3. Least singular value estimate

To estimate the logarithmic potential of the random matrix

$$Y = USV^* + A,$$

we need to estimate the singular value of all the translates $Y - \lambda$, $\lambda \in \mathbb{C}$. The least singular values of $Y - \lambda$ have the same distribution as those of

$$U_1|A - \lambda|V_1 + U_2\Sigma V_2$$

where $U_1, V_1, U_2, V_2$ are independent Haar-distributed unitary matrices.

Since $|A - \lambda|$ is a nonnegative matrix, in this section, we estimate the least singular value of a random matrix of the form

$$U_1\Sigma_1 V_1 + U_2\Sigma_2 V_2$$

where $\Sigma_1$ and $\Sigma_2$ are deterministic nonnegative diagonal matrices and $U_1, V_1, U_2, V_2$ are independent Haar-distributed unitary matrices. Since we assume $A$ and $\Sigma$ satisfy Assumption [1][1] so we make the following assumption for $\Sigma_1$ and $\Sigma_2$.

**Assumption 3.1.** There exists $\alpha > 0$ such that if we denote

$$\gamma_1 = \# \{ i : (\Sigma_1)_{ii} \geq N^{-\alpha} \}, \quad \gamma_2 = \# \{ i : (\Sigma_2)_{ii} \geq N^{-\alpha} \},$$

then

$$\gamma_1 + \gamma_2 \geq N.$$ 

(3.1)

Given a matrix $L$, denote by $s_{\min}(L)$ be the least singular value of $L$. The following is the main theorem of this section. This theorem generalizes Theorem 1.1 in [40].
Theorem 3.2. Under the notation and assumptions in the preceding paragraphs, for some positive constants $c$ and $c'$ independent of $N$,
\[ P(s_{\min}(U_1 \Sigma_1 V_1 + U_2 \Sigma_2 V_2) \leq t) \leq t^c N^{c'}, \quad t > 0. \]
As a consequence, Theorem 1.6 holds.

Remark 3.3. Since $U_1, V_1, U_2, V_2$ are independent Haar-distributed unitary matrices, the distribution of $s_{\min}(U_1 \Sigma_1 V_1 + U_2 \Sigma_2 V_2)$ remains unchanged if we permute the diagonal entries of $\Sigma_1$ or $\Sigma_2$. By (3.1) in our assumptions, we may assume for every $j$, at least one of $(\Sigma_1)_{jj}$ or $(\Sigma_2)_{jj}$ is at least $N^{-\alpha}$.

We adapt the arguments in [40, Theorem 1.1] to prove Theorem 3.2. By the min-max theorem,
\[ s_{\min}(U_1 \Sigma_1 V_1 + U_2 \Sigma_2 V_2) = \inf_{\|x\|_2 = 1} \|(U_1 \Sigma_1 V_1 + U_2 \Sigma_2 V_2)x\|_2. \]
For each $x$ such that $\|x\|_2 = 1$, there exists at least one entry $x_j$ such that $|x_j| \geq 1/\sqrt{N}$. Let $S(1) = \{ x \in \mathbb{C}^N : |x_1| \geq 1/\sqrt{N} \}$. Then by symmetry,
\[ P(s_{\min}(U_1 \Sigma_1 V_1 + U_2 \Sigma_2 V_2) \leq t) \leq N P \left( \inf_{x \in S(1)} \|(U_1 \Sigma_1 V_1 + U_2 \Sigma_2 V_2)x\|_2 \leq t \right). \]
Thus, it suffices to bound $P \left( \inf_{x \in S(1)} \|(U_1 \Sigma_1 V_1 + U_2 \Sigma_2 V_2)x\|_2 \leq t \right)$. To this end, observe that
\[ s_{\min}(U_1 \Sigma_1 V_1 + U_2 \Sigma_2 V_2) = s_{\min}(U_1 \Sigma_1 V_1^* V_2 + U_2 \Sigma_2). \]
Without loss of generality, we assume $(\Sigma_2)_{11} \geq N^{-\alpha}$ because at least one of $(\Sigma_1)_{jj}$ or $(\Sigma_2)_{jj}$ is at least $N^{-\alpha}$ by Remark 3.3. For notational convenience, write $D = U_1 \Sigma_1 V_1^* V_2$, $\Sigma = \Sigma_2$, and $U = U_2$. For the rest of this section, we estimate
\[ s_{\min}(D + U \Sigma), \]
where $U$ is a Haar-distributed unitary random matrix.

Define the skew-Hermitian random matrix
\[ S = \begin{pmatrix} is & -Z^T \\ Z & 0 \end{pmatrix}, \]
where $s$ is a real-valued standard normal random variable and $Z$ is an $(n - 1)$-dimensional standard normal random vector. For any $\varepsilon > 0$, write the singular value decomposition $I + \varepsilon S = U_0 \Sigma V_0$. Define a random matrix $W \in U(N)$ by
\[ W = U_0 V_0. \]
We now recall the following lemma from [40, Lemma 3.1].

Lemma 3.4. With $S$ and $W$ defined as above,
\[ \|W - (I + \varepsilon S)\| \leq 2\varepsilon^2 \|S^2\| \]
whenever $\varepsilon^2 \|S^2\| \leq 1/4$.

Suppose that
1. $V$ is a Haar unitary random matrix on $U(N)$;
2. $W$ is as in (3.3);
3. $R = \text{diag}(r, 1, \ldots, 1)$ where $r$ is uniformly distributed on the unit circle.
4. $V$, $W$ and $R$ are independent.
Then $V^{-1} R^{-1} W$ is Haar-distributed on $U(N)$ and we set $U = V^{-1} R^{-1} W$. Since $\|\Sigma\| \leq M$, as a consequence of Lemma 3.4, we have
\begin{align*}
\|(D + U\Sigma)x\|_2 &= \|(RV D + (I + \varepsilon S)\Sigma)x + (W - (I + \varepsilon S))\Sigma x\|_2 \\
&\geq \|(RV D + (I + \varepsilon S)\Sigma)x\|_2 - 2M\varepsilon^2 \|S\|
\end{align*}
whenever $\varepsilon^2 \|S\| \leq 1/4$.

Lemma 3.5. Write $\Lambda = RV D + (I + \varepsilon S)\Sigma$. Choose $\varepsilon \in (0, 1)$, $K_0 > 1$ and $\mu \in (0, 1)$ such that
\[ \varepsilon^2 K_0^2 N \leq 1/4, \quad \mu \geq 2M\varepsilon K_0^2 N. \]

Then for some constant $c$,
\[ \mathbb{P}\left( \inf_{x \in \mathbb{S}(1)} \|(D + U\Sigma)x\|_2 \leq \mu\varepsilon \right) \leq \mathbb{P}\left( \left\{ \inf_{x \in \mathbb{S}(1)} \|Ax\|_2 \leq 2\mu\varepsilon \right\} \cap \mathcal{E}_S \right) + 2e^{-cK_0^2 N} \]
where $\mathcal{E}_S$ denotes the event $\{\|S\| \leq K_0\sqrt{N}\}$.

Proof. This follows from a similar computation as in Section 3.1.3 of [10]. Note that the term $2e^{-cK_0^2 N}$ comes from $\mathbb{P}(\mathcal{E}_S^c) \leq 2e^{-cK_0^2 N}$. $\square$

We use the subscript $(1, 1)$ of a matrix to mean the submatrix obtained by deleting the first row and the first column of the matrix. Recall that $R = \text{diag}(r, 1, \ldots, 1)$ and we use the notation
\[ \Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{(1,1)} \end{pmatrix}; \quad VD = \begin{pmatrix} (VD)_{11} & (VD)_{1,1}^T \\ u & (VD)_{(1,1)} \end{pmatrix}; \quad S = \begin{pmatrix} is & -Z^T \\ Z & 0 \end{pmatrix}. \]
We assume $\Sigma_{(1,1)} \neq 0$, because we assume $\mu_{\Sigma_N}$ converges to a probability measure different from $\delta_0$. With these notation, we have
\[ \Lambda = RV D + (I + \varepsilon S)\Sigma = \begin{pmatrix} r(VD)_{11} + \Sigma_{11}(1 + i\varepsilon s) & (rv - \varepsilon \Sigma_{(1,1)}Z)^T \\ u + \varepsilon \Sigma_{11}Z & (VD + \Sigma)_{(1,1)} \end{pmatrix}, \quad (3.4) \]
In the following lemma, we prove that $(VD + \Sigma)_{(1,1)}$ is invertible almost surely.

Lemma 3.6. The submatrix $(VD + \Sigma)_{(1,1)}$ is invertible almost surely.

Proof. While in this section we have mostly dealt with the randomness of $U$ but not $D$, in this proof, it is helpful to recall $D = U_1 \Sigma_1 V_1 V_2^*$ is a random matrix. Since we write $\Sigma_1$ in this proof, we write $\Sigma$ in place of $\Sigma_2$.

We view the statement of the lemma as to show that the function defined by
\[ \Phi(V, U_1, V_1, V_2) \mapsto (VU_1 \Sigma_1 V_1 V_2^* + \Sigma_2)_{(1,1)} \]
is invertible almost everywhere with respect to the product of the Haar measures on the unitary group $U(N)$. This product measure is the Haar measure on $U(N)^4$.

Recall Remark 3.3 that we can without loss of generality assume for every $j$, at least one of $(\Sigma_1)_{jj}$ or $(\Sigma_2)_{jj}$ is at least $N^{-\alpha}$; in particular, at least one of $(\Sigma_1)_{jj}$ or $(\Sigma_2)_{jj}$ is positive. Then $\Phi(I, I, I, I)$ is invertible. This shows that the set $\{(V, U_1, V_1, V_2) : \Phi(V, U_1, V_1, V_2) \text{ is not invertible}\}$ is the solution of a non-zero polynomial, namely the determinant of $\Phi(V, U_1, V_1, V_2)$, on $U(N)^4$.

Since the zeros of a polynomial on $\mathbb{C}^m$ form a Lebesgue-measure-zero set for any $m > 0$, by that the Haar measure on a local chart is absolutely continuous with respect to the
Lebesgue measure on the chart, we can see that \( \Phi(V, U_1, V_1, V_2) \) is invertible for Haar-almost every \((V, U_1, V_1, V_2)\) in a local chart around \((I, I, I, I)\); in particular, there exists an open set containing \((I, I, I, I)\) such that \( \Phi(V, U_1, V_1, V_2) \) is invertible for Haar-almost every \((V, U_1, V_1, V_2)\) in the open set. We claim that if \( O \) is an open set containing \((I, I, I, I)\) such that \( \Phi(V, U_1, V_1, V_2) \) is invertible for Haar-almost every \((V, U_1, V_1, V_2)\) in \( O \), then either \( O = U(N)^4 \) or there exists another open set \( O_1 \) such that \( O \) is a proper subset of \( O_1 \) and \( \Phi(V, U_1, V_1, V_2) \) is invertible for Haar-almost every \((V, U_1, V_1, V_2)\) in \( O_1 \). Suppose \( O \) is not \( U(N)^4 \). Since \( U(N)^4 \) is connected, there exists a point \( W \in U(N)^4 \) such that \( W \in \partial O \). By a similar argument as above, there exists a chart \( O' \) containing \( W \) such that \( \Phi \) is invertible Haar-almost everywhere in \( O' \). Thus \( O \) is a proper subset of the open set \( O_1 = O \cup O' \) in which \( \Phi \) is invertible Haar-almost everywhere.

We are now already to conclude the lemma. Consider \( \mathcal{F} \) to be the family of open sets containing \( W^* \) such that \( \Phi \) is invertible Haar-almost everywhere. The lemma will be established if we show \( U(N)^4 \in \mathcal{F} \). Let

\[
\tilde{O} = \bigcup_{O \in \mathcal{F}} O.
\]

Then \( \tilde{O} \) is a maximal element in \( \mathcal{F} \). Since \( \tilde{O} \) is maximal, by the claim proved in the previous paragraph, we must have \( \tilde{O} = U(N) \), completing the proof of the lemma. \( \square \)

By Lemma 3.5, it suffices to estimate

\[
\mathbb{P}\left( \inf_{x \in S(1)} \|Ax\|_2 \leq 2\mu \right).
\]

To estimate \( \inf_{x \in S(1)} \|Ax\|_2 \), write \( \Lambda_1, \ldots, \Lambda_n \) be the columns of \( \Lambda \). Let \( h \in \mathbb{C}^N \) be a unit vector such that

\[
h^T \Lambda_j = 0
\]

for all \( j = 2, \ldots, N \). Then by Cauchy–Schwarz inequality, for all \( x \in S(1) \),

\[
\|Ax\|_2 \geq \left| h^T \sum_{j=1}^N x_j \Lambda_j \right| \geq \frac{1}{\sqrt{N}} |h^T \Lambda_1|.
\]

The following lemma follows from a direct application of Lemma 3.2 of [40]. We will then estimate \( |h^T \Lambda_1| \) case-by-case.

