SUBORDINATION FOR THE SUM OF TWO RANDOM MATRICES

By V. Kargin

University of Cambridge

This paper is about the relation of random matrix theory and the subordination phenomenon in complex analysis. We find that the resolvent of the sum of two random matrices is approximately subordinated to the resolvents of the original matrices. We estimate the error terms in this relation and in the subordination relation for the traces of the resolvents. This allows us to prove a local limit law for eigenvalues and a delocalization result for eigenvectors of the sum of two random matrices. In addition, we use subordination to determine the limit of the largest eigenvalue for the rank-one deformations of unitary-invariant random matrices.

1. Introduction.

1.1. *Subordination.* Much of the modern approach to random matrices is based on the analysis of how the resolvent of a matrix A, that is, the function $G_A(z) = (A - zI)^{-1}$, behaves when A is modified by a random perturbation (see [23], e.g.). In this paper, we investigate what happens with the resolvent if an independent rotationally invariant random matrix B is added to A. We find that the resolvent of the sum $A + B$ is (approximately) subordinated to the resolvent of the original matrix A.

The concept of subordination comes from the complex analysis. If $f(z)$ and $g(z)$ are two functions which are analytic in the upper half-plane $\mathbb{C}^+ := \{z : \operatorname{Im} z > 0\}$, then $f(z)$ is *subordinated* to $g(z)$ if there exists an analytic function $\omega(z) : \mathbb{C}^+ \to \mathbb{C}^+$, such that $f(z) = g(\omega(z))$ and $\operatorname{Im} \omega(z) \geq \operatorname{Im} z$ for all $z \in \mathbb{C}^+$. In this definition, $f(z)$ and $g(z)$ can be vector or operator valued functions.

Voiculescu and Biane [12, 44] have discovered that the subordination holds for the resolvent of the sum of two free operators in a von Neumann algebra.

(See also [6] and [19] for different proofs of these results.) This subordination result can be formulated as follows (cf. Theorem 3.1 in [12]). Let $\mathcal{A}$ be a von Neumann operator algebra with the normal faithful trace $\tau : \mathcal{A} \to \mathbb{C}$. If two self-adjoint operators $A, B \in \mathcal{A}$ are free in the sense of Voiculescu (see [34]), then the following identity holds:

$$(1) \qquad \tau(G_{A+B}(z)|A) = G_A(\omega_B(z)),$$

where $\tau(\cdot|A)$ denotes the conditional expectation on the subalgebra generated by operator A, and $\omega_B(z)$ is a function analytic in $\mathbb{C}^+$ and such that $\operatorname{Im}\omega_B(z) \geq \operatorname{Im} z$. In other words, $\tau(G_{A+B}(z)|A)$ is subordinated to $G_A(z)$.

This subordination result is very useful since it implies results about the smoothness of the spectral distribution of the sum $A + B$.

Since large independent random matrices are asymptotically free [41, 43], it is natural to ask whether subordination holds in the context of random matrices. Some results in this direction have been recently obtained in [18, 33] and [15]. In [18] (which builds on an earlier work in [17]), the authors study the matrix $A_N + W_N/\sqrt{N}$, where A_N and W_N are N-by-N Hermitian matrices, A_N is deterministic and W_N is Wigner. It is assumed that the eigenvalue distribution of A_N weakly converges to a measure ν as $N \to \infty$ and that the largest r eigenvalues of A_N ("spikes") are fixed and are outside of the support of ν. The authors are interested in the behavior of r largest eigenvalues of $A_N + W_N/\sqrt{N}$ and this question leads them to the study of the subordination for the trace of the resolvent of $A_N + W_N/\sqrt{N}$.

In further developments, in [33] and in [15], the setup of [18] is generalized for perturbations of the block random matrices and sample covariance matrices, respectively.

We are interested in a somewhat different setup. Let $\widetilde{A}$ and $\widetilde{B}$ be two N-by-N diagonal matrices with real entries. Define the random matrices $A := V\widetilde{A}V^*$ and $B := U\widetilde{B}U^*$ where U and V are two N-by-N random independent uniformly distributed unitary matrices, and define $H := A + B$. Note that the distribution of eigenvalues of H is the same as that of $\widetilde{A} + U\widetilde{B}U^*$, however, it will be convenient to treat A and B symmetrically. The *resolvent* of H is defined as $G_H(z) := (H - zI)^{-1}$ and the *Stieltjes transform* of H is defined as the normalized trace of the resolvent:

$$m_H(z) := \frac{1}{N} \operatorname{Tr}(G_H(z)).$$

The resolvents and the Stieltjes transforms of matrices A and B are defined similarly.

Is it true that $G_H(z)$ is subordinated to $G_A(z)$ and $G_B(z)$ for sufficiently large N?

First, we need to define a candidate subordination function. Let

$$(2) \qquad \omega_B(z) := z - \frac{\mathbb{E}f_B(z)}{\mathbb{E}m_H(z)} \quad \text{and} \quad \omega_A(z) := z - \frac{\mathbb{E}f_A(z)}{\mathbb{E}m_H(z)},$$

where

$$f_B(z) := N^{-1} \operatorname{Tr}(BG_H(z)) \quad \text{and} \quad f_A(z) := N^{-1} \operatorname{Tr}(AG_H(z)).$$

We claim $\omega_A(z)$ and $\omega_B(z)$ are "almost" subordination functions.

THEOREM 1.1.   *Assume that* $\eta := \operatorname{Im} z \in (0,1)$ *and* $|\operatorname{Re} z| \leq K(A,B) :=$ $\max\{\|A\|, \|B\|\}$. *Then for all* $N \gg \eta^{-5}$,

$$\min\{\operatorname{Im}(\omega_A(z)), \operatorname{Im}(\omega_B(z))\} \geq \eta - \frac{c}{N\eta^7},$$

with $c > 0$ *that depends only on* $K(A,B)$.

In other words, for all sufficiently large N, the excess of the imaginary parts of functions $\omega_A(z)$ and $\omega_B(z)$ over $\operatorname{Im} z$ is almost nonnegative. The proof of this theorem is postponed to the next section.

Now we are able to formulate the main result.

[We use the following notation. The average of a random variable X over U is denoted by $\mathbb{E}_U(X) := \mathbb{E}(X|V)$ and the average over V is denoted by $\mathbb{E}_V(X) := \mathbb{E}(X|U)$. The unconditional expectation value is denoted by $\mathbb{E}(X)$. Similar notation will be used for conditional probabilities and variances. For example, $\mathbb{V}\mathrm{ar}_V(X) = \mathbb{E}_V((X - \mathbb{E}_V X)^2)$. The notation $x \ll y$ and $x = O(y)$ mean that there exists a constant $C > 0$ such that $\|x\| \leq Cy$. The constants in these inequalities may depend on $K(A,B) := \max\{\|A\|, \|B\|\}$. The norm $\|\cdot\|$ is the usual uniform norm on matrices.]

THEOREM 1.2.   *Assume that* $\eta := \operatorname{Im} z \in (0,1)$ *and* $|\operatorname{Re} z| \leq K(A,B) :=$ $\max\{\|A\|, \|B\|\}$. *Suppose that* $N \gg \eta^{-7}$. *Then we have:*

(i)

$$\text{(3)} \qquad \mathbb{E}_U G_H(z) - G_A(\omega_B(z)) = O\left(\frac{1}{N\eta^6}\right),$$

(ii)

$$\mathbb{E}m_H(z) - m_A(\omega_B(z)) = O\left(\frac{1}{N^2\eta^6}\right),$$

(iii)

$$\mathbb{P}_U\{|(G_H(z))_{ij} - (G_A(\omega_B(z)))_{ij}| \geq \delta\} \leq \exp(-c\delta^2 N^{3/7}),$$

for all $N \geq N_0$, *where* N_0 *can depend on* $K(A,B)$ *and on* δ.

Estimates similar to estimates in parts (i) and (iii) hold for the conditional expectations with respect to V. The expectation in part (ii) is unconditional since $\mathbb{E} m_H(z) = \mathbb{E}_U m_H(z) = \mathbb{E}_V m_H(z)$ (by an application of Lemma B.2 in the Appendix). An estimate similar to the estimate in part (ii) holds for $\mathbb{E} m_H(z) - m_B(\omega_A(z))$.

It is easy to check (see Lemma B.1) that if the basis is chosen in such a way that A is diagonal then the matrix $\mathbb{E}_U G_H(z)$ is also diagonal. Hence, parts (i) and (iii) of the theorem essentially say that if A is diagonal then $G_H(z)$ is approximately diagonal and its diagonal entries satisfy the formula

$$(G_H(z))_{ii} \approx \frac{1}{A_{ii} - \omega_B(z)}.$$

Part (ii) says that taking the trace of the resolvent makes the error in the approximate formula even smaller, $O(N^{-2}\eta^{-6})$ instead of $O(N^{-1}\eta^{-6})$.

It is interesting to compare this result with the results in [18]. Let $H = A + W/\sqrt{N}$ where W is the N-by-N Wigner matrix with i.i.d. Gaussian entries of variance σ^2. Let $m_H(z) := N^{-1}\mathbb{E}\operatorname{Tr} G_H(z)$. (Note that $\mathbb{E}$ has a slightly different meaning here. It is the expectation taken with respect to the randomness in the Wigner matrix.) It was proved in [18] that for all $z \in \mathbb{C}^+$,

$$(4) \qquad m_H(z) = m_A(z + \sigma^2 m_H(z)) + O(N^{-2}),$$

with the constant in the O-term that depends on z. In addition, if W is a non-Gaussian Wigner matrix, then the same formula is proved in [18] with an additional term on the right of the form $L(z)/N$.

Formula (4) gives a subordination result with the subordination function $\omega_B = z + \sigma^2 m_H(z)$. Our Theorem 1.2 holds for a more general matrix model and gives estimates for resolvents as well as for the Stieltjes transforms. These estimates give the explicit dependence of the error term on z unlike formula (4). These advantages are crucial for the applications to the local distribution of eigenvalues and delocalization of eigenvectors.

Now we turn to these applications.

1.2. *Delocalization.* *Delocalization* of eigenvectors generally refers to the situation when all individual coordinates of a normalized eigenvector v_α in a specific basis are not greater than $N^{-\kappa/2}$ with a high probability. Because of normalization, the eigenvector is forced to be spread over at least N^κ coordinates and it is customary to say that the delocalization length of eigenvectors is at least N^κ.

The question about delocalization of eigenvectors frequently occurs in physics. For example, a famous open problem is to show that for $d \geq 3$ the eigenvectors of random Schrödinger operators on $\mathbb{Z}^d$ are completely delocalized for small disorder. Recently, there was some progress on delocalization

of eigenvectors in simpler models, for instance, in the case of random Wigner matrices and in the case of random band matrices. In the former case, the complete delocalization has been established recently (see [20] for a review) and the method is similar to the method that is used in this paper. In the case of band matrices, it is expected that complete delocalization holds for matrices with the band width W greater than $\sqrt{N}$. What was actually shown in this case is that the delocalization length is greater than $W^{1+d/6}$ ([21], and an improvement was recently achieved in [22]). The method is based on quantum diffusion and different from the method that is used in this paper.

In our model, we say that the eigenvectors $\nu_a^{(N)}$ of a sequence of matrices $H_N = A_N + U B_N U^*$ are *delocalized* at length N^κ in the interval I, if there exists $\delta > 0$ such that

$$\mathbb{P}\{|v_a^{(N)}(i)|^2 > N^{-\kappa} \log N\} \le \exp(-N^{-\delta}),$$

for all sufficiently large N, all $i \in \{1, \ldots, N\}$ and all $\nu_a^{(N)}$ such that the corresponding eigenvalues are in the interval I.

Let μ_{A_N} be the empirical measure of eigenvalues of A_N, that is, $\mu_{A_N} := N^{-1} \sum_{k=1}^N \delta_{\lambda_k}$, where λ_k are eigenvalues of μ_{A_N}. Define μ_{B_N} similarly. We are going to prove that the eigenvalues of H_N are delocalized at a certain scale if μ_{A_N} and μ_{B_N} are close enough to a couple of measures that satisfy a regularity conditions.

As a measure of closeness between probability measures μ and ν, we use the *Lévy distance*

$$d_L(\mu, \nu) = \sup_x \inf\{s \ge 0 : F_\nu(x - s) - s \le F_\mu(x) \le F_\nu(x + s) + s\},$$

where $F_\mu(t)$ and $F_\nu(t)$ are the cumulative distribution functions of μ and ν. Note that $\mu^{(N)} \to \mu$ in distribution if and only if $d_L(\mu^{(N)}, \mu) \to 0$. (See Theorem III.1.2 on page 314 and Exercise III.1.4 on page 316 in [40].)

THEOREM 1.3. *Assume that* (i) *a pair of probability measures* (μ_α, μ_β) *is smooth in a closed interval* I, *and* (ii) *for a sequence of* A_N *and* B_N, $\max\{\|A_N\|, \|B_N\|\} \le K$ *for all* N. *Then there exists* $s > 0$ *such that if*

$$(5) \qquad \max\{d_L(\mu_{A_N}, \mu_\alpha), d_L(\mu_{B_N}, \mu_\beta)\} \le s$$

for N *large enough, then eigenvectors of* $H_N = A_N + U B_N U^*$ *are delocalized at scale* $N^{1/7}$ *in the interval* I.

Note that we do not require the measures μ_{A_N} and μ_{B_N} to converge to μ_α and μ_β. It is enough that they are sufficiently close to μ_α and μ_β for all large N. The reason for this is that this weaker condition is enough to ensure that the subordination functions of the pair (μ_{A_N}, μ_{B_N}) are separated

from zero in the region $\eta \geq N^{-1/7}$. On the other hand, if μ_{A_N} and μ_{B_N} do converge to μ_α and μ_β, then condition (5) is automatically satisfied. This is perhaps the most important case in applications.

We still need to explain what is meant by the smoothness of a pair (μ_α, μ_β). Let μ_α and μ_β be two probability measures with bounded support, and let $m_\alpha(z) := \int (t-z)^{-1} \mu_\alpha(dt)$ and $m_\beta(z) := \int (t-z)^{-1} \mu_\beta(dt)$. The system of equations

$$
\begin{aligned}
m(z) &= m_\alpha(\omega_\beta(z)), \\
(6) \qquad m(z) &= m_\beta(\omega_\alpha(z)) \quad \text{and} \\
z - \frac{1}{m(z)} &= \omega_\alpha(z) + \omega_\beta(z)
\end{aligned}
$$

has a unique solution $(m(z), \omega_\alpha(z), \omega_\beta(z))$ in the class of functions that are analytic in $\mathbb{C}^+ = \{z : \operatorname{Im} z > 0\}$ and that have the following expansions at infinity:

$$
\begin{aligned}
(7) \qquad m(z) &= -z^{-1} + O(z^{-2}), \\
\omega_\alpha(z) &= z + O(1) \quad \text{and} \quad \omega_\beta(z) = z + O(1).
\end{aligned}
$$

The function $m(z)$, which we denote as $m_{\mu_\alpha \boxplus \mu_\beta}(z)$, is the Stieltjes transform of a probability measure which is called the *free convolution* of measures μ_α and μ_β and denoted $\mu_\alpha \boxplus \mu_\beta$. The functions $\omega_\alpha(z)$ and $\omega_\beta(z)$ are the subordination functions for the free convolution.