**Lemma 3.7.** Suppose that \((VD + \Sigma)_{(1,1)}\) is invertible and write \( L = ((VD + \Sigma)_{(1,1)})^{-T} \). We can compute \( |h^T \Lambda_1| \) as

\[
|\sqrt{1 + \|rLv - \varepsilon L\Sigma_{(1,1)}\|^2_2} |.
\]

3.1. **First case:** under \( \varepsilon_{\text{denom}} = \{\|rLv - \varepsilon L\Sigma_{(1,1)}\|^2_2 \leq K\} \). Let \( K \geq 1 \) be to be chosen later. The event

\[
\varepsilon_{\text{denom}} = \{\|rLv - \varepsilon L\Sigma_{(1,1)}\|^2_2 \leq K\}
\]

is independent of \( s \) (recall that \( s \) is real standard normal). We then have the following estimate of the probability of the event \( \varepsilon_{\text{denom}} \); in Section 3.2.3 we will analyze \( \varepsilon_{\text{denom}}^c \) by dividing into several cases.
Lemma 3.9. The event $E$ is likely; more precisely, we have
\[ \mathbb{P} \left\{ \inf_{x \in \mathbb{S}(1)} \| \Lambda x \|_2 \leq \frac{\lambda \varepsilon}{K \sqrt{N}} \text{ and } \mathcal{E}_{\text{denom}} \right\} \leq C_1 N^\alpha \lambda, \quad \lambda > 0 \]
for some constant $C_1 > 0$.

Proof. By conditioning on $r$ and $Z$ which satisfy $\mathcal{E}_{\text{denom}}$, $|h^T \Lambda_1|$ in Lemma 3.7, the following inequality
\[ |h^T \Lambda_1| = \frac{|a + ib + i\varepsilon \sum_{11} |}{1 + K^2} \geq \frac{|a + i(b + \varepsilon \sum_{11})|}{2K} \geq \frac{|b + \varepsilon \sum_{11}|}{2K}, \]
for some real numbers $a$ and $b$ independent of $s$. Recall that we assume $\sum_{11} \geq N^{-\alpha}$. Therefore, by denoting $P_s$ the probability of conditioning on all random variables but $s$, we have, for all $\lambda > 0$,
\[ \mathbb{P}_s \left\{ |h^T \Lambda_1| \leq \frac{\lambda \varepsilon}{K} \right\} \leq \mathbb{P}_s \left\{ \frac{|b + \varepsilon \sum_{11}|}{2K} \leq \frac{\lambda \varepsilon}{K} \right\} = \mathbb{P}_s \left\{ \frac{|b + \varepsilon |}{\varepsilon + \sum_{11}} \leq 2\lambda \right\} \leq C_1 N^\alpha \lambda \]
where the last inequality comes from the density of $\sum_{11}$ is at most $1/\sqrt{2\pi \sum_{11}^2} \leq N^\alpha/\sqrt{2\pi}$. By (3.5), we have the estimate for the unconditional probability
\[ \mathbb{P} \left\{ \inf_{x \in \mathbb{S}(1)} \| \Lambda x \|_2 \leq \frac{\lambda \varepsilon}{K \sqrt{N}} \text{ and } \mathcal{E}_{\text{denom}} \right\} \leq \mathbb{P} \left\{ |h^T \Lambda_1| \leq \frac{\lambda \varepsilon}{K} \text{ and } \mathcal{E}_{\text{denom}} \right\} \leq C_1 N^\alpha \lambda, \]
as desired. \hfill \square

3.2. Dividing $\mathcal{E}_{\text{denom}}$ into cases.

3.2.1. Two likely events $\mathcal{E}_{L \Sigma(1,1)Z}$ and $\mathcal{E}_{\text{num}}$. Before we analyze $\mathcal{E}_{\text{denom}}$, we introduce two likely events. If $\mathcal{E}_{\text{denom}}$ does not occur, then we must have
\[ \sqrt{1 + K^2} < \sqrt{1 + \|rL_\nu - \varepsilon L_{\Sigma(1,1)}Z\|_2^2} \leq 2\|L_\nu\|_2 + \varepsilon \|L_{\Sigma(1,1)}Z\|_2. \]
Note that $\mathbb{E} \|L_{\Sigma(1,1)}Z\|_2^2 = \|L_{\Sigma(1,1)}\|_{\text{HS}}^2$; thus, we consider the event
\[ \mathcal{E}_{L \Sigma(1,1)Z} = \{ \|L_{\Sigma(1,1)}Z\|_2 \leq K_1 \|L_{\Sigma(1,1)}\|_{\text{HS}} \}. \]
where $K_1 \geq 1$ will be chosen. The other independent event that we consider is
\[ \mathcal{E}_{\text{num}} = \{ Q \geq \lambda_1 |(VD)_{11}^T - u^T L_\nu - \varepsilon \Sigma_{11} (L_\nu)^T Z| \}\]
where $Q$ is the numerator of the expression of $|h^T \Lambda_1|$ in Lemma 3.7.

Lemma 3.9. The event $\mathcal{E}_{L \Sigma(1,1)Z}$ is likely; more precisely, we have
\[ \mathbb{P}(\mathcal{E}_{L \Sigma(1,1)Z}) \geq 1 - \exp(-cK_1^2). \]

Proof. The function $f(Z) = \|L_{\Sigma(1,1)}Z\|_2$ defined on $\mathbb{R}^{N-1}$ has Lipschitz norm bounded by $\|L_{\Sigma(1,1)}\|$. We then have the concentration inequality
\[ \mathbb{P}(\mathcal{E}_{L \Sigma(1,1)Z}) \geq 1 - \exp \left( - \frac{cK_1^2 \|L_{\Sigma(1,1)}\|_{\text{HS}}^2}{\|L_{\Sigma(1,1)}\|^2_2} \right) \geq 1 - \exp(-cK_1^2). \]
The lemma is established. \hfill \square
Lemma 3.10. There exists a constant $C_1 > 0$ such that
\[ \mathbb{P}(\mathcal{E}_{\text{num}}) \geq 1 - C_1\lambda_1, \quad \lambda_1 \in (0, 1). \]

By choosing a larger constant if necessary, we assume this $C_1$ is the same as the one in Proposition 3.8.

Proof. Recall that $r$ is uniformly distributed in the unit circle. Given any $a, b \in \mathbb{C}$ and $\lambda_1 \in (0, 1)$, $\mathbb{P}(|ar + b| < \lambda_1|a|)$ is maximized precisely when either $a = 0$ or $|b/a| = 1$, where the latter means that the circle $\{ae^{i\theta} + b : \theta \in (-\pi, \pi)\}$ passes through the origin. Thus we have
\[ \mathbb{P}(|ar + b| \geq \lambda_1|a|) \geq 1 - C_1\lambda_1, \quad \lambda_1 \in (0, 1) \]
for some constant $C_1 > 0$ independent of $a$ and $b$. The lemma is proved, because, by conditioning on all random variables except $r$, the event $\mathcal{E}_{\text{num}}$ is of the form $\{|ar + b| \geq \lambda_1|a|\}$. \qed

3.2.2. The event $\mathcal{E}_L = \{\|L\|_{\text{HS}} \leq K/2\varepsilon MK_1\}$. Recall that $\Sigma$ is a diagonal matrix and $\|\Sigma\| \leq M$; as a consequence, $\|\Sigma(1,1)\| \leq M$. On the event $\mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{LS(1,1)Z}$, we have
\[
\sqrt{1 + \|rLv - \varepsilon L\Sigma(1,1)Z\|^2_2} \leq 2\|Lv\|_2 + \varepsilon K_1\|\Sigma(1,1)\|\|L\|_{\text{HS}} \\
\leq 2\|Lv\|_2 + \varepsilon K_1 M\|L\|_{\text{HS}}.
\]

We look at the event
\[ \mathcal{E}_L = \left\{ \|L\|_{\text{HS}} \leq \frac{K}{2\varepsilon MK_1} \right\}. \]

Then on the event $\mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{LS(1,1)Z} \cap \mathcal{E}_L$,
\[ K \leq \sqrt{1 + \|rLv - \varepsilon L\Sigma(1,1)Z\|^2_2} \leq 2\|Lv\|_2 + \frac{K}{2}, \]
which shows
\[ \sqrt{1 + \|rLv - \varepsilon L\Sigma(1,1)Z\|^2_2} \leq 4\|Lv\|_2. \quad (3.6) \]

This inequality will be useful to give a lower bound for $|h^T\Lambda_1|$ in the event $\mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{LS(1,1)Z} \cap \mathcal{E}_L$.

3.2.3. Estimating $\mathcal{E}_{\text{denom}}^c$ case-by-case. On the event $\mathcal{E}_{\text{denom}}^c$, we divide into cases:

1. $\mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{LS(1,1)Z}$
2. $\mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{LS(1,1)Z} \cap \mathcal{E}_L \cap \mathcal{E}_{\text{num}}$
3. $\mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{LS(1,1)Z} \cap \mathcal{E}_L \cap \mathcal{E}_{\text{num}}$
4. $\mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{LS(1,1)Z} \cap \mathcal{E}_{\text{num}}$

The first two events have small probabilities by Lemmas 3.9 and 3.10. We will analyze the last two events in this subsection.

We first look at the event $\mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{LS(1,1)Z} \cap \mathcal{E}_L \cap \mathcal{E}_{\text{num}}$. Recall that the event $\mathcal{E}_{\text{num}}$ depends on the parameter $\lambda_1$.

Proposition 3.11. Given any $\lambda_1 \in (0, 1),
\[ \mathbb{P}(h^T\Lambda_1 \leq \lambda\lambda_1\varepsilon \text{ and } \mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{LS(1,1)Z} \cap \mathcal{E}_L \cap \mathcal{E}_{\text{num}}) \leq C_1 N^\alpha \lambda, \quad \lambda > 0. \]
Therefore, we also have
\[ \mathbb{P}\left( \inf_{x \in \mathbb{S}(1)} \|Ax\|_2 \leq \frac{\lambda_1 \varepsilon}{\sqrt{N}} \text{ and } \mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{\Sigma_{(1,1)}Z \cap \mathcal{E}_{L} \cap \mathcal{E}_{\text{num}}} \right) \leq C_1 N^\alpha \lambda, \quad \lambda > 0. \]

We postpone the proof of Proposition 3.11. By combining the probability estimates under the first three events listed in the beginning of this subsection: \( \mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{\Sigma_{(1,1)}Z \cap \mathcal{E}_{L} \cap \mathcal{E}_{\text{num}}} \), \( \mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{\Sigma_{(1,1)}Z \cap \mathcal{E}_{L} \cap \mathcal{E}_{\text{num}}} \), and \( \mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{\Sigma_{(1,1)}Z \cap \mathcal{E}_{L} \cap \mathcal{E}_{\text{num}}} \), we get the following estimate under the event \( \mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{L} \).

**Corollary 3.12.** For all \( \lambda, \lambda_1 \in (0, 1) \),
\[ \mathbb{P}\left( \inf_{x \in \mathbb{S}(1)} \|Ax\|_2 \leq \frac{\lambda_1 \varepsilon}{\sqrt{N}} \text{ and } \mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{L} \right) \leq C_1 (\lambda_1 + N^\alpha \lambda) + \exp(-cK^2). \]

**Proof.** Applying Lemma 3.10 and Proposition 3.11 gives us the estimate
\[ \mathbb{P}\left( \inf_{x \in \mathbb{S}(1)} \|Ax\|_2 \leq \frac{\lambda_1 \varepsilon}{\sqrt{N}} \text{ and } \mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{\Sigma_{(1,1)}Z \cap \mathcal{E}_{L}} \right) \leq C_1 (\lambda_1 + N^\alpha \lambda). \]

We further apply Lemma 3.9 to conclude the corollary.

**Proof of Proposition 3.11.** By writing \( d = \lambda_1 \frac{(VD)_{11} - u^T L v}{4\|L v\|_2} \) and \( w = -\frac{L v}{\|L v\|_2} \), on \( \mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{\Sigma_{(1,1)}Z \cap \mathcal{E}_{L} \cap \mathcal{E}_{\text{num}}} \), we have
\[ |h^T \Lambda_1| \geq \frac{\lambda_1 |(VD)_{11} - u^T L v - \varepsilon \Sigma_{11} (Lv)^T Z|}{4\|L v\|_2} = \left| d + \frac{\lambda_1 \varepsilon \Sigma_{11}}{4} w^T Z \right| \]
where the estimate of the numerator comes from the definition of \( \mathcal{E}_{\text{num}} \) and that of the denominator comes from (3.6).