By Theorem 3.3 in [5], the limits $\omega_j(x) = \lim_{\eta \downarrow 0} \operatorname{Im} \omega_j(x + i\eta)$ exist for $j = \alpha, \beta$, and we make the following definition. A pair of probability measures on the real line (μ_α, μ_β) is said to be *smooth* at $x \in \mathbb{R}$ if the following two conditions hold:

(A) $\operatorname{Im} \omega_j(x) > 0$ for $j = \alpha, \beta$, and
(B)

$$
(8) \quad k_\mu(x) := \frac{1}{m'_{\mu_\alpha}(\omega_\beta(x))} + \frac{1}{m'_{\mu_\beta}(\omega_\alpha(x))} - (\omega_\alpha(x) + \omega_\beta(x) - x)^2 \neq 0.
$$

We say that the pair (μ_α, μ_β) is *smooth in interval* $I \subset \mathbb{R}$ if $\omega_\alpha(z)$ and $\omega_\beta(z)$ are continuous in a rectangle $\{z = x + i\eta | x \in I, 0 \leq \eta \leq \varepsilon\}$ where ε is a positive constant, and if the pair (μ_α, μ_β) is smooth at every point of I.

The proof of Theorem 1.3 is based on part (iii) of Theorem 1.2 which imply that $\operatorname{Im}(G_H)_{kk}(\lambda_a + i\eta) \leq \operatorname{Im}(G_A(\omega_B(\lambda_a + i\eta)))_{kk} + \delta$. Then the assumption of smoothness leads (after some work) to the conclusion that the quantity on the right is bounded for all k and all $N \gg \eta^{-1/7}$ with high probability. Therefore, the components of the eigenvector corresponding to λ_a can be estimated by using the bound on the resolvent

$$
(9) \qquad |v_a(k)|^2 \leq \eta \operatorname{Im} G_{kk}(\lambda_a + i\eta) \leq C\eta \leq CN^{-1/7} \log N.
$$

To get the last inequality, η is chosen as $N^{-1/7}\log N$ so that Theorem 1.2 is applicable. The details are postponed to Section 3.

Let us add some comments about the assumption of smoothness. Condition (B) is technical and holds for a generic point $x \in \mathbb{R}$. It ensures that the solution of the system (6) at x is stable with respect to a small perturbation in the system. Condition (A) is essential and closely related to regularity properties of the measure $\mu_\alpha \boxplus \mu_\beta$ at x. Here are some cases when it holds:

(i) If $\mu_\alpha = \mu_\beta = \mu$, and $\mu \boxplus \mu$ is absolutely continuous with positive density at x, then (A) is satisfied at the point x. In particular, if μ is an arbitrary measure that does not have an atom with the mass greater than $1/2$, then condition (A) is satisfied at every point inside the support of $\mu \boxplus \mu$.

(ii) If one of the probability measures has the semicircle distribution with the density $f_{\mathrm{sc}}(x) = \frac{1}{2\pi}\sqrt{(4-x^2)_+}$, the density of $\mu_{\mathrm{sc}} \boxplus \mu$ is positive at x, and $|m_{\mu_{\mathrm{sc}}\boxplus\mu}(x)| \neq 1$, then condition (A) is satisfied. (For a more detailed discussion of these examples, the reader can see Propositions 1.4 and 1.5 in [29].)

In fact, smoothness is likely to be a typical situation for pairs (μ_α, μ_β). This is because the free convolution operation has very strong smoothing properties. Even if we start with two discrete measures μ_α and μ_β, the free convolution $\mu_\alpha \boxplus \mu_\beta$ is absolutely continuous provided that the masses of an atom of μ_α and an atom of μ_β do not add up to more than 1.

How does one find pairs which are not smooth? A pair of measures (μ_α, μ_β) is not smooth at a point where the density of $\mu_\alpha \boxplus \mu_\beta$ vanishes, in particular at the boundary of the support of $\mu_\alpha \boxplus \mu_\beta$. One other example occurs when both μ_α and μ_β have an atom, and the sum of the atoms' masses is greater than 1. In this case, the free convolution $\mu_\alpha \boxplus \mu_\beta$ also has an atom and, therefore, the pair (μ_α, μ_β) is not smooth at the location of this atom.

The result in Theorem 1.3 is certainly not optimal. The true localization length is probably of order N under assumptions of the theorem, that is, eigenvectors are likely to be completely delocalized.

1.3. *Local limit for eigenvalue distribution.* Another consequence of Theorem 1.2 is the convergence of the eigenvalue counting measure on the local scale.

Let $\mathcal{N}_{\eta^*}(x)$ be the number of eigenvalues of H_N in the interval $I^* = [x - \eta^*, x + \eta^*]$. What can be said about $\mathcal{N}_{\eta^*}(x)/(2\eta^* N)$ when $N \to \infty$? If η^* is fixed, then it is known [43] and [41] that the limit approaches $\mu_\alpha \boxplus \mu_\beta(I^*)/(2\eta^*)$. Local limit theorems address the question of what happens if η^* is not fixed but approaches 0 when $N \to \infty$.

THEOREM 1.4.   *Assume that* (i) $\max\{d_L(\mu_{A_N}, \mu_\alpha), d_L(\mu_{B_N}, \mu_\beta)\} \to 0$, (ii) *the pair of probability measures* (μ_α, μ_β) *is smooth on interval* I *and*

(iii) $\max\{\|A_N\|, \|B_N\|\} \leq K$ *for all* N. *Let* $\rho_{\mu_\alpha \boxplus \mu_\beta}$ *denote the density of* $\mu_\alpha \boxplus \mu_\beta$, *and let* $\eta^* = cN^{-1/7} \log N$. *Then, for every* $x \in I$,

$$\frac{\mathcal{N}_{\eta^*}(x)}{2\eta^* N} \to \rho_{\mu_\alpha \boxplus \mu_\beta}(x)$$

in probability.

The theorem improves the local limit law in [28], where it was found that it holds for the window size $\eta^* \sim (\log N)^{-1/2}$. The optimal result is probably $\eta^* \sim N^{-1+\varepsilon}$ with arbitrarily small positive ε, similar to the case of classical Gaussian ensembles and the case of Wigner/sample covariance matrices. The proof of Theorem 1.4 will be given in Section 4.

1.4. *Largest eigenvalues of finite rank deformations of unitarily-invariant matrices.* The largest eigenvalues of finite-rank deformations of Wigner matrices have been recently received much attention and studied in [15, 17, 18, 24, 30–33, 36, 38, 39] and [37]. This study is closely connected to the study of spiked population models in [2–4, 14] and [13].

The idea that the subordination identities are useful in the context of matrix deformations has first appeared in the work of Capitaine, Donati-Martin, Feral and Fevrier (see [18] and [15]). We will use this idea to give a different proof for a result of Benaych-Georges and Nadakuditi in [10]. They considered the largest eigenvalue of $H_N = A_N + U_N B_N U_N^*$, where A_N is a finite rank Hermitian matrix, and found a formula for the limits of the largest eigenvalues. (See also [8, 9] and [11] for further developments.)