We now condition on all random variables except \( Z \). Since \( w \) is a unit vector independent of \( Z \), \( w^T Z = oX \) in conditional distribution for some \( \alpha \in \mathbb{T} \) and standard real normal random variable \( X \). Since \( N^{-\alpha} \leq \Sigma_{11} \leq \|\Sigma\| \leq M \) and the density of \( X \) is bounded above by \( 1/\sqrt{2\pi} \), we have
\[ \mathbb{P}_Z \left( \left| d + \frac{\lambda_1 \varepsilon \Sigma_{11}}{4} w^T Z \right| \leq \lambda_1 \varepsilon \right) \leq C_1 N^\alpha \lambda, \quad \lambda > 0. \]

By choosing a larger constant, we assume this \( C_1 \) is the same as the one in Lemma 3.10.

Finally we compute an estimate for the event \( \mathcal{E}_{\text{denom}}^c \cap \mathcal{E}_{\Sigma_{(1,1)}Z \cap \mathcal{E}_{L}} \). In fact, we compute an estimate for the (larger) event \( \mathcal{E}_{L}^c \).

**Proposition 3.13.**
\[ \mathbb{P}\left( \inf_{x \in \mathbb{S}(1)} \|Ax\|_2 \leq \varepsilon \lambda \sqrt{N} - \frac{2\varepsilon MK_1 \sqrt{N}}{K} \text{ and } \mathcal{E}_{L}^c \right) \leq C_1 N^\alpha \lambda \]

By choosing a larger constant, we assume this \( C_1 \) is the same as the one in Corollary 3.12.

**Proof.** For notational convenience, write \( B = (VD + \Sigma)_{(1,1)} \), so that \( B^{-T} = L \). On \( \mathcal{E}_{L}^c \), we have
\[ \|B^{-1}\| = \|L\| \geq \frac{1}{\sqrt{N}} \|L\|_{HS} \geq \frac{K}{2\varepsilon MK_1 \sqrt{N}}. \]
Choose $w_1 \in \mathbb{C}^{n-1}$ with $\|w_1\|_2 = 1$ such that $\left\| \frac{B^{-1}}{\|B^{-1}\|} w_1 \right\| = 1$. Then $\tilde{w} = \frac{B^{-1}}{\|B^{-1}\|} w_1$ is a unit vector and

$$\|B\tilde{w}\| = \frac{1}{\|B^{-1}\|} \leq \frac{2\varepsilon MK_1 \sqrt{N}}{K}.$$ 

Note that $\tilde{w}$ is chosen using only $B$ and thus independent of the random vector $Z$.

Let $x \in S(1)$ be of the form

$$x = \begin{pmatrix} x_1 \\ \bar{x} \end{pmatrix}$$

where, by the definition of $S(1)$, $|x_1| \geq \frac{1}{\sqrt{N}}$. Let $w \in \mathbb{C}^{n-1}$ be

$$w = \begin{pmatrix} 0 \\ \tilde{w} \end{pmatrix}.$$ 

It follows from (3.4) that

$$\|\Lambda x\|_2 \geq |w^T \Lambda x| \geq \|w^T \Lambda \bar{x}\| = \|w^T \bar{x}\| - \|B\tilde{w}\|_2 \geq \frac{1}{\sqrt{N}} |\tilde{w}^T u + \varepsilon \Sigma_{11} \tilde{w}^T Z| - \frac{2\varepsilon MK_1 \sqrt{N}}{K}.$$ 

Now, $\tilde{w}^T u$ is independent of $Z$, and $\tilde{w}^T Z = \alpha X$ in distribution where $\alpha \in \mathbb{C}$, $|\alpha| = 1$, and $X$ is a standard real normal random variable. Since $N^{-\alpha} \leq \Sigma_{11} \leq \|\Sigma\| \leq M$, when we condition on all random variables except $Z$,

$$\mathbb{P}(\|\tilde{w}^T u + \varepsilon \Sigma_{11} \tilde{w}^T Z\| \leq \varepsilon \lambda) \leq C_1 N^\alpha \lambda, \quad \lambda > 0.$$ 

Therefore, we have the following estimate for the unconditional probability

$$\mathbb{P}\left( \inf_{x \in S(1)} \|\Lambda x\|_2 \leq \frac{\varepsilon \lambda}{\sqrt{N}} - \frac{2\varepsilon MK_1 \sqrt{N}}{K} \text{ and } \varepsilon_L \right) \leq C_1 N^\alpha \lambda, \quad \lambda > 0.$$ 

This completes the proof of the proposition. \hfill \Box

3.3. **Proof of Theorem 3.2.**

Proof of Theorem 3.2. We need to choose $\mu$ such that $\mu$ satisfies the condition in Lemma 3.5.

Set

$$\mu = \frac{1}{2} \min \left( \frac{\lambda}{K \sqrt{N}}, \frac{\lambda \lambda_1}{\sqrt{N}}, \frac{\lambda}{K \sqrt{N}} - \frac{2MK_1 \sqrt{N}}{K} \right).$$ 

We have, by Proposition 3.8, Corollary 3.12 and Proposition 3.13

$$\mathbb{P}\left( \inf_{x \in S(1)} \|\Lambda x\|_2 \leq 2\mu \varepsilon \right) \leq C_1 N^\alpha \lambda + C_1 (\lambda_1 + N^\alpha \lambda) + \exp(-cK_1^2) + C_1 N^\alpha \lambda.$$ 

By Lemma 3.5

$$\mathbb{P}\left( \inf_{x \in S(1)} \|(D + U \Sigma)x\|_2 \leq \mu \varepsilon \right) \leq 3C_1 N^\alpha \lambda + C_1 \lambda_1 + \exp(-cK_1^2) + 2\exp(-cK_0^2 N).$$ 

Therefore,

$$\mathbb{P}(s_{\min}(D + U \Sigma) \leq \mu \varepsilon) \leq N[3C_1 N^\alpha \lambda + C_1 \lambda_1 + \exp(-cK_1^2) + 2\exp(-cK_0^2 N)].$$
We proceed to choose the parameters. Let $\varepsilon > 0$ be a constant to be chosen later. Choose $\lambda = \lambda_1 = \varepsilon^{0.1} K_0 = \frac{\log(1/\varepsilon)}{\sqrt{M}}, K_1 = \frac{\log(1/\varepsilon)}{M}, K = \frac{4\log(1/\varepsilon) N}{\varepsilon^{0.1}}$.

Then

$$\mu = \frac{1}{2} \frac{\varepsilon^{0.1}}{\sqrt{N}} \min \left( \frac{\varepsilon^{0.1}}{4 N \log(1/\varepsilon)}, \frac{1}{2} \right).$$

If $\varepsilon = \delta/N^4$ for a sufficiently small positive constant $\delta$, then

$$\mu = \frac{1}{2} \frac{\delta^{0.2}}{4 N^{3/2} \log(N^4/\delta)} \geq \frac{2\delta}{N^3} \log^2 \left( \frac{N^4}{\delta} \right) = 2M \varepsilon K_0^2 N$$

and

$$\varepsilon^2 K_0^2 N \leq \frac{1}{4},$$

so that the conditions in Lemma 3.5 are satisfied.

With the choices of parameters in the preceding paragraph, we have

$$\mathbb{P} \left( s_{\min}(D + U\Sigma) \leq \frac{1}{2} \frac{\delta^{1.2}}{4 N^{11/2} \log(N^4/\delta)} \right)$$

$$\leq N \left[ 3C_1 \frac{N^a \delta^{0.1}}{N^{0.4}} + C_1 \frac{\delta^{0.1}}{N^{0.4}} + \exp \left( -\frac{c}{M^2} \log^2 (N^4/\delta) \right) + 2 \exp \left( -\frac{c}{M} \log^2 (N^4/\delta) N \right) \right].$$

Since this is true for all sufficiently small $\delta$, the above estimate establishes Theorem 3.2. $\square$

4. The deformed local single ring theorem

4.1. Convergence of Cauchy transforms in the bulk. Consider the random matrix model $Y$ as in (1.2) and operator $y$ in (1.3). To study the Cauchy transform of $Y$, we apply the random matrix model $X$ in (2.7) with $\Xi = |A - \lambda|$. In this section, we try to understand the rate of the convergence of the Cauchy transform of $Y$. For any $\lambda$, we write (following the Girko’s trick)

$$H^\lambda = \begin{pmatrix} 0 & Y - \lambda \\ Y^* - \overline{\lambda} & 0 \end{pmatrix}. \quad (4.1)$$

The eigenvalues of $H^\lambda$ are exactly $\{z^\lambda_i, -z^\lambda_i, i = 1, \ldots, N\}$, where $z^\lambda_i$ are singular values of $Y - \lambda$. Denote by $G^\lambda$ the Cauchy transform of $H^\lambda$. Then,

$$G^\lambda(z) = \frac{1}{2N} \left( \frac{1}{z - z^\lambda_i} + \frac{1}{z + z^\lambda_i} \right).$$

We denote

$$G_{\Sigma, |A - \lambda|} = G_{\mu_{\Sigma, |A - \lambda|}} \quad (4.2)$$

where $\mu_{\Sigma, |A - \lambda|}$ is the free convolution of $\mu_\Sigma$ and $\mu_{|A - \lambda|}$, consistent to the notation (2.2).

**Theorem 4.1.** For any $L_0 > 0$, there exists $N_0$ depending on $\varepsilon, T, a$ and $M$, the estimate

$$\sup_{\lambda \in D^{(\varepsilon)}(T,a)} |G^\lambda(i\eta) - G_{\Sigma, |A - \lambda|}(i\eta)| \leq \frac{1}{N\eta}, \quad (4.3)$$

holds uniformly in $\eta > N^{-L_0}$ for all $N \geq N_0$. 
Proof. For any \( \lambda \in \mathbb{C} \), let
\[
A - \lambda = U_1^* \Xi^\lambda V_1
\]
be the singular value decomposition of \( A - \lambda \), where \( U_1, V_1 \) are unitary matrices and \( \Xi^\lambda = \text{diag}(\xi_1^\lambda, \cdots, \xi_N^\lambda) \) is a diagonal matrix consisting of the singular values \( \xi_j^\lambda \) of \( A - \lambda \) on the diagonal. Then, the singular values of \( U \Sigma \Sigma^* + A - \lambda \) has the same singular values as \( U_1 U \Sigma \Sigma^* V_1 + \Xi^\lambda \). Then \( U_1 U \) and \( V_1 V \) are again two independent Haar random unitary matrices. Hence, \( U \Sigma \Sigma^* + A - \lambda \) has the same singular values as \( U \Sigma \Sigma^* + \Xi^\lambda \). Moreover, since the matrices \( A_N - \lambda \) converge in \( * \)-moments to \( a - \lambda \), it follows that
\[
\mu_{|A_N - \lambda|} \rightarrow \mu_{|a - \lambda|}
\]
weakly. Recall that our random matrix model assumes \( \mu_{\Sigma} \rightarrow \mu_\sigma \) weakly.

If \( |z| \gg 1 \), we note that for a symmetric probability measure \( \mu \) on \( \mathbb{R} \), we can expand
\[
G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z - u} \, d\mu(u) = \sum_{k=0}^{\infty} \frac{m_{2k}(\mu)}{z^{2k+1}},
\]
where \( m_n(\mu) = \int_{\mathbb{R}} u^n \, d\mu(u) \) is the \( n \)-th moment of \( \mu \). Consequently, for any \( \lambda \in D^{(e)}(T, a) \), we have
\[
G^\lambda(i\eta) = \frac{1}{i\eta} + O\left( \frac{1}{\eta^3} \right), \quad G_{\Sigma,|A-\lambda|}(i\eta) = \frac{1}{i\eta} + O\left( \frac{1}{\eta^3} \right)
\]
provided that \( \eta > N^L \) and \( L \geq L_0 \) is sufficiently large. The above approximation is uniform for all \( \lambda \in D^{(e)}(T, a) \). Hence, (4.3) holds uniformly in \( \lambda \) for \( \eta > N^L \). In the rest of the proof, we show that (4.3) also holds uniformly in \( \lambda \) for \( N^{-L_0} \leq \eta \leq N^L \).