Theorem 1.2 allows us to obtain a different proof of Benaych-Georges and Nadakuditi's result. While their method is based on analysing the zeros of determinants of certain matrix-valued functions, our method uses the singularities of the resolvent traces. In particular, we use the description that Theorem 1.2 gives for the resolvent behavior in the upper half-plane.

We consider the simplest case when matrix A_N has rank one. The ideas of the proof can be applied similarly in the case when A_N is a finite-rank matrix with the rank fixed and N approaching infinity.

Let $\rho_\mu(\theta)$ be the largest real solution of the equation $\theta m_\mu(x) + 1 = 0$, and let $\lambda_1(X)$ denote the largest eigenvalue of Hermitian matrix X.

THEOREM 1.5. *Let* $H_N = A_N + U_N B_N U_N^*$ *where* A_N *is a rank-one Hermitian matrix with the eigenvalue* $\theta_0 > 0$, *and* B_N *is a Hermitian matrix with the empirical eigenvalue distribution* μ_{B_N}. *Let* $\lambda_1(B_N) \to L$ *in probability. Assume that matrices* B_N *are uniformly bounded almost surely and that* μ_{B_N} *weakly converges to a probability measure* μ. *Then*

$$\lambda_1^{(H_N)} \to \begin{cases} \rho_\mu(\theta_0), & \text{if } \rho_\mu(\theta_0) > L, \\ L, & \text{otherwise,} \end{cases}$$

where convergence is in probability.

The proof of Theorem 1.5 is based on the subordination-like formula, which we will prove in Proposition 5.1:

$$(10) \quad \mathbb{E} m_{H_N}(z) = m_{B_N}(z) + \frac{1}{N} \frac{m'_{B_N}(z)}{m_{B_N}(z)} \left(\frac{1}{\theta m_{B_N}(z) + 1} - 1 \right) + O_\eta \left(\frac{1}{N^2} \right),$$

where $O_\eta(N^{-2})$ denotes a function $f(z)$ such that $N^2 |f(z)| \leq C (\operatorname{Im} z)^{-k}$ for some $k > 0$ and $C > 0$. This formula explicitly shows the correction term to the Stieltjes transform $m_{B_N}(z)$ that results from adding matrix A_N. In particular, this correction term has an additional pole to the right of L if and only if a zero of $\theta m_{B_N}(z) + 1$ is located to the right of L. This implies (after some additional work) that $\rho_\mu(\theta_0)$ is the only possible limit point for the largest eigenvalue of H_N.

After the preprint of this paper has appeared, the method of subordination functions was used in [7] to generalize the results [10]. The main innovation in [7] is that the matrix A_N is no longer required to be finite rank. It is only required that it has sufficiently large fixed eigenvalues ("spikes").

Two standard examples in the deformation theory are Gaussian–Hermitian matrices and Gaussian–Wishart matrices as B_N. In these examples, Theorem 1.5 gives the results in agreement with available in the literature [3] and [17], with $\rho_H(\theta) = \theta + \sigma^2/\theta$ and $\rho_W(\theta) = \theta + \lambda\theta/(\theta - 1)$.

Another example, which seems to be new, is provided by random projection matrices.

EXAMPLE.   Consider matrices $B_N = U_N P_N U_N^*$ where P_N is a projection matrix of rank p_N. If $p_N/N \to p > 0$ as $N \to \infty$, then the empirical eigenvalue distribution of B_N converges to the Bernoulli distribution $\mu_b = p\delta_1 + q\delta_0$, where $q = 1 - p$. One computes that

$$\rho(\theta) = \frac{\theta}{2} + \frac{1}{2}(1 + \sqrt{(1+\theta)^2 - 4q\theta}),$$

and this is the limit of the largest eigenvalue for the matrices $A_N + B_N$ when $N \to \infty$. This formula for the limit of the largest eigenvalue is valid for all $\theta_0 > 0$.

In the context of this example, an interesting phenomenon is uncovered by numerical evidence, which is not explained by Theorem 1.5. Namely, adding a rank one projection A_N with eigenvalue θ results in a creation of *two* "new" eigenvalues. (See Figure 1 for a numerical example with the size of the matrices fixed at $N = 100$.) One new eigenvalue is given by $\rho(\theta)$, and another one by the other solution of the equation $\theta m(x) + 1 = 0$:

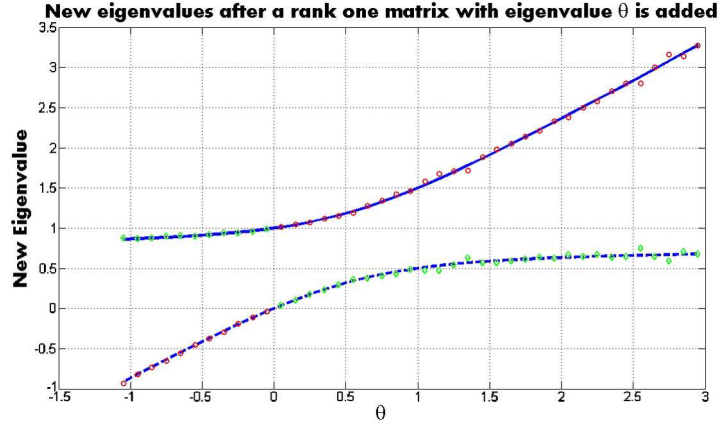$$\overline{\rho}(\theta) = \frac{\theta}{2} + \frac{1}{2}(1 - \sqrt{(1+\theta)^2 - 4q\theta}).$$

FIG. 1. *New eigenvalues of the rank one perturbation of the projection matrix model with $p = 1/4$. Circles are eigenvalues outside of $[0,1]$, diamonds are eigenvalues inside $[0,1]$. The solid line is $\rho(\theta)$, the dashed line is its conjugate.*

1.5. *Brief overview.* In this paper, we consider the resolvent of the matrix $H = A + B$, where $A := V\widetilde{A}V^*$ and $B := U\widetilde{B}U^*$, $\widetilde{A}$ and $\widetilde{B}$ are two N-by-N Hermitian diagonal matrices, and U and V are two N-by-N random independent uniformly distributed unitary matrices.

We have showed that there exist two functions $\omega_A(z)$ and $\omega_B(z)$ that depend only on the sets of eigenvalues of A and B and have the following properties:

(i) $\omega_A(z)$ and $\omega_B(z)$ are analytic in $\mathbb{C}^+$;
(ii) if $\operatorname{Im} z \gg N^{-1/5}$, then

$$\min\{\operatorname{Im}\omega_A(z), \operatorname{Im}\omega_B(z)\} \geq \operatorname{Im} z - \frac{c}{N(\operatorname{Im} z)^7}.$$

Moreover, if $\mu_A = \mu_B$, then $\min\{\operatorname{Im}\omega_A(z), \operatorname{Im}\omega_B(z)\} \geq \operatorname{Im} z$ for all $z \in \mathbb{C}^+$;
(iii) if $\operatorname{Im} z \gg N^{-1/7}$, then

$$\mathbb{E}_U G_H(z) - G_A(\omega_B(z)) = O\left(\frac{1}{N\eta^6}\right)$$

and

$$\mathbb{E} m_H(z) - m_A(\omega_B(z)) = O\left(\frac{1}{N^2\eta^6}\right),$$

and similar estimates hold for $\mathbb{E}_V G_H(z) - G_B(\omega_A(z))$ and $\mathbb{E} m_H(z) - m_B(\omega_A(z))$.

This can be thought of as a subordination property for the resolvent of the sum $A + B$ with respect to resolvents of A and B.

We have used the subordination property to show that the localization length of eigenvectors is greater than N^κ, where $\kappa = 1/7$. The probable actual localization length is $O(N)$.