Recall that we write \( \mu_1 = \tilde{\mu}_{|a-\lambda|} \) and \( \mu_2 = \tilde{\mu}_{|T|} = \tilde{\mu}_\sigma \). There are subordination functions in the sense of (2.1) such that \( \omega_1(\lambda, z) + \omega_2(\lambda, z) = F_{\mu_1 \oplus \mu_2}(z) + z \). Hence, for \( \lambda \in D^{(e)}(T, a) \),
\[
f_{\mu_1 \oplus \mu_2}(0) = -\frac{1}{\pi} \lim_{\eta \to 0} \Im G_{\mu_1 \oplus \mu_2}(i\eta) \in \left( \frac{\varepsilon}{2\pi}, \frac{1}{2\pi\varepsilon} \right).
\]
Hence by definition of \( \Omega \) (see also Theorem 2.3), \( 0 \in \mathcal{B}_{\mu_1 \oplus \mu_2} \). By applying Theorem 2.8 for the choice \( J = \{0\} \), (4.3) holds uniformly for \( 0 \leq \eta \leq N^L \) for \( \lambda \in D^{(e)}(T, a) \) fixed.

Fix some large \( L \) so that (4.4) holds for any \( \lambda \in D^{(e)}(T, a) \). We next adapt the approach in the proof of [2] Theorem 2.5 for \( 0 \leq \eta \leq N^L \). By the definition of stochastic domination and (4.3) for any \( \lambda \in D^{(e)}(T, a) \) fixed, the number of elements in the lattice \( D^{(e)}(T, a) \cap \{N^{-L_1} + iN^{-L_1} \} \) is of order \( N^{2L_1} \), hence we have
\[
\max_{\lambda \in D^{(e)}(T, a) \cap N^{-L_1} \{Z \times i\mathbb{Z} \}} |G^\lambda(i\eta) - G_{\Sigma,|A-\lambda|}(i\eta)| < \frac{1}{N\eta}.
\]
That is, the inequality (4.3) holds for \( \lambda \) at the lattice points in \( D^{(e)}(T, a) \), where the coordinates of these points are multiple of \( N^{-L_1} \).

It remains to prove some Lipschitz type inequality with respect to the parameter \( \lambda \). Let \( L_1 \) be a large positive number such that \( L_1 \geq 2L \). We claim that, for \( \lambda_1, \lambda_2 \in D^{(e)}(T, a) \) with \( |\lambda_1 - \lambda_2| \leq N^{-L_1} \), we have
\[
|G^{\lambda_1}(i\eta) - G^{\lambda_2}(i\eta)| < \frac{1}{N\eta} \quad (4.5)
\]
and
\[
|G_{\Sigma,|A-\lambda_1|}(i\eta) - G_{\Sigma,|A-\lambda_2|}(i\eta)| < \frac{1}{N\eta} \quad (4.6)
\]
uniformly in \( N^{-L} \leq N^{-L_0} \leq \eta \leq N^L \). Recall the definition of the random matrix \( H^\lambda \) (4.1). By the resolvent identity, if \( \eta \geq N^{-L} \), we have
\[
|G^{\lambda_1}(i\eta) - G^{\lambda_2}(i\eta)| = \left| \frac{1}{2N} \text{Tr}(i\eta - H^{\lambda_1})^{-1} \frac{1}{2N} \text{Tr}(i\eta - H^{\lambda_2})^{-1} \right|
\leq \frac{|\lambda_1 - \lambda_2|}{2N} |\text{Tr}( (i\eta - H^{\lambda_1})^{-1}(i\eta - H^{\lambda_2})^{-1})| \\
\leq |\lambda_1 - \lambda_2| \cdot \| (H^{\lambda_1} - i\eta)^{-1} \cdot (H^{\lambda_2} - i\eta)^{-1} \| \\
\leq \frac{|\lambda_1 - \lambda_2|}{\eta^2} \leq \frac{C}{\eta} \frac{1}{N} \leq \frac{1}{N\eta}.
\]
This establishes (4.5).
We now apply [4, Equation (2.20)] to get
\[
|G_{\Sigma, |A-\lambda|}(i\eta) - G_{\Sigma, |A-\lambda_2|}(i\eta)| \leq \frac{C}{\eta} \left( 1 + \frac{1}{\eta} \right) d_L(\tilde{\mu}_{|A-\lambda|}, \tilde{\mu}_{|A-\lambda_2|}),
\]
for all \( \eta > 0 \) and some constant \( C \) independent from \( \eta \). Note that \( \tilde{\mu}_{|A-\lambda|} \) has the same distribution as the eigenvalue distribution of the matrix \( \begin{bmatrix} 0 & A - \lambda \\ A^* - \frac{1}{\lambda_1} & 0 \end{bmatrix} \). By some standard inequality for spectral measure as in [21] Proposition 1.6 (iii), we have
\[
d_L(\tilde{\mu}_{|A-\lambda|}, \tilde{\mu}_{|A-\lambda_2|}) \leq \left\| \begin{bmatrix} 0 & A - \lambda_1 \\ A^* - \frac{1}{\lambda_1} & 0 \end{bmatrix} - \begin{bmatrix} 0 & A - \lambda_2 \\ A^* - \frac{1}{\lambda_2} & 0 \end{bmatrix} \right\| \\
\leq |\lambda_1 - \lambda_2| \leq \frac{1}{N L_1} \leq \frac{1}{N}.
\]
This proves (4.6). We conclude that (4.3) holds uniformly in \( \lambda \) for \( 0 \leq \eta \leq N^L \). Since we already prove that (4.3) holds uniformly in \( \eta > N^L \) in the first two paragraph of the proof, the theorem is established.

4.2. Proof of the local convergence. In this section, we prove a deformed local single ring theorem for the random matrix model \( Y \) in (1.2). Write \( \lambda_1, \ldots, \lambda_N \) to be the random eigenvalues of \( Y \). The random matrix \( Y \) converges in \(*\)-distribution to \( y \) in (1.3). Recall that we assume \( \|A\|, \|\Sigma\| \leq M \) for all \( N \). We denote
\[
H^\lambda = \begin{pmatrix} 0 & Y - \lambda \\ (Y - \lambda)^* & 0 \end{pmatrix}.
\]

Theorem 4.2. Consider the random matrix model \( Y \) and the operator \( y \in \mathcal{A} \) as in (1.2) and (1.3). Write \( \lambda_1, \ldots, \lambda_N \) be the (random) eigenvalues of \( Y \). Assume further that the set \( \Gamma_N \) in Assumption (1) has Lebesgue measure zero for all \( N \).

Let \( R > 0 \). For any compact set \( K \subset D(T, a), \beta \in (0, 1/2), w_0 \in K \) and any \( f : \mathbb{C} \to \mathbb{R} \) be a smooth function supported in the disk centered at 0 of radius \( R \), define the function by
\[
f_{w_0}(w) = N^{2\beta} f(N^\beta (w - w_0)).
\]
Let \( a_N, \sigma_N \in \mathcal{A} \) be such that \( a_N \) and \( \sigma_N \) have the same \(*\)-distribution as \( A_N \) and \( \Sigma_N \). Also let \( T_N \in \mathcal{A} \) be an \( R \)-diagonal operator such that \( \mu_{|T_N|} = \mu_{\Sigma_N} \). Write
\[
y_N = a_N + T_N.
\]
Then
\[
\frac{1}{N} \sum_{k=1}^{N} f_{w_0}(\lambda_k) - \int_{\mathbb{C}} f_{w_0}(w) \, d\mu_{y_N}(w) \prec N^{-1+2\beta} \|\Delta f\|_{L^1(\mathbb{C})},
\]
(4.7)
where \(\mu_{y_N}\) is the Brown measure of \(y_N\). This convergence is uniform in \(f\) and in \(w_0 \in K\), for all large enough \(N\) depending on \(K, R, M, a\) and \(T\). (The constant \(M\) is defined in (1.1).)

The test bump function \(f_{w_0}\) is supported in a disk of radius \(RN^{-\beta}\), whose Lebesgue measure goes to 0 as \(N \to \infty\). Therefore, in this theorem, we make a stronger assumption that \(\Gamma_N\) has Lebesgue measure zero for all \(N\). There are examples where this assumption holds; a trivial example is \(\|\Sigma\| \geq N^{-\alpha}\) for some \(\alpha > 0\).

We need several lemmas before we can prove Theorem 4.2.

Lemma 4.3. Let \(K \subset \mathbb{C}\) be compact and let \(\mu_\sigma\) be a probability measure supported on \([0, \infty)\). For any \(\lambda \in \mathbb{C}\), denote by \(G_{\sigma,|a-\lambda|}(i\eta)\) the Cauchy transform of \(\tilde{\mu}_\sigma \boxplus \tilde{\mu}_{|a-\lambda|}\). If \(G_{\sigma,|a-\lambda|}(0)\) is uniformly bounded for all \(\lambda \in K\), then there exists \(L_1 > 1\) such that for any \(\lambda \in K\)
\[
\int_0^{N-L_1} G_{\sigma,|a-\lambda|}(i\eta) \, d\eta \leq \frac{1}{N}
\]
(4.8)
for all large enough \(N\).

Proof. By Proposition 2.1, \(G_{\sigma,|a-\lambda|}(i\eta)\) is uniformly bounded for \(\lambda \in K\) and \(\eta \in [0, 1]\). Thus, we can choose \(L_1\) large enough such that (4.8) holds. \(\square\)

Lemma 4.4. There exist positive constants \(c\) and \(\tilde{c}\) such that for any \(\lambda \in \Gamma_N^c\) and \(L_1 > 0\),
\[
E \left| \int_0^{N-L_1} G^\lambda(i\eta) \, d\eta \right| \leq N^{-cL_1/2+c}.
\]

Proof. Denote \(\sigma_1^\lambda\) to be the least singular value of \(Y - \lambda\); that is, the least nonnegative eigenvalue of \(H^\lambda\). Then since \(G^\lambda\) is the Cauchy transform of the eigenvalues of \(H^\lambda\),
\[
E \left| \int_0^{N-L_1} G^\lambda(i\eta) \, d\eta \right| \leq E \int_0^{N-L_1} \frac{\eta}{(\sigma_1^\lambda)^2 + \eta^2} \, d\eta
= \frac{1}{2} E[\log(1 + (N^{L_1} \sigma_1^\lambda)^{-2})]
= \frac{1}{2} \int_0^{\infty} \mathbb{P}(\log(1 + (N^{L_1} \sigma_1^\lambda)^{-2}) \geq s) \, ds
= \frac{1}{2} \int_0^{\infty} \mathbb{P} \left( \sigma_1^\lambda \leq N^{-L_1} \frac{1}{\sqrt{e^s - 1}} \right) \, ds.
\]
(4.9)
We decompose the integral into three parts \(\int_0^{N-L_1} + \int_{N-L_1}^1 + \int_1^{\infty}\). For the first integral, it is straightforward to see that
\[
\int_0^{N-L_1} \mathbb{P} \left( \sigma_1^\lambda \leq N^{-L_1} \frac{1}{\sqrt{e^s - 1}} \right) \, ds \leq N^{-L_1}.
\]
For the second integral, we estimate that \( \frac{1}{\sqrt{e^s - 1}} \leq \frac{1}{\sqrt{s}} \leq N^{-L1/2} \) for all \( N^{-L1} \leq s \leq 1 \). Since \( \lambda \in \Gamma_N \), we can apply Theorem 3.2. There exist positive constants \( c \) and \( c' \)

\[
\int_{N^{-L1}}^{1} \mathbb{P} \left( \sigma_1^{\lambda} \leq N^{-L1} \frac{1}{\sqrt{e^s - 1}} \right) \, ds \leq \int_{N^{-L1}}^{1} \mathbb{P}(\sigma_1^{\lambda} \leq N^{-L1/2}) \leq N^{-cL1/2 + c'}.
\]

Finally, for the third integral, using \( e^s - 1 > \frac{1}{2} e^s \) for all \( s \geq 1 \), we have

\[
\int_{1}^{\infty} \mathbb{P} \left( \sigma_1^{\lambda} \leq N^{-L1} \frac{1}{\sqrt{e^s - 1}} \right) \, ds \leq \int_{1}^{\infty} \mathbb{P}(\sigma_1^{\lambda} \leq N^{-L1/2} \sqrt{2} e^{-s/2}) \, ds
\]

\[
\leq \int_{1}^{\infty} 2^{c/2} N^{-cL1 + c'} e^{-cs/2} \, ds
\]

\[
= \frac{2^{1+c/2} e^{-c/2}}{c} N^{-cL1 + c'}
\]

for some positive constants \( c \) and \( c' \). Put these three estimates of integrals to (4.9), we have, for some positive constants \( c \) and \( \bar{c} \), the conclusion holds. \( \Box \)

We are now ready to prove Theorem 4.2.

**Proof of Theorem 4.2.** We do a change of variable \( \lambda = N^{\beta}(w - w_0) \) and write

\[
\frac{1}{N} \sum_{k=1}^{N} f_{w_0}(\lambda_k) - \int_{C} f_{w_0}(w) \, d\mu_{yN}(w)
\]

\[
= N^{2\beta} \frac{1}{2\pi} \int_{C} (\Delta f)(\lambda) \left( \frac{1}{2N} \text{Tr} \log |H^w| - \int_{\mathbb{R}} \log |u| d\mu_{\Sigma|A-w|}(u) \right) \, d^2\lambda. \tag{4.10}
\]

By modifying the approach in [3], we will prove that

\[
\int_{C} (\Delta f)(\lambda) \left( \frac{1}{2N} \text{Tr} \log |H^w| - \int_{\mathbb{R}} \log |u| d\mu_{\Sigma|A-w|}(u) \right) \, d^2\lambda \ll \frac{\|\Delta f\|_{L^1(\mathbb{C})}}{N} \tag{4.11}
\]

uniformly in \( w \in \overline{W} \) where \( W \) is any fixed neighborhood of \( K \) such that \( \overline{W} \subset D(T, a) \) is compact. Note that for all \( N \) large enough, \( f_{w_0} \) is supported in \( \overline{W} \) for all \( w_0 \in K \).

We now estimate (4.11). For any \( L > 0 \) and \( w \), write

\[
\frac{1}{2N} \text{Tr} \log |H^w| = \frac{1}{2N} \text{Tr} \log |(H^w - iN^L)| + \text{Im} \int_{0}^{N^L} G^w(i\eta) \, d\eta
\]

\[
\int_{\mathbb{R}} \log |u| d\mu_{\Sigma|A-w|}(u) = \int_{\mathbb{R}} \log |u - iN^L| d\mu_{\Sigma|A-w|}(u) + \text{Im} \int_{0}^{N^L} G_{\Sigma|A-w|}(i\eta) \, d\eta. \tag{4.12}
\]

It is clear that there is a constant \( C \) such that \( \|H^w\| \leq C \) for all \( w \) in any ball of finite radius. Since the support of \( f_{w_0} \) lies in a ball of radius \( RN^{-\alpha} \), we can then choose \( L \) large enough so that

\[
\left| \frac{1}{2N} \text{Tr} \log |H^w - iN^L| - \int_{\mathbb{R}} \log |u - iN^L| d\mu_{\Sigma|A-w|}(u) \right| \ll \frac{1}{N}. \tag{4.13}
\]

Thus, it remains to estimate the second terms in (4.12); we will show

\[
\left| \int_{C} (\Delta f)(\lambda) \left( \int_{0}^{N^L} (G^w(i\eta) - G_{\Sigma|A-w|}(i\eta)) \, d\eta \right) \, d^2\lambda \right| \ll \frac{\|\Delta f\|_{L^1(\mathbb{C})}}{N}. \tag{4.14}
\]
We will choose large enough constants $L_1$ and $L$ and decompose the integral with respect to $\eta$ into
\[
\int_0^{N_L} = \int_0^{N_{-L_1}} + \int_{N_{-L_1}}^{N_L}.
\]

We first analyze the integral $\int_0^{N_{-L_1}}$. For the fixed neighborhood $W \subset D(T, a)$ of $K$, recall that there exists $\varepsilon > 0$ such that $\overline{W} \subset D^{(\varepsilon)}(T, a)$. If $N$ is large enough, the support of $f_{w_0}$ must also lie in $\overline{W}$ for all $w_0 \in K$. We then derive from Theorem 4.1 that the stochastic domination
\[
\left| \int_C (\Delta f)(\lambda) \left( \int_0^{N_{-L_1}} (G^w(\eta)) - G_{\Sigma, [A-w]}(\eta) \right) \right| d^2 \lambda \leq \int_C |(\Delta f)(\lambda)| \left( \int_{N_{-L_1}}^{N_L} \frac{1}{N} d\eta \right) d^2 \lambda < \frac{||\Delta f||_{L^1(C)}}{N}.
\]

Next, we analyze the integral $\int_{N_{-L_1}}^{N_L}$. Our strategy is to choose $L_1$ large enough so that
\[
\left| \int_C (\Delta f)(\lambda) \left( \int_0^{N_{-L_1}} G_{\Sigma, [A-w]}(\eta) d\eta \right) \right| d^2 \lambda \leq \frac{||\Delta f||_{L^1(C)}}{N} \quad (4.15)
\]
and
\[
\mathbb{P} \left( \left| \int_C (\Delta f)(\lambda) \left( \int_0^{N_{-L_1}} G^w(\eta) d\eta \right) \right| d^2 \lambda \geq \frac{||\Delta f||_{L^1(C)}}{N} \right) \ll \frac{1}{N}. \quad (4.16)
\]

By Proposition 2.2 and Definition 2.4, $G_{\Sigma, [A-w]}(0)$ is uniformly bounded for all $w \in \overline{W}$ and all large enough $N$. Applying Lemma 4.3 with $\overline{W}$ in place of $K$ and $\Sigma$ in place of $\xi$, we can choose $L_1$ large enough such that (4.15) holds uniformly for test function $f$ supported in the disk centered at 0 of radius $R$.

To show (4.16), we use Markov’s inequality to deduce
\[
\mathbb{P} \left( \left| \int_C (\Delta f)(\lambda) \left( \int_0^{N_{-L_1}} G^w(\eta) d\eta \right) \right| d^2 \lambda \geq \frac{||\Delta f||_{L^1(C)}}{N} \right) \leq \frac{N}{||\Delta f||_{L^1(C)}} \mathbb{E} \left| \int_C (\Delta f)(\lambda) \left( \int_0^{N_{-L_1}} G^w(\eta) d\eta \right) \right| d^2 \lambda. \quad (4.17)
\]

We can apply Lemma 4.4 except for $w \in \Gamma_N$ which has Lebesgue measure 0. By Lemma 4.3, there exist positive constants $c$ and $\tilde{c}$ such that we can estimate (4.17) by
\[
\mathbb{E} \left| \int_C (\Delta f)(\lambda) \left( \int_0^{N_{-L_1}} G^w(\eta) d\eta \right) \right| d^2 \lambda \leq \int_C |(\Delta f)(\lambda)| \mathbb{E} \left| \int_0^{N_{-L_1}} G^w(\eta) d\eta \right| d^2 \lambda \leq ||\Delta f||_{L^1(C)} N^{-cL_1/2 + \tilde{c}}.
\]

Therefore, using (4.17), we can choose $L_1$ large enough such that (4.16) also holds.

We have proved (4.14) by decomposing the integral into $\int_0^{N_{-L_1}}$ and $\int_{N_{-L_1}}^{N_L}$. Combining with the estimate (4.13), using (4.10) and (4.12), we conclude that (4.7) holds. □


5. The deformed single ring theorem

In this section, we prove a deformed single ring theorem for the random matrix model $Y$ in \((1.2)\). Recall that $Y$ converges in \(\ast\)-distribution to $y$ in \((1.3)\). Recall also that we assume \(\|\Sigma_N\| \leq M\) for all $N$ and $\Omega$ is defined in \((2.4)\). We denote

\[
H^\lambda = \begin{pmatrix} 0 & Y - \lambda^* \\ (Y - \lambda) & 0 \end{pmatrix}
\]

and write $m$ to be the Lebesgue measure on $\mathbb{C}$.

**Theorem 5.1.** Consider the random matrix model $Y$ and the operator $y \in A$ as in \((1.2)\) and \((1.3)\). Suppose that Assumption \(1.1\) holds. Also assume at least one of the following is true:

1. there exists a Lebesgue measurable set $E \subset \Omega^c$ with $m(E) = 0$ satisfying the following property: for any compact set $S \subset \Omega^c \cap E^c$, there exist constants $\kappa_1, \kappa_2 > 0$ such that

\[
|G_{\tilde{\mu}_{A-\lambda}}(i\eta)| \leq \kappa_2
\]

for all $\eta > N^{-\kappa_1}$ and $\lambda \in S$; or

2. there exist constants $\kappa_1, \kappa_2 > 0$ such that

\[
|G_{\tilde{\mu}_\sigma}(i\eta)| \leq \kappa_2
\]

for all $\eta > N^{-\kappa_1}$.

Then the empirical eigenvalue distribution of $Y$ converges weakly to the Brown measure of $y$ in probability.

Before we prove Theorem 5.1, we first show in the following theorem that Theorem 5.1 applies to a large family of operators that is of particular interest. By taking $A = 0$, the following theorem recovers the single ring theorem by Guionnet et al. \cite{25} with weaker conditions.

**Theorem 5.2.** Consider the random matrix model $Y$ and the operator $y \in A$ as in \((1.2)\) and \((1.3)\). If $A$ is Hermitian for all $N$ or if $A$ is unitary for all $N$, then the empirical eigenvalue distribution of $Y$ converges weakly to the Brown measure of $y$ in probability.

More generally, if $A$ is a normal matrix and there is a closed set $F$ with Lebesgue measure zero independent of $N$ such that all the eigenvalues of $A$ lie inside $F$ for all $N$, then the empirical eigenvalue distribution of $Y$ converges weakly to the Brown measure of $y$ in probability.

**Proof.** We prove the more general statement for normal matrix $A$ satisfying the hypothesis. Assumption \(1.1\) holds under the given hypothesis, by Remark \(1.2\). We then show that Assumption \(1.1\) in Theorem 5.1 holds. Take $E = F \cap \Omega^c$. Then any compact set $S \subset \Omega^c \cap E^c$ has positive distance from $E$. The conclusion then follows from Theorem 5.1. 

Now we proceed to prove Theorem 5.1. We prove the following lemmas.

**Lemma 5.3.** Let $\varphi$ be a $C^\infty_c(\mathbb{C})$ function and denote the Lebesgue measure on $\mathbb{C}$ by $m$. For any $\varepsilon > 0$, there exists $\delta > 0$ independent of $N$ such that whenever $E$ is a Lebesgue
measurable set satisfying \( m(E) < \delta \), we have
\[
\left| \int_E \varphi(\lambda) \frac{1}{N} \text{Tr}[\log |Y - \lambda|] d^2\lambda \right| < \varepsilon \quad \text{(5.1)}
\]
\[
\left| \int_E \varphi(\lambda) \tau[\log |y - \lambda|] d^2\lambda \right| < \varepsilon. \quad \text{(5.2)}
\]
In other words, the logarithmic potentials of \( Y \) and \( y \) are uniformly locally integrable for all \( N \).

**Proof.** Let \( \varepsilon > 0 \) be given. The function \( \lambda \mapsto \log |x - \lambda| \) are uniformly locally integrable for all \( |x| \leq 2M \); that is, there exists \( \delta > 0 \) such that
\[
\left| \int_E \varphi(\lambda) \log |x - \lambda| d^2\lambda \right| < \varepsilon \quad \text{(5.3)}
\]
whenever \( m(E) < \delta \) and \( |x| \leq M \). By replacing \( x \) in (1.3) by \( y \), an application of Fubini–Tonelli theorem shows that, with the same \( \delta > 0 \),
\[
\left| \int_E \varphi(\lambda) \tau[\log |y - \lambda|] d^2\lambda \right| = \int_{|z| \leq 2M} \left| \int_E \varphi(\lambda) \log |z - \lambda| d^2\lambda \right| d\mu_y(z) < \varepsilon
\]
whenever \( m(E) < \delta \). We have used the fact that \( ||y|| \leq 2M \).

Now, we first note that
\[
\frac{1}{N} \text{Tr}[\log |Y - \lambda|] = \frac{1}{N} \sum_{j=1}^N \log |\lambda_j - \lambda|
\]
where \( \lambda_j \) are the eigenvalues of \( Y \). By assumption, \( ||Y|| \leq ||A|| + ||\Sigma|| \leq 2M \), we must have \( \max_j |\lambda_j| \leq 2M \). By (5.3),
\[
\left| \int_E \varphi(\lambda) \frac{1}{N} \text{Tr}[\log |Y - \lambda|] d^2\lambda \right| \leq \frac{1}{N} \sum_{j=1}^N \left| \int_E \varphi(\lambda) \log |x - \lambda| d^2\lambda \right| < \varepsilon
\]
whenever \( m(E) < \delta \). This shows (5.1) and establishes the lemma.