Next, we have showed that a local limit law holds for the empirical eigenvalue measure μ_{H_N} with the window length $N^{-1/7}$. This result improves over the result in [28]. However, it is still far from the probable optimal result with the window length $N^{-1+\varepsilon}$.

Finally, by using our results about subordination we studied the rank-one deformations of unitarily-invariant random matrices, and derived explicit formulas for the limit of their largest eigenvalues.

The rest of the paper is organized as follows. Section 2 is devoted to the proof of Theorems 1.1 and 1.2 regarding the subordination. Section 3 is about delocalization of eigenvectors (Theorem 1.3). Section 4 proves Theorem 1.4 about the local law for eigenvalues. Section 5 proves Theorem 1.5 about rank-one deformations of unitarily-invariant ensembles. And three appendices contain various auxiliar results.

2. Approximate subordination. Before we start the proof of Theorems 1.1 and 1.2, note that the definitions imply the following useful identity:

$$(11) \qquad \omega_A(z) + \omega_B(z) = z - \frac{1}{\mathbb{E}m_H(z)}.$$

Indeed,

$$\omega_A(z) + \omega_B(z) = 2z - \frac{\mathbb{E}[f_A(z) + f_B(z)]}{\mathbb{E}m_H(z)}$$

and

$$\begin{aligned} f_A(z) + f_B(z) &= N^{-1}\operatorname{Tr}((A+B)(A+B-zI)^{-1}) \\ &= 1 + zN^{-1}\operatorname{Tr}((A+B-zI)^{-1}) \\ &= 1 + zm_H(z), \end{aligned}$$

which implies (11).

Now we start proving Theorem 1.1. First, write

$$(12) \qquad \mathbb{E}_U G_H(z) = G_A(\omega_B(z)) + R_A(z).$$

The error term in subordination formula (12) can be written as follows:

$$(13) \qquad R_A(z) := \frac{1}{\mathbb{E}m_H}(A - zI)G_A(\omega_B(z))\mathbb{E}_U\Delta_A,$$

where

$$\Delta_A := -(m_H - \mathbb{E}m_H)G_H - G_A(f_B - \mathbb{E}f_B)G_H.$$

In order to derive formulae (12) and (13), one starts by calculating dG_t/dt where $G_t = (A + e^{iXt}Be^{-iXt})^{-1}$ and X is an Hermitian matrix. Since B has a rotationally invariant distribution, hence $\mathbb{E}_U(dG_t/dt) = 0$, and one can find by using different generator matrices X that this implies that $\mathbb{E}_U(G_H \otimes BG_H) = \mathbb{E}_U(G_H B \otimes G_H)$. After taking the trace over the first component of the tensor product, one gets $\mathbb{E}_U(m_H BG_H) = \mathbb{E}_U(f_B G_H)$. This can be rewritten as $\mathbb{E}_U(m_H G_H) = G_A \mathbb{E}_U(m_H I - f_B G_H)$. Next, one writes $\mathbb{E}_U(m_H G_H) = \mathbb{E}(m_H)\mathbb{E}_U(G_H) + e_1$ and $\mathbb{E}_U(f_B G_H) = \mathbb{E}(f_B)\mathbb{E}_U(G_H) + e_2$, where e_1 and e_2 are error terms. After substituting these expressions, one can manipulate the previous identity so that $\mathbb{E}_U(G_H)$ is on the left-hand side and everything else is on the right-hand side. The resulting expression is equivalent to (12) with the error term given by (13). See Appendix A for a more complete derivation, and [35] or proof of Theorem 7 in [28] for details.

We can also rewrite formula (12) as follows:

$$\mathbb{E}_U G_H = G_A(\omega_B(z))\left(I + \frac{1}{\mathbb{E}m_H}(A - zI)\mathbb{E}_U\Delta_A \right).$$

Hence,

$$(\mathbb{E}_U G_H)^{-1} = \left(I + \frac{1}{\mathbb{E}m_H}(A - zI)\mathbb{E}_U\Delta_A \right)^{-1}(A - \omega_B(z)I)$$

and

$$\begin{aligned}
(14) \quad \omega_B(z)I &= -(\mathbb{E}_U G_H)^{-1} + A \\
&\quad + \left[\left(I + \frac{1}{\mathbb{E}m_H}(A - zI)\mathbb{E}_U\Delta_A \right)^{-1} - I \right](A - \omega_B(z)I).
\end{aligned}$$

Let us consider the first two terms in this expression. Later, we are going to show that the third term is small.

Define

$$\Omega_B(z, A) := -(\mathbb{E}_U G_H(z))^{-1} + A.$$

The matrix function $\Omega_B(z, A)$ has a property which is similar to the subordination property.

LEMMA 2.1.   *Let $\lambda(z)$ be an eigenvalue of $\Omega_B(z, A)$. Then $\operatorname{Im}\lambda(z) \geq \operatorname{Im} z$.*

PROOF.   Let $z = x + i\eta$ and $\eta > 0$. Then every matrix $(H - x - i\eta)^{-1}$ is normal and its eigenvalues are on the border of a disc D_η with the center at $i/(2\eta)$ and the radius equal to $1/(2\eta)$. Hence, by Lemma B.3 in Appendix B, the eigenvalues of $\mathbb{E}_U(H - x - i\eta)^{-1}$ belong to the disc D_η. It follows that eigenvalues of $-[\mathbb{E}_U(H - x - i\eta)^{-1}]^{-1}$ are in $\mathbb{H}_\eta = \{w : \operatorname{Im} w \geq \eta\}$.

If we take the basis in which A is diagonal, then $(\mathbb{E}_U G_H(z))^{-1} = [\mathbb{E}_U(H - x - i\eta)^{-1}]^{-1}$ is diagonal by Lemma B.1 in Appendix B. Since A is Hermitian, therefore its eigenvalues are real. Hence, the imaginary parts of eigenvalues of $\Omega_B(z, A)$ coincide with imaginary parts of eigenvalues of $-[\mathbb{E}_U(H - x - i\eta)^{-1}]^{-1}$, and we arrive at the claim of the lemma. $\square$

Now we are going to estimate the size of the third term in the right-hand side in (14). First, we estimate the size of $\mathbb{E}_U \Delta_A$. We use concentration inequalities.

LEMMA 2.2.   *Assume that $\eta := \operatorname{Im} z \in (0, 1)$ and $|\operatorname{Re} z| \leq K(A, B)$. Then*

$$\mathbb{E}_U \Delta_A(z) = O\left(\frac{1}{\eta^4 N}\right).$$

PROOF.   Since $\|G_H\| \leq 1/\eta$, hence by using Lemma C.1 in Appendix C, we obtain

$$\mathbb{P}\{\|(m_H(z) - \mathbb{E}m_H(z))G_H\| \geq \delta/\eta\} \leq \exp\left[-c\frac{\delta^2\eta^4}{\|B\|^2}N^2\right],$$

$$\mathbb{P}\{\|G_A(f_B(z) - \mathbb{E}f_B(z))G_H\| \geq \delta/\eta^2\} \leq \exp\left[-c\frac{\delta^2\eta^4}{\|B\|^2}N^2\right].$$

Set $\varepsilon = \delta/\eta$ and $\varepsilon = \delta/\eta^2$ in the first and the second inequalities, respectively, and use the triangle inequality for norms in order to obtain that

$$\mathbb{P}\{\|\Delta_A(z)\| \geq \varepsilon\} \leq \exp\left[-\frac{c\varepsilon^2 N^2}{\|B\|^2}\min\{\eta^6, \eta^8\}\right]$$

$$\leq \exp[-c\varepsilon^2\eta^8 N^2].$$

Next, note that $\|\mathbb{E}_U \Delta_A\| \leq \mathbb{E}_U\|\Delta_A\|$ by the convexity of norm, and $\mathbb{E}_U\|\Delta_A\|$ can be estimated by using the equality $\mathbb{E}X = \int_0^\infty (1 - \mathcal{F}_X(t))\,dt$, valid for every positive random variable X and its cumulative distribution function $\mathcal{F}_X(t)$. In our case, we obtain

$$\mathbb{E}_U\|\Delta_A\| \leq \int_0^\infty \exp[-ct^2\eta^8 N^2]\,dt = \frac{c'}{N\eta^4}.   \square$$

Next, Lemma C.2 in Appendix C says that $(\mathbb{E}m_H(z))^{-1} \leq c/\eta$. Hence, Lemma 2.2 implies that

$$(15) \qquad \left\|\frac{1}{\mathbb{E}m_H}(A - z)\mathbb{E}\Delta_A\right\| \leq \frac{c}{\eta^5 N},$$

where $c > 0$ depends only on K and R.

It is easy to prove that if $\|X\| \leq \varepsilon < 1/2$, then $\|(I+X)^{-1} - I\| \leq 2\varepsilon$. In particular, for all $N \gg \eta^{-5}$, we have

$$\left\| \left(I + \frac{1}{\mathbb{E}m_H}(A-z)\mathbb{E}\Delta_A \right)^{-1} - I \right\| \leq \frac{c}{\eta^5 N}.$$

Next, note that by definition $\omega_B(z) = z - \mathbb{E}f_B(z)/\mathbb{E}m_H(z)$. From Lemma C.2, $|(\mathbb{E}m_H(z))^{-1}| < c/\eta$. In addition,

$$|\mathbb{E}f_B(z)| = \left| \mathbb{E}\frac{1}{N}\operatorname{Tr}\left(B\frac{1}{H-z} \right) \right| \leq \|B\|\mathbb{E}\left(\left\| \frac{1}{H-z} \right\| \right) \leq c\frac{1}{\eta}.$$

Hence, $|\omega_B(z) - z| \leq c/\eta^2$. It follows that

$$(16) \qquad \left\| \left(\left(I + \frac{1}{\mathbb{E}m_H}(A-z)\mathbb{E}\Delta_A \right)^{-1} - I \right)(A - \omega_B(z)) \right\| \leq \frac{c}{N\eta^7}.$$

LEMMA 2.3.   *Let Ω be a diagonal matrix and R be an arbitrary matrix. Then for every eigenvalue $\widehat{\lambda}_i$ of $\Omega + R$, there exists an eigenvalue λ_i of Ω such that $|\widehat{\lambda}_i - \lambda_i| \leq \|R\|$.*

(See Theorem 6.3.2 on page 365 in [26].)

Formulae (14) and (16), and Lemmas 2.1 and 2.3 imply that

$$\operatorname{Im}(\omega_B(z)) \geq \operatorname{Im} z - \frac{c}{N(\operatorname{Im} z)^7}.$$

This completes the proof of Theorem 1.1.

PROOF OF THEOREM 1.2.   Since $N \gg 1/\eta^7$, Theorem 1.1 implies that $\operatorname{Im}(\omega_B(z)) \geq \eta/2$. Hence, $\|G_A(\omega_B(z))\| \leq c/\eta$. Since

$$R_A(z) = G_A(\omega_B(z))\frac{1}{\mathbb{E}m_H}(A-z)\mathbb{E}_U\Delta_A,$$

we can use (15) in order to obtain

$$\|R_A(z)\| \leq \|G_A(\omega_B(z))\| \left\| \frac{1}{\mathbb{E}m_H}(A-z)\mathbb{E}_U\Delta_A \right\| \leq \frac{c}{N\eta^6},$$

which yields the first point of Theorem 1.2.

In order to estimate the error term in the second part of the theorem, we note that by definition of R_A it is enough to show that

$$\mathbb{E}\frac{1}{N}\operatorname{Tr}[(A-z)G_A(\omega_B(z))\Delta_A] = O\left(\frac{1}{\eta^6 N^2} \right).$$

By using the definition of Δ_A, we can write the modulus of the expression on the left-hand side as follows:

$$|\mathbb{E}(m_H - \mathbb{E}m_H)(\varphi - \mathbb{E}\varphi) + \mathbb{E}(f_B - \mathbb{E}f_B)(\psi - \mathbb{E}\psi)|,$$

where

$$\varphi := \frac{1}{N} \operatorname{Tr}[(A - z)G_A(\omega_B(z))G_H(z)]$$

and

$$\psi := \frac{1}{N} \operatorname{Tr}[(A - z)G_A(\omega_B(z))G_A(z)G_H(z)].$$

By the Cauchy–Schwarz inequality, we estimate this from above by

(17) $$\sqrt{\operatorname{Var}(m_H)\operatorname{Var}(\varphi)} + \sqrt{\operatorname{Var}(f_B)\operatorname{Var}(\psi)}.$$

By applying Lemma C.1 from Appendix C and the estimate $\|G_A(\omega_B)\| \leq c/\eta$ in order to bound the variances, we find that for sufficiently large N, $\operatorname{Var}(m_H) = O(\eta^{-4}N^{-2})$, $\operatorname{Var}(\varphi) = O(\eta^{-6}N^{-2})$, $\operatorname{Var}(f_B) = O(\eta^{-4}N^{-2})$ and $\operatorname{Var}(\psi) = O(\eta^{-8}N^{-2})$. Hence, the expression in (17) is smaller than $c/(\eta^6 N^2)$, provided that z is in the region where $\omega_B(z)$ increases the imaginary part. This completes the proof of the second part of the theorem.

The third part of Theorem 1.2 immediately follows from the first part and Lemma C.1 in Appendix C if we take $\eta \gg N^{-1/7}$.

Indeed, if $\operatorname{Im} z = \eta \gg N^{-1/7} \gg N^{-1/6}$, then the first part of Theorem 1.2 implies that $|(\mathbb{E}_U G_H(z))_{ij} - (G_A(\omega_B(z)))_{ij}| \leq \delta/2$ for all sufficiently large N. For these N, we have

$$\mathbb{P}_U\{|(G_H(z))_{ij} - (G_A(\omega_B(z)))_{ij}| \geq \delta\}$$
$$\leq \mathbb{P}_U\{|(G_H(z))_{ij} - (\mathbb{E}_U G_H(z))_{ij}| \geq \delta/2\}$$
$$\leq \exp\left(-\frac{c\delta^2\eta^4}{\|B\|^2}N\right) \leq \exp(-c\delta^2 N^{3/7}). \qquad \square$$

3. Delocalization. The essential part of the proof is to show that ω_{A_N} and ω_{B_N} are close to ω_α and ω_β, respectively. Namely, let

(18) $$r(z) := \max\{|r_A(z)|, |r_B(z)|\}$$

and

(19) $$s(A, B) := \max\{d_L(\mu_A, \mu_\alpha), d_L(\mu_B, \mu_\beta)\}.$$

PROPOSITION 3.1. *Assume that a pair of probability measures (μ_α, μ_β) is smooth in a closed interval I. Then for some positive $\overline{r}$, $\overline{s}$ and $\overline{\eta}$, if $r(z) \leq \overline{r}$, $s(A, B) \leq \overline{s}$, $\operatorname{Re} z \in I$ and $\operatorname{Im} z \in (0, \overline{\eta}]$, then*

$$\max(|\omega_\alpha(z) - \omega_A(z)|, |\omega_\beta(z) - \omega_B(z)|) = O(r + s),$$

where the constant in the O-term may depend on the pair (μ_α, μ_β) and on $\max\{\|A\|, \|B\|\}$.