**Lemma 5.4.** Fix a compact set \( K \subset \Omega^c \) and a bounded nonnegative Borel function \( \varphi \) on \( \mathbb{C} \). Given any \( \varepsilon > 0 \), there exists \( t > 0 \) such that
\[
\int_K \varphi(\lambda) \int_{[0,t]} |\log |x|| d\mu_{|y-\lambda|}(x)d^2\lambda < \varepsilon.
\]

**Proof.** By Proposition 2.5 \( G_{\bar{\mu}_{|y-\lambda|}}(ix) \) is uniformly bounded for all \( \lambda \in K \) and \( x \geq 0 \). By \[25\] Lemma 15], there exists a constant \( C > 0 \) such that
\[
\bar{\mu}_{|y-\lambda|}([-x, x]) \leq 2x \text{ Im} G_{\bar{\mu}_{|y-\lambda|}}(ix) \leq Cx.
\]
By \[8\] Lemma 4.1(a)], there is a constant \( C' > 0 \) such that
\[
\int_{[0,t]} |\log |x||^2 d\mu_{|y-\lambda|}(x) \leq C't(1 - \log t).
\]
It follows that
\[
\int_K \varphi(\lambda) \int_{[0,t]} |\log |x|| d\mu_{|y-\lambda|}(x)d^2\lambda \leq C'\|\varphi\|_\infty m(K)t(1 - \log t)
\]
where \( m(K) \) is the Lebesgue measure of \( K \). Now, it is evident that we can choose small enough \( t \) such that the conclusion of the lemma holds. □

**Lemma 5.5.** Fix any \( K \subset D(T, a) \). There exists a constant \( C > 0 \) such that whenever \( w \in K \),
\[
\left| \int \log |u| \, d\mu_{\sigma, |a-w|}(u) - \int \log |u| \, d\mu_{\Sigma, |A-w|}(u) \right| \leq C(d_L(\mu_\sigma, \mu_\Sigma) + d_L(\mu_{|a-w|}, \mu_{|A-w|}))
\]
for all large enough \( N \).

**Proof.** We use the following observation due to [33] (see also (3.6) and (3.7) from [3]) to write
\[
\int \log |u| \, d\mu_{\sigma, |a-w|}(u) = \int \log |u - i| \, d\mu_{\sigma, |a-w|}(u) + \text{Im} \int_0^1 G_{\sigma, |a-w|}(i\eta) \, d\eta.
\]
(5.4)

Since \( \log |u - i| \) is a smooth function, by applying the continuity of free convolution [10, Theorem 4.13], we have
\[
\left| \int \log |u - i| \, d\mu_{\sigma, |a-w|}(u) - \int \log |u - i| \, d\mu_{\Sigma, |A-w|}(u) \right| \leq C(d_L(\mu_\sigma, \mu_\Sigma) + d(\mu_{|a-w|}, \mu_{|A-w|})).
\]
(5.5)

By the local stability of Cauchy transform in the bulk [4, Theorem 2.7], we have
\[
\left| \int_0^1 G_{\sigma, |a-w|}(i\eta) \, d\eta - \int_0^1 G_{\Sigma, |A-w|}(i\eta) \, d\eta \right| \leq C(d_L(\mu_\sigma, \mu_\Sigma) + d_L(\mu_{|a-w|}, \mu_{|A-w|})),
\]
(5.6)

for any \( w \) such that \( 0 \in B_{\mu_1 \oplus \mu_2} \), where \( \mu_1 = \tilde{\mu}_{|a-w|} \) and \( \mu_2 = \tilde{\mu}_{|\sigma|} = \tilde{\mu}_{|T|} \). Note that there exists \( \varepsilon > 0 \) such that \( K \subset D(\varepsilon)(T, a) \). By the proof of [4, Theorem 2.7] (see also [4, Lemma 3.4] and [4, Lemma 5.1]), the constant \( C \) in (5.5) and (5.6) can be chosen independently from \( w \in K \). □

**Lemma 5.6.** Let \( K \subset \mathbb{C} \) be a compact set. Then
\[
d_L(\mu_{|a-w|}, \mu_{|A-w|}) \to 0
\]
uniformly in \( w \in K \).

**Proof.** Since \( w \mapsto d_L(\mu_{|a-w|}, \mu_{|A-N-w|}) \) is a continuous function in \( w \), we can choose \( w_N = \arg \max_{w \in K} d_L(\mu_{|a-w|}, \mu_{|A-N-w|}) \). We claim that
\[
d_L(\mu_{|a-w_N|}, \mu_{|A-N-w_N|}) \to 0.
\]
Suppose that the claim is false. There is a subsequence \( w_{N_k} \) of \( w_N \) such that
\[
d_L(\mu_{|a-w_{N_k}|}, \mu_{|A-w_{N_k}|}) \geq \varepsilon_0
\]
for some \( \varepsilon_0 > 0 \). By dropping to a subsequence if necessary, we may assume \( w_{N_k} \to w_0 \) for some \( w_0 \in K \). However, this contradicts that
\[
d_L(\mu_{|a-w_{N_k}|}, \mu_{|A-w_{N_k}|}) \to 0
\]
because the moments of \( \mu_{|A-w_{N_k}|} \) converges to \( \mu_{|a-w_0|} \). This proves the lemma. □
Proof of Theorem 5.1. Let $f$ be a $C_c^\infty(\mathbb{C})$ smooth function and write $K$ to be the support of $f$. Let $\varepsilon > 0$. We want to show that

$$\mathbb{P}\left(\left|\int_K (\Delta f)(\lambda) \frac{1}{N} \text{Tr}[\log |Y - \lambda|] d^2\lambda - \int_K (\Delta f)(\lambda) \tau |\log |y - \lambda|| d^2\lambda \right| > \varepsilon\right) \to 0. \quad (5.7)$$

Recall that $S(T,a)$ is finite. Since log-potentials are uniformly locally integrable in the sense of Lemma 5.3, there exists an open set $W$ such that $E \cup S(T,a) \subset W$ and

$$\left|\int_{W^1} (\Delta f)(\lambda) \frac{1}{N} \text{Tr}[\log |Y - \lambda|] d^2\lambda \right| + \left|\int_{W^1} (\Delta f)(\lambda) \tau |\log |y - \lambda|| d^2\lambda \right| < \frac{\varepsilon}{4}. \quad (5.8)$$

By doing a compact exhaustion, there exist compact sets $F^{(1)} \subset F^{(2)} \subset \ldots \subset \Omega$ such that

$$\bigcup_{k=1}^{\infty} F^{(k)} = \Omega.$$ 

Since $\Omega$ is a bounded set, by Lemma 5.3, there exists $n$ such that $5.8$ holds with $\Omega \setminus F^{(n)}$ in place of $W_1$. We decompose $(K \cap W^1) \cap \Omega$ into $W_2 \cup K_1$ where $K_1 = F^{(n)}$ and $W_2 = ((K \cap W^1) \cap \Omega) \setminus F^{(n)}$. By writing $K_2 = (K \cap W^1) \cap \Omega^c$, we decompose $K$ into the disjoint union

$$K = (K \cap W_1) \cup W_2 \cup K_1 \cup K_2.$$ 

By construction, $K_1 \subset \Omega$ and $K_2 \subset \Omega^c$ are compact sets, and $5.8$ holds with $W_2$ in place of $W_1$ in the equation. In the following paragraphs, we will estimate the integral on the left-hand side of $(5.7)$ over $K_1$ and $K_2$ instead of $K$ separately.

We first look at the integral on the left-hand side of $(5.7)$ over $K_1$. We want to show that

$$\int_{K_1} (\Delta f)(\lambda) \frac{1}{N} \text{Tr}[\log |Y - \lambda|] d^2\lambda - \int_{K_1} (\Delta f)(\lambda) \tau |\log |y - \lambda|| d^2\lambda \to 0$$

in probability. Since $m(\Gamma_N) \to 0$ as $N \to \infty$ where $\Gamma_N$ is defined in Assumption 1.1, we must have

$$\int_{K_1 \cap \Gamma_N} (\Delta f)(\lambda) \frac{1}{N} \text{Tr}[\log |Y - \lambda|] d^2\lambda - \int_{K_1 \cap \Gamma_N} (\Delta f)(\lambda) \tau |\log |y - \lambda|| d^2\lambda \to 0 \quad (5.9)$$

Thus, for $K_1$, it suffices to show that

$$\int_{K_1 \setminus \Gamma_N} (\Delta f)(\lambda) \left(\frac{1}{2N} \text{Tr} \log |H^\lambda| - \int_{\mathbb{R}} \log |u| d\mu_{\Sigma,|A-w|}\right) d^2\lambda \lesssim \frac{\|\Delta f\|_{L^1(\mathbb{C})}}{N}, \quad (5.10)$$

$$\int_{K_1 \setminus \Gamma_N} (\Delta f)(\lambda) \left(\int_{\mathbb{R}} \log |u| d\mu_{\Sigma,|A-w|}^2 d\lambda - \int_{\mathbb{R}} \log |u| d\mu_{\sigma,|A-w|}^2 d\lambda\right) \to 0. \quad (5.11)$$

We first note that the convergence $(5.11)$ is deterministic and it follows from Lemmas 5.5 and 5.9.

We now prove $(5.10)$. For any $L > 0$ and $\lambda \in K_1$, write the logarithmic potentials $\frac{1}{2N} \text{Tr} \log |H^\lambda|$ and $\int_{\mathbb{R}} \log |u| d\mu_{\Sigma,|A-w|}$ as in $(1.12)$. Since $K_1$ is compact, the estimate $(1.13)$ holds uniformly for all $\lambda \in K_1$. As in the proof of Theorem 1.2 in order to prove $(5.10)$, it suffices to show

$$\left|\int_{K_1 \setminus \Gamma_N} (\Delta f)(\lambda) \left(\int_0^{N^L} (G^\lambda(i\eta) - G_{\Sigma,|A-\lambda|}(i\eta)) d\eta\right) d\lambda\right| < \frac{\|\Delta f\|_{L^1(\mathbb{C})}}{N}, \quad (5.12)$$
The strategy is similar to that in Theorem 4.2. We decompose the integral with respect to $\eta$ into

$$\int_{0}^{N_L} = \int_{0}^{N_L-1} + \int_{N_L}^{N_L}.$$  

Since $K_1$ is compact, there exists $\delta > 0$ such that $K_1 \setminus \Gamma_N \subset K_1 \subset D^{(\delta)}(T,a)$. By Theorem 4.1, we have the stochastic domination

$$\int_{K_1 \setminus \Gamma_N} (\Delta f)(\lambda) \left( \int_{0}^{N_L} (G^\lambda(i\eta) - G_{\Sigma,|A-\lambda|}(i\eta)) d\eta \right) d^2 \lambda \leq \frac{\| \Delta f \|_{L^1(\mathbb{C})}}{N}.$$  

(5.13)

We now estimate the integral $\int_{0}^{N_L-1}$. By Proposition 2.2 and Definition 2.4, $G_{\Sigma,|A-\lambda|}(0)$ is uniformly bounded for all $\lambda \in K_1$ and for all large enough $N$. Applying Lemma 4.3 with $K_1$ in place of $K$ and $\sigma$ in place of $\xi$, we see that we can choose $L_1$ large enough such that

$$\int_{K_1 \setminus \Gamma_N} (\Delta f)(\lambda) \left( \int_{0}^{N_L-1} G_{\Sigma,|A-\lambda|}(i\eta) d\eta \right) d^2 \lambda \leq \frac{\| \Delta f \|_{L^1(\mathbb{C})}}{N}.$$  

(5.14)

Meanwhile, we apply Lemma 4.4 and use Markov’s inequality to get the approximation

$$\mathbb{P} \left( \int_{K_1 \setminus \Gamma_N} (\Delta f)(\lambda) \left( \int_{0}^{N_L-1} G^\lambda(i\eta) d\eta \right) d^2 \lambda \geq \frac{\| \Delta f \|_{L^1(\mathbb{C})}}{N} \right)$$

$$\leq \frac{N}{\| \Delta f \|_{L^1(\mathbb{C})}} \mathbb{E} \left| \int_{K_1 \setminus \Gamma_N} (\Delta f)(\lambda) \left( \int_{0}^{N_L-1} G^\lambda(i\eta) d\eta \right) d^2 \lambda \right|$$

$$\leq N^{cL_1/2+\bar{c}+1}$$  

(5.15)

for some positive constants $c$ and $\bar{c}$. By choosing large enough $L_1$, we can then combine (5.14), (5.15) and (5.16) to prove (5.12).