Let us postpone the proof and show how this result implies Theorem 1.3.

PROOF OF THEOREM 1.3.   Let N be the size of matrices A and B and assume that N is sufficiently large so that $\max\{d_L(\mu_A, \mu_\alpha), d_L(\mu_B, \mu_\beta)\} \leq s < \overline{s}$. By definition, $r_A(z) = \mathbb{E}m_H(z) - m_A(\omega_B(z))$, hence the second part of Theorem 1.2 says that if $N \gg \eta^{-7}$, then $|r_A(z)| = O(\frac{1}{N^2\eta^6})$. A similar bound holds for $|r_B(z)|$. Hence, we can take $r = O(\frac{1}{N^2\eta^6})$ in Proposition 3.1 and conclude that

$$(20) \qquad \omega_\beta(x + i\eta) - \omega_B(x + i\eta) = O\left(\frac{1}{N^2\eta^6} + s\right).$$

Hence, if s is sufficiently small, then $\operatorname{Im}\omega_B(z) \geq c > 0$ for all z with $\operatorname{Re}z \in I$ and $cN^{-2/6} \leq \operatorname{Im}z \leq \overline{\eta}$. It follows that $[G_A(\omega_B(z))]_{kk}$ is bounded, say, $|[G_A(\omega_B(z))]_{kk}| < C$. By using the third part of Theorem 1.2, we find that

$$\mathbb{P}\{|[G_H(z)]_{kk}| \geq C + \delta\} \leq \exp(-c\delta^2 N^{3/7}).$$

Now let $\{v_a\}_{a=1}^N$ denote an orthonormal basis of eigenvectors of H and let λ_a be the corresponding eigenvalues. Let $v_a(j)$ denote the jth component of vector v_a in the standard basis. Since

$$G_H(z) = \sum_{a=1}^N \frac{|v_a\rangle\langle v_a|}{\lambda_a - z},$$

hence

$$\operatorname{Im}G_{kk}(x + i\eta) = \sum_{a=1}^N \frac{\eta|v_a(k)|^2}{(\lambda_a - x)^2 + \eta^2}.$$

Let us set $x = \lambda_a$ for a particular value of a, then

$$\operatorname{Im}G_{kk}(\lambda_a + i\eta) \geq \frac{|v_a(k)|^2}{\eta},$$

and, therefore,

$$(21) \qquad |v_a(k)|^2 \leq \eta\operatorname{Im}G_{kk}(\lambda_a + i\eta) \leq C\eta \leq CN^{-1/7}\log N. \qquad \square$$

Before starting the proof of Proposition 3.1, let us exclude $m(z)$ from the free probability system (6):

$$(22) \qquad \begin{aligned} m_\alpha(\omega_\beta(z)) + \frac{1}{\omega_\alpha(z) + \omega_\beta(z) - z} &= 0, \\[2mm] m_\beta(\omega_\alpha(z)) + \frac{1}{\omega_\alpha(z) + \omega_\beta(z) - z} &= 0. \end{aligned}$$

A similar system can be written in the matrix case for $\omega_A(z)$ and $\omega_B(z)$:

$$m_A(\omega_B(z)) + \frac{1}{\omega_A(z) + \omega_B(z) - z} = -r_A(z),$$

(23)

$$m_B(\omega_A(z)) + \frac{1}{\omega_A(z) + \omega_B(z) - z} = -r_B(z),$$

where $r_A(z) := N^{-1}\operatorname{Tr}(R_A(z)) = N^{-1}\operatorname{Tr}(\mathbb{E}R_A(z))$, $r_B(z) := N^{-1}\operatorname{Tr}(R_B(z)) = N^{-1}\operatorname{Tr}(\mathbb{E}R_B(z))$. Here, $R_A(z)$ is defined in (12), and $R_B(z) := \mathbb{E}_V G_H(z) - G_B(\omega_A(z))$.

The proof of Proposition 3.1 is done by an application of the Kantorovich–Newton method that allows us to study how the perturbation of the system for ω_α and ω_β affects the solution. The role of Theorem 1.2 in the proof is to ensure that the size of the perturbation is small.

Let us briefly recall the Newton–Kantorovich method of successive approximations [27]. The method is quite general and works for perturbabions of maps acting on Banach spaces. We will use it for the maps defined on pairs of functions $w_1(z)$, $w_2(z)$ which are holomorphic in a compact domain Ω. However, since the maps can be considered for every z separately, we will essentially consider them as maps from $\mathbb{C}^2$ to $\mathbb{C}^2$ with the norm $\|(w_1, w_2)\| = (|w_1|^2 + |w_2|^2)^{1/2}$.

The general setup is as follows. Let $F(w) = 0$ be a nonlinear functional equation where F is a nonlinear operator that sends elements of a Banach space W to itself. Let F be twice differentiable, and assume that in a neighborhood of a point w_0 the operator $F'(w)$ has an inverse $[F'(w)]^{-1} \in L(W)$ where $L(W)$ denotes the space of bounded linear operators from W to W. Consider the iterations

$$w_{n+1} = w_n - [F'(w_n)]^{-1}F(w_n).$$

The Kantorovich theorem (i) gives the sufficient conditions for the convergence of this process to a solution w^* of equation $F(w) = 0$, (ii) estimates the speed of convergence, and (iii) estimates the distance of the solution w^* from the initial point w_0. We give the statement of the theorem omitting the claim about the speed of convergence, which is not important for us.

THEOREM 3.2 (Kantorovich).   *Suppose that the following conditions hold:*

(i) for an initial approximation w_0, the operator $F'(w_0)$ possesses an inverse operator $\Gamma_0 = [F'(w_0)]^{-1}$ whose norm has the following estimate: $\|\Gamma_0\| \leq C_0$,
    (ii) $\|\Gamma_0 F(w_0)\| \leq \delta_0$,
    (iii) the second derivative $F''(w)$ is bounded in the domain determined by inequality (24) below, namely, $\|F''(w)\| \leq M$,

(iv) *the constants* C_0, δ_0, M *satisfy the relation* $h_0 = C_0 \delta_0 M \leq 1/2$.

Then equation $F(w) = 0$ *has a solution* w^*, *which lies in a neighborhood of* w_0 *determined by the inequality*

$$(24) \qquad \|w - w_0\| \leq \frac{1 - \sqrt{1 - 2h_0}}{h_0} \delta_0,$$

and the successive approximations w_n *of the Newton method converge to* w^*.

PROOF OF PROPOSITION 3.1. We want to prove that the solutions of systems (22) and (23) are close to each other in a certain region of $\mathbb{C}^+$.

Write system (23) as $F(w) = 0$, where

$$F : \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \to \begin{pmatrix} (w_1 + w_2 - z)^{-1} + m_A(w_2) + r_A(z) \\ (w_1 + w_2 - z)^{-1} + m_B(w_1) + r_B(z) \end{pmatrix}.$$

Theorem 3.2 requires estimating three norms, $\|\Gamma_0\|$, $\|\Gamma_0 F(w_0)\|$ and $\|F''(w)\|$. We start by estimating the norm $\|F(w_0)\|$ with $w_0 = (\omega_\alpha(z), \omega_\beta(z))$. We have

$$\|F(w_0)\| = \left\| \begin{pmatrix} m_A(\omega_\beta(z)) - m_\alpha(\omega_\beta(z)) + r_A(z) \\ m_B(\omega_\alpha(z)) - m_\beta(\omega_\alpha(z)) + r_B(z) \end{pmatrix} \right\|.$$

By assumption, $\|(r_A(z), r_B(z))\| \leq r$. To complete the estimate, we also need a lemma.