We now proceed to estimate the integral on the left-hand side of (5.7) over $K_2$ in place of $K$. Similar to the estimate for $K_1$, since $m(\Gamma_N) \to 0$ as $N \to \infty$ where $\Gamma_N$ is defined in Assumption 1.1, we must have

$$\int_{K_2 \cap \Gamma_N} (\Delta f)(\lambda) \frac{1}{N} \text{Tr}[\log |Y - \lambda|] d^2 \lambda - \int_{K_2 \cap \Gamma_N} (\Delta f)(\lambda) \tau \log |y - \lambda| d^2 \lambda \to 0.$$  

(5.17)

Hence, it suffices to estimate the integral on the left-hand side of (5.7) over $K_2 \setminus \Gamma_N$ instead of $K_2$. If Assumption 1 or 2 holds, without loss of generality, we assume $0 < \kappa_1 < 1/8$. We want to show that under one of these assumptions, there exists $\kappa_3, \kappa_4 > 0$ such that for all $N$ large enough, we have

$$\|EG^\lambda(i\eta)\| \leq \kappa_4 \quad \text{for all } \eta > N^{-\kappa_3} \text{ and } \lambda \in K_2.$$  

(5.18)

We first assume Assumption 1 holds. In this case, by Theorem 2.9, there are functions $\omega_A$ and $r_A$ such that

$$\mathbb{E}G^\lambda(i\eta) = \mathbb{E}G_{\bar{\mu},|A-\lambda|}(\omega_A(i\eta)) + r_A(i\eta).$$

We apply Assumption 1 with $S = K_2$. Fix a $\kappa_3$ such that $0 < \kappa_3 < \kappa_1 < 1/8$. If $N$ is large enough, by Lemma 2.10, $\omega_A(i\eta)$ is purely imaginary with $\text{Im} \omega_A(i\eta) > N^{-\kappa_1}$ for all $\eta > N^{-\kappa_3}$ and $\lambda \in K_2$; moreover, $r_A(i\eta) \to 0$ uniformly in $\eta > N^{-\kappa_3}$ and $\lambda \in K_2$. 

Assumption (1) tells us that $|\mathbb{E} G_{\rho_{A - \lambda}}(\omega_A(i\eta))| \leq \kappa_2$; thus, (5.18) holds for some $\kappa_4 > 0$. If Assumption (2) holds instead of Assumption (1), we can show (5.18) in a similar manner, by applying Theorem 2.8 that there are functions $\omega_B$ and $\tau_B$ such that

$$
\mathbb{E} G^\lambda(i\eta) = \mathbb{E} G^\lambda_{\rho_{\lambda}}(\omega_B(i\eta)) + \tau_B(i\eta).
$$

By (5.18) and [25, Lemma 15], for all $\lambda \in K_2$, $\mathbb{E} \mu_{[Y - \lambda]}([-x, x]) \leq 2\kappa_4 \max\{x, N^{-\kappa_3}\}$. For any $t > 0$, apply the Cauchy–Schwarz inequality to $\int_{[N^{-\kappa_3}, t]} \log |x| d\mu_{[Y - \lambda]}(x)$; by [8, Lemma 4.1(c)], there exists a constant $\kappa_5 > 0$ such that

$$
\mathbb{E} \int_{[N^{-\kappa_3}, t]} \log |x| d\mu_{[Y - \lambda]}(x) \leq \kappa_5 \sqrt{t} \log t + N^{-\kappa_3} \log N^2. \tag{5.19}
$$

Furthermore, given any $L_1 > 0$ which will be chosen later,

$$
\mathbb{E} \int_{[N^{-L_1}, N^{-\kappa_3}]} \log |x| d\mu_{[Y - \lambda]}(x) \leq L_1 \log N \mathbb{E} \left[ \mathbb{E} \int_{[N^{-L_1}, N^{-\kappa_3}]} \log |x| d\mu_{[Y - \lambda]}(x) \right] 
$$

$$
\leq 2\kappa_4 L_1 N^{-\kappa_3} \log N. \tag{5.20}
$$

We now choose $L_1 > 0$. Denote by $s_{\min}$ the least singular value of $Y - \lambda$. We compute

$$
\mathbb{E} \int_{[0, N^{-L_1}]} \log |x| d\mu_{[Y - \lambda]}(x) \leq \mathbb{E} \left[ \log s_{\min} \mathbb{1}_{s_{\min} \leq N^{-L_1}} \right] 
$$

$$
= \int_{0}^{N^{-L_1}} \mathbb{P}(s_{\min} \leq t) \frac{1}{t} dt + \mathbb{P}(s_{\min} \leq N^{-L_1}) \log N^{-L_1}.
$$

The above estimates are valid for $\lambda \in K_2$. We then apply Theorem 3.2 to $\lambda \in K_2 \setminus \Gamma_N$. There are positive constants $c$ and $c'$ such that for any $\lambda \in \Gamma_N^c$,

$$
\mathbb{E} \int_{[0, N^{-L_1}]} \log |x| d\mu_{[Y - \lambda]}(x) \leq \int_{0}^{N^{-L_1}} t^{-c'} N^c t \log N
$$

$$
= \left( \frac{1}{c'} + L_1 \log N \right) N^{-c' - L_1}. \tag{5.21}
$$

Hence, we choose $L_1 > 0$ such that $c' - L_1 = -\kappa_3$. By (5.19), (5.20) and (5.21), there is a constant $C > 0$ such that for all $\lambda \in K_2 \setminus \Gamma_N$,

$$
\mathbb{E} \int_{[0, t]} \log |x| d\mu_{[Y - \lambda]}(x) \leq \kappa_5 \sqrt{t} \log t^2 + N^{-\kappa_3} \log N^2 + CN^{-\kappa_3} \log N. \tag{5.22}
$$

By Lemma 5.4 given any $\delta > 0$, we can find $t$ such that the integral

$$
\int_{K_2 \setminus \Gamma_N} |\Delta f(\lambda)| \int_{[0, t]} \log |x| d\mu_{[Y - \lambda]}(x) d^2 \lambda < \delta
$$

for all $\lambda \in K_2 \setminus \Gamma_N$. This, together with (5.22), shows that, by Markov inequality, for any $\delta' > 0$, we can choose $t$ small enough such that

$$
\mathbb{P} \left( \left| \int_{K_2 \setminus \Gamma_N} (\Delta f)(\lambda) \left( \int_{[0, t]} \log x d\mu_{[Y - \lambda]}(x) - \int_{[0, t]} \log x d\mu_{[Y - \lambda]}(x) \right) d^2 \lambda \right| > \delta' \right) \to 0. \tag{5.23}
$$
Finally, recall that \(\|Y\| \leq 2M\), \(K_2\) is a bounded set in \(\mathbb{C}\). By weak convergence of \(\mu_{|Y-\lambda|}\) to \(\mu_{|y-\lambda|}\) in probability and by the dominated convergence theorem, for any \(t > 0\) and \(\delta' > 0\).

\[
P \left( \left| \int_{K_2 \setminus \Gamma_N} (\Delta f)(\lambda) \left( \int_{[t,\infty)} \log x \, d\mu_{|Y-\lambda|}(x) - \int_{[t,\infty)} \log x \, d\mu_{|y-\lambda|}(x) \right) \, d^2\lambda \right| > \delta' \right) \to 0. \tag{5.24}
\]

Combining (5.17), (5.23) and (5.24) with suitable \(\delta' > 0\), by triangle inequality, we see that

\[
P \left( \left| \int_{K_2} (\Delta f)(\lambda) \frac{1}{N} \text{Tr}[\log |Y - \lambda|] \, d^2\lambda - \int_{K_2} (\Delta f)(\lambda) \tau[\log |y - \lambda|] \, d^2\lambda \right| > \frac{\varepsilon}{4} \right) \to 0. \tag{5.25}
\]

By our choice of \(W_1\) and \(W_2\), (5.7) follows from (5.9), (5.10) and (5.25). This also completes the proof of the theorem. \(\square\)

6. Example: The Jordan Block Matrix

Consider a sequence of \(N \times N\) Jordan block matrices

\[
A = \begin{pmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 \\
0 & 0 & 0 & \cdots & 0
\end{pmatrix}. \tag{6.1}
\]

This matrix \(A\) is a well-known example that \(A\) converges in \(*\)-distribution to a Haar unitary operator \(a\) but the eigenvalue distribution does not converge to the Brown measure of the limit operator. In this section, we will show that, however, the empirical eigenvalue distribution of the random matrix \(U \Sigma V + A\) does converge to the Brown measure of the operator of its limit in \(*\)-distribution.

In this section, we will apply Theorem 5.1, we will show that Assumption 11 in the theorem statement holds for \(A\). The main result of the section is Proposition 6.5. In the following lemma, we study the singular values of \(A - \lambda\) for \(|\lambda| \neq 1\).

**Lemma 6.1.** For \(|\lambda| \neq 1\), the magnitudes of the singular values of \(A - \lambda\) can be described as follows.

1. If \(|\lambda| < 1\), all but one singular values of \(A - \lambda\) are at least \(1 - |\lambda|\).
2. If \(|\lambda| > 1\), all the singular values of \(A - \lambda\) are at least \(|\lambda| - 1\).

We first prove a more general result on estimating singular values. While we only state the result for matrices, the result, with the same proof, indeed holds for compact operators on a Hilbert space. For an \(N \times N\) matrix \(L\), we denote the singular values of \(L\) by

\[
s_1(L) \geq s_2(L) \geq \ldots \geq s_N(L).
\]

**Proposition 6.2.** Suppose that \(L\) and \(U\) are \(N \times N\) matrices such that \(U\) is unitary and \(L - U\) has rank \(k\). Then, for any \(\lambda \in \mathbb{C}\),

\[
s_n(\lambda - L) \geq s_{n+k}(\lambda - U).
\]

for all \(0 \leq n \leq n + k \leq N\). In particular, if \(|\lambda| < 1\),

\[
s_1(\lambda - L) \geq \ldots \geq s_{N-k}(\lambda - L) \geq 1 - |\lambda|.
\]
Proof. By [24, Theorem 2.1], for any $n$, the $n$th-largest singular value of any matrix $L$ can be computed as
\[ s_n(L) = \inf \{ \| L - X \| : \text{rank}(X) < n \}. \]
Since $L - U$ has rank $k$, $L - U + X$ has at most rank $n + k - 1$ for any $X$ has rank at most $n - 1$. We then apply the above formula to $\lambda - L$ and get
\[ s_n(\lambda - L) = \inf \{ \| \lambda - L - X \| : \text{rank}(X) < n \} \]
\[ = \inf \{ \| (\lambda - U) - (L - U + X) \| : \text{rank}(X) < n \} \]
\[ \geq \inf \{ \| (\lambda - U) - Z \| : \text{rank}(Z) < n + k \} \]
\[ = s_{n+k}(\lambda - U). \]
For the last assertion, since the above computation shows $s_{N-k}(\lambda - L) \geq s_N(\lambda - U)$, it suffices to show that $s_N(\lambda - U) \geq 1 - |\lambda|$. But note that $s_N(\lambda - U) = 1/\| (\lambda - U)^{-1} \|$ and
\[ \| (\lambda - U)^{-1} \| = \| (\lambda U^* - 1)^{-1} U^* \| \]
\[ \leq \sum_{l=0}^{\infty} \| U^* \|^l \]
\[ = \frac{1}{1 - |\lambda|}. \]
This completes the proof of the proposition. \hfill \Box

Proof of Lemma 6.1. Let $E_{N1}$ be the matrix that has all entries equal to 0 except the $(N, 1)$-entry, in which the entry is 1. Then $A + E_{N1}$ is a unitary matrix; in fact, this matrix cyclically permutes the standard basis. Point (1) of the lemma follows from applying Proposition 6.2 with $L = A$, $U = A + E_{N1}$ and $k = 1$.

For Point (2), we first note that $\| A \| = 1$ and assume $|\lambda| > 1$. We expand $(\lambda - A)^{-1}$ into a power series to get the norm estimate
\[ \| (\lambda - A)^{-1} \| \leq \sum_{l=0}^{\infty} \frac{\| A \|^l}{|\lambda|^{l+1}} \]
\[ = \frac{1/|\lambda|}{1 - (1/|\lambda|)}. \]
Using the least singular value of $\lambda - A$ is given by $1/\| (\lambda - A)^{-1} \|$ again, we conclude that all the singular values of $\lambda - A$ are at least $|\lambda| - 1$. \hfill \Box

Corollary 6.3. The matrix $A$ satisfies Assumption [7] with any sequence of deterministic nonnegative matrices $\Sigma$ whose eigenvalue distribution converges weakly to a probability measure on $\mathbb{R}$ that is not the Dirac delta measure $\delta_0$.

Proof. Let $R = \{ re^{i\theta} \in \mathbb{C} : 1 - 1/N \leq r \leq 1 + 1/N \}$. We claim that with this Jordan block matrix $A$, if we choose $\alpha = 1$, $\Gamma_N \subset R$. Let $\lambda \in R^c$. If $|\lambda| > 1$, then all the singular values of $A - \lambda$ are at least $|\lambda| - 1 \geq N^{-1}$. If $|\lambda| < 1$, then all except one singular values of $A - \lambda$ are at least $1 - |\lambda| \geq N^{-1}$. Since the eigenvalue distribution of $\Sigma$ converges weakly to a probability measure that is not the Dirac delta measure at $\delta_0$, for all large enough $N$, there must be at least one eigenvalue (which is a singular value as well) away from 0. This shows $\gamma_1(1) + \gamma_2(1) \leq N$ for all $\lambda \in R^c$ and thus $\Gamma_N \subset R$. We can then complete the proof by noting that $R$ has Lebesgue measure $4\pi/N$. \hfill \Box
Corollary 6.4. Given any compact set $K$ that is either $K \subset \mathbb{D}$ or $K \subset (\mathbb{D})^c$, there exists a constant $\kappa > 0$ such that

$$G_{\tilde{\mu}(A-\lambda)}(i\eta) \leq \kappa$$

for all $\eta > 1/N$ and $\lambda \in K$.

Proof. We look at the case that $K \subset \mathbb{D}$. Let $d$ be defined by

$$d = 1 - \sup_{\lambda \in K} |\lambda|.$$

Since $K \subset \mathbb{D}$ is compact, we must have $d > 0$. Now, by Lemma 6.1, $A - \lambda$ has $N - 1$ singular values that are at least $d$. Hence, for any $\lambda \in K$,

$$G_{\tilde{\mu}(A-\lambda)}(i\eta) \leq \frac{1}{N} \frac{\eta}{s_N(A-\lambda)^2 + \eta^2} + \frac{N-1}{N} \frac{\eta}{d^2 + \eta^2} \leq \frac{1}{N} \frac{\eta}{\eta} + \frac{1}{2d} < 1 + \frac{1}{2d}.$$

This shows that, in this $K \subset \mathbb{D}$, $G_{\tilde{\mu}(A-\lambda)}(i\eta) < 1 + 1/2d$ for all $\eta > 1/N$ and $\lambda \in K$.

The case $K \subset (\mathbb{D})^c$ is simpler. Since all the singular values of $A - \lambda$ are at least $d' - 1$, which is positive, where $d' = \inf_{\lambda \in K} |\lambda|$. This shows that, in this case,

$$G_{\tilde{\mu}(A-\lambda)}(i\eta) \leq \frac{1}{2(d' - 1)}$$

for all $\eta > 0$ and $\lambda \in K$. \hfill \Box

Proposition 6.5. Consider the random matrix model $Y$ as in (1.2) and the operator $y \in A$ as in (1.3) with $A$ defined in (6.1). Then the limiting eigenvalue distribution of $Y$ is the Brown measure of $y$.

We run computer simulations where $\Sigma$ is a deterministic diagonal matrix such that half of the diagonal entries are 1 and half of the diagonal entries are 0. Figure 1 shows computer simulations of the eigenvalues of $U\Sigma V + A$ and $U\Sigma V + W$, where $W$ is a Haar-distributed unitary random matrix independent of $U$ and $V$. Both $A$ and $W$ converge in $*$-distribution to a Haar unitary operator. We can see that the eigenvalues distributions of the two simulations are about the same.

Proof. Recall that the matrix $A$ converges to a Haar unitary operator $a$ in $*$-distribution as $N \to \infty$. By Corollary 6.3, Assumption (1.1) holds with this $A$ and $\Sigma$. By Corollary 6.4, Assumption (1) in Theorem 5.1 holds if we take $E = \mathbb{T}$. The conclusion of this proposition then follows from Theorem 5.1. \hfill \Box

Acknowledgments. The authors thank Hari Bercovici for useful conversations. We particularly thank him for showing us the proof of Proposition 6.2. The first author is partially supported by the MoST grant 111-2115-M-001-011-MY3. The second author is supported in part by Collaboration Grants for Mathematicians from Simons Foundation.
Figure 1. Eigenvalue simulations of $U\Sigma V + A$ (left) and $U\Sigma V + W$ (right) where $A$ is the Jordan block matrix, $\Sigma$ is a deterministic diagonal matrix such that $\mu_\Sigma = (1/2)(\delta_0 + \delta_1)$, and $W$ is a Haar-distributed unitary matrix independent of $U$ and $V$.

References

[1] Bai, Z. D. Circular law. *Ann. Probab.* 25, 1 (1997), 494–529.
[2] Bao, Z., Erdős, L., and Schnelli, K. Local law of addition of random matrices on optimal scale. *Comm. Math. Phys.* 349, 3 (2017), 947–990.
[3] Bao, Z., Erdős, L., and Schnelli, K. Local single ring theorem on optimal scale. *Ann. Probab.* 47, 3 (2019), 1270–1334.
[4] Bao, Z., Erdős, L., and Schnelli, K. Local stability of the free additive convolution. *J. Funct. Anal.* 271, 3 (2016), 672–719.
[5] Belinschi, S., Bercovici, H., and Ho, C.-W. On the convergence of Denjoy-Wolff points. arXiv:2203.16728 (2022).
[6] Belinschi, S. T. The Lebesgue decomposition of the free additive convolution of two probability distributions. *Probab. Theory Related Fields* 142, 1-2 (2008), 125–150.
[7] Belinschi, S. T., Šniady, P., and Speicher, R. Eigenvalues of non-Hermitian random matrices and Brown measure of non-normal operators: Hermitian reduction and linearization method. *Linear Algebra and its Applications* 537 (2018), 48 – 83.
[8] Benaych-Georges, F. Local single ring theorem. *Ann. Probab.* 45, 6A (2017), 3850–3885.
[9] Bercovici, H. *Free convolution*, Free probability and operator algebras - Chapter V, *Münster Lectures in Mathematics*. European Mathematical Society (EMS), Zürich, 2016. Lecture notes from the masterclass held in Münster.
[10] Bercovici, H., and Voiculescu, D. Free convolution of measures with unbounded support. *Indiana Univ. Math. J.* 42, 3 (1993), 733–773.
[11] Bercovici, H., and Voiculescu, D. Regularity questions for free convolution. In *Nonselfadjoint operator algebras, operator theory, and related topics*, vol. 104 of *Oper. Theory Adv. Appl.* Birkhäuser, Basel, 1998, pp. 37–47.
[12] Bercovici, H., and Zhong, P. The Brown measure of a sum of two free random variables, one of which is $R$-diagonal. arXiv:2209.12379 (2022).
[13] Bordenave, C., Caputo, P., and Chafaï, D. Spectrum of Markov generators on sparse random graphs. *Comm. Pure Appl. Math.* 67, 4 (2014), 621–669.
DEFORMED SINGLE RING THEOREMS 37

[14] Bourgade, P., Yau, H.-T., and Yin, J. Local circular law for random matrices. Probab. Theory Related Fields 159, 3–4 (2014), 545–595.

[15] Bourgade, P., Yau, H.-T., and Yin, J. The local circular law II: the edge case. Probab. Theory Related Fields 159, 3–4 (2014), 619–660.

[16] Brown, L. G. Lidskiǐ’s theorem in the type II case. In Geometric methods in operator algebras (Kyoto, 1983), vol. 123 of Pitman Res. Notes Math. Ser. Longman Sci. Tech., Harlow, 1986, pp. 1–35.

[17] Collins, B., and Male, C. The strong asymptotic freeness of Haar and deterministic matrices. Ann. Sci. Éc. Norm. Supér. (4) 47, 1 (2014), 147–163.

[18] Erdős, L., Knowles, A., and Yau, H.-T. Averaging fluctuations in resolvents of random band matrices. Ann. Henri Poincaré 14, 8 (2013), 1837–1926.

[19] Erdős, L., Schlein, B., and Yau, H.-T. Local semicircle law and complete delocalization for Wigner random matrices. Comm. Math. Phys. 287, 2 (2009), 641–655.

[20] Erdős, L., Schlein, B., and Yau, H.-T. Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. Ann. Probab. 37, 3 (2009), 815–852.

[21] Fack, T. Sur la notion de valeur caractéristique. J. Operator Theory 7, 2 (1982), 307–333.

[22] Feinberg, J., and Zee, A. Non-Gaussian non-Hermitian random matrix theory: phase transition and addition formalism. Nuclear Phys. B 501, 3 (1997), 643–669.

[23] Ginibre, J. Statistical Ensembles of Complex, Quaternion and Real Matrices. J. Math. Phys. 6 (1965), 440–449.

[24] Gohberg, I. C., and Kreĭn, M. G. Introduction to the theory of linear nonselfadjoint operators. Translations of Mathematical Monographs, Vol. 18. American Mathematical Society, Providence, R.I., 1969. Translated from the Russian by A. Feinstein.

[25] Gualtieri, A., Krishnapur, M., and Zeitouni, O. The single ring theorem. Ann. of Math. (2) 174, 2 (2011), 1189–1217.

[26] Haagerup, U., and Larsen, F. Brown’s spectral distribution measure for R-diagonal elements in finite von Neumann algebras. J. Funct. Anal. 176, 2 (2000), 331 – 367.

[27] Haagerup, U., and Schultz, H. Brown measures of unbounded operators affiliated with a finite von Neumann algebra. Math. Scand. 100, 2 (2007), 209–263.

[28] Haagerup, U., and Schultz, H. Invariant subspaces for operators in a general II1-factor. Publ. Math. Inst. Hautes Études Sci., 109 (2009), 19–111.

[29] Heins, M. H. On the iteration of functions which are analytic and single-valued in a given multiply-connected region. Amer. J. Math. 63 (1941), 461–480.

[30] Hø, C.-W. The Brown measure of the sum of a self-adjoint element and an elliptic element. Electron. J. Probab. 27 (2022), 1 – 32.

[31] Hø, C.-W., and Zhong, P. Brown measures of free circular and multiplicative brownian motions with self-adjoint and unitary initial conditions. arXiv:1908.08150, to appear in J. Eur. Math. Soc. (JEMS) (2020).

[32] Horn, A. Eigenvalues of sums of Hermitian matrices. Pacific J. Math. 12 (1962), 225–241.

[33] Kargin, V. Subordination for the sum of two random matrices. Ann. Probab. 43, 4 (2015), 2119–2150.

[34] Knutson, A., and Tao, T. The honeycomb model of GLn(C) tensor products. I. Proof of the saturation conjecture. J. Amer. Math. Soc. 12, 4 (1999), 1055–1090.

[35] Knutson, A., Tao, T., and Woodward, C. The honeycomb model of GLn(C) tensor products. II. Puzzles determine facets of the Littlewood-Richardson cone. J. Amer. Math. Soc. 17, 1 (2004), 19–48.

[36] Mingo, J. A., and Speicher, R. Free probability and random matrices, vol. 35 of Fields Institute Monographs. Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2017.

[37] Nica, A., and Speicher, R. R-diagonal pairs—a common approach to Haar unitaries and circular elements. In Free probability theory (Waterloo, ON, 1995), vol. 12 of Fields Inst. Commun. Amer. Math. Soc., Providence, RI, 1997, pp. 149–188.

[38] Nica, A., and Speicher, R. Commutators of free random variables. Duke Math. J. 92, 3 (1998), 553–592.

[39] Nica, A., and Speicher, R. Lectures on the Combinatorics of Free Probability. London Mathematical Society Lecture Note Series. Cambridge University Press, 2006.

[40] Rudelson, M., and Vershynin, R. Invertibility of random matrices: unitary and orthogonal perturbations. J. Amer. Math. Soc. 27, 2 (2014), 293–338.
[41] **Shapiro, J. H.** *Composition operators and classical function theory*. Universitext: Tracts in Mathematics. Springer-Verlag, New York, 1993.

[42] **Śniady, P.** Random regularization of Brown spectral measure. *J. Funct. Anal.* 193, 2 (2002), 291–313.

[43] **Tao, T., and Vu, V.** Random matrices: universality of ESDs and the circular law. *Ann. Probab.* 38, 5 (2010), 2023–2065. With an appendix by Manjunath Krishnapur.

[44] **Voiculescu, D.** Symmetries of some reduced free product $C^*$-algebras. In *Operator algebras and their connections with topology and ergodic theory (Bușteni, 1983)*, vol. 1132 of *Lecture Notes in Math.* Springer, Berlin, 1985, pp. 556–588.

[45] **Voiculescu, D.** Limit laws for random matrices and free products. *Invent. Math.* 104, 1 (1991), 201–220.

[46] **Yin, J.** The local circular law III: general case. *Probab. Theory Related Fields* 160, 3-4 (2014), 679–732.

[47] **Zhong, P.** Brown measure of the sum of an elliptic operator and a free random variable in a finite von neumann algebra. *arXiv:2108.09844* (2021).

[48] **Zhong, P.** Brown measure of $R$-diagonal operators, revisited. *arXiv:2204.01896* (2022).

**Ching-Wei Ho:** Institute of Mathematics, Academia Sinica, Taipei 10617, Taiwan  
*Email address:* chwho@gate.sinica.edu.tw

**Ping Zhong:** Department of Mathematics and Statistics, University of Wyoming, Laramie, WY 82071, United States  
*Email address:* pzhong@uwyo.edu