LEMMA 3.3. *Let* $m_1(z)$ *and* $m_2(z)$ *denote the Stieltjes transforms of measures* μ_1 *and* μ_2, *respectively. Let* $d_L(\mu_1, \mu_2) = s$ *and* $z = x + i\eta$, *where* $\eta > 0$. *Then:*

(a) $|m_1(z) - m_2(z)| < cs\eta^{-1} \max\{1, \eta^{-1}\}$ *where* $c > 0$ *is a numeric constant, and*

(b) $|\frac{d^r}{dz^r}(m_1(z) - m_2(z))| < c_r s \eta^{-1-r} \max\{1, \eta^{-1}\}$ *where* $c > 0$ *are numeric constants.*

This lemma was proved as Lemma 2.2 in [29].

By assumption of smoothness on the closed inteval I, we know that $\operatorname{Im}\omega_\alpha(x)$ and $\operatorname{Im}\omega_\beta(x)$ are uniformly bounded away from zero on I. Moreover, since $\omega_\alpha(z)$ and $\omega_\beta(z)$ are continuous in a rectangle $R_\varepsilon := \{z = x + i\eta | x \in I, 0 \leq \eta \leq \varepsilon\}$ (again by assumption of smoothness in the interval I), hence $\operatorname{Im}\omega_\alpha(z)$ and $\operatorname{Im}\omega_\beta(z)$ are uniformly bounded away from zero on this rectangle provided that ε is sufficiently small. We will use this fact repeatedly below.

In particular, together with Lemma 3.3 this implies that

$$\left\| \begin{pmatrix} m_A(\omega_\beta(z)) - m_\alpha(\omega_\beta(z)) \\ m_B(\omega_\alpha(z)) - m_\beta(\omega_\alpha(z)) \end{pmatrix} \right\| \leq cs$$

on R_ε.

Hence, $\|F(w_0)\|$ is bounded by $O(r+s)$ uniformly for every point $z \in R_\varepsilon$.

The next step is to estimate the norm of the inverse derivative. We compute

$$F' = \begin{pmatrix} -(w_1+w_2-z)^{-2} & -(w_1+w_2-z)^{-2}+m'_A(w_2) \\ -(w_1+w_2-z)^{-2}+m'_B(w_1) & -(w_1+w_2-z)^{-2} \end{pmatrix}.$$

The determinant of this matrix is

$$[m'_A(w_2)+m'_B(w_1)](w_1+w_2-z)^{-2} - m'_A(w_2)m'_B(w_1).$$

Recall that condition (B) in the assumption of smoothness requires that

$$(25)\quad k_\mu(x) := \frac{1}{m'_{\mu_\alpha}(\omega_\beta(x))} + \frac{1}{m'_{\mu_\beta}(\omega_\alpha(x))} - (\omega_\alpha(x)+\omega_\beta(x)-x)^2 \neq 0.$$

By continuity of $\omega_\alpha(z)$, $\omega_\beta(z)$ and $(\omega_\alpha(z)+\omega_\beta(z)-z)^{-2}$ in the rectangle R_ε, we have

$$(26)\quad \left| \frac{m'_{\mu_\alpha}(\omega_\beta(z))+m'_{\mu_\beta}(\omega_\alpha(z))}{(\omega_\alpha(z)+\omega_\beta(z)-z)^2} - m'_{\mu_\alpha}(\omega_\beta(z))m'_{\mu_\beta}(\omega_\alpha(z)) \right| \geq c > 0,$$

everywhere in R_ε, provided that ε is chosen sufficiently small.

By using Lemma 3.3, we conclude that the determinant

$$|[m'_A(w_2)+m'_B(w_1)](w_1+w_2-z)^{-2} - m'_A(w_2)m'_B(w_1)| \geq c > 0,$$

where $w_1 = \omega_\alpha(z)$, $w_2 = \omega_\beta(z)$ and $z \in R_\varepsilon$.

It follows (with some additional help from Lemma 3.3), that the entries of the matrix $[F']^{-1}$ are bounded at $(\omega_\alpha, \omega_\beta)$ if $\operatorname{Im} z$ is sufficiently small [so that $\operatorname{Im}(w_1+w_2-z)$ is bounded away from zero]. By compactness of R_ε and continuity of entries of the matrix $[F']^{-1}$, this shows that the operator norm of $[F']^{-1}$ is bounded at $(\omega_\alpha, \omega_\beta)$ uniformly for $z \in R_\varepsilon$.

By a similar argument, an application of Lemma 3.3 shows that the operator norm of F'' is bounded for all (w_1, w_2) in a fixed neighborhood of $(\omega_\alpha, \omega_\beta)$, and the bound is uniform on R_ε. For example, we can compute

$$\frac{\partial^2 F_1}{(\partial w_2)^2} = 2(w_1+w_2-z)^{-3} + m''_A(w_2),$$

and this is uniformly bounded in a certain neighborhood of $w_1 = \omega_\alpha(z)$, $w_2 = \omega_\beta(z)$ if $z \in R_\varepsilon$ and ε is sufficiently small. The crucial fact here is that the imaginary parts of $\omega_\alpha(z)$ and $\omega_\beta(z)$ are uniformly bounded away from zero for all $z \in R_\varepsilon$.

This shows that conditions (i) and (iii) of the Kantorovich theorem are satisfied with some C_0 and M_0. Since $\|\Gamma_0 F(x_0)\| \leq C_0\|F(x_0)\|$, we define $\delta_0 := C_0\|F(x_0)\|$ and note that $\delta_0 = O(r+s)$. By appropriate choice of $\bar{r}$

and $\overline{s}$, one can make sure that $h_0 = C_0\delta_0 M_0 < 1/2$ and, therefore, that conditions (ii) and (iv) are satisfied. Moreover, one can make sure that h_0 is arbitrarily small, and therefore that the neighborhood in the conclusion of the Kantorovich theorem has the form $\|x - x_0\| \leq \delta_0 = O(r + s)$.

It follows by the Newton–Kantorovich theorem that there exists a solution of the equation $F(w) = 0$ which satisfies the inequalities

$$|w_1(z) - \omega_\alpha(z)| = O(r + s) \quad \text{and} \quad |w_2(z) - \omega_\beta(z)| = O(r + s).$$

The functions $\omega_A(z)$ and $\omega_B(z)$ defined by (2) satisfy equation $F(w) = 0$, and one can show that for every fixed z they approach $\omega_\alpha(z)$ and $\omega_\beta(z)$ as $N \to \infty$. Hence, for sufficiently small $\overline{r}$ and $\overline{s}$ the solution of $F(w) = 0$ found by the Newton–Kantorovich method coincide with the pair $(\omega_A(z), \omega_B(z))$ and we can conclude that

$$|\omega_A(z) - \omega_\alpha(z)| = O(r + s) \quad \text{and} \quad |\omega_B(z) - \omega_\beta(z)| = O(r + s).$$

This completes the proof of Proposition 3.1 and Theorem 1.3. □

4. Local law for eigenvalues. Let $\eta^* = M\eta$ and $I_{\eta^*} = [x - \eta^* + i\eta, x + \eta^* + i\eta]$. Recall that

$$s(A, B) := \max\{d_L(\mu_A, \mu_\alpha), d_L(\mu_B, \mu_\beta)\}.$$

PROPOSITION 4.1. *Assume that a pair of probability measures (μ_α, μ_β) is smooth in a closed interval I. Assume that $s(A, B) \leq \overline{s}$ where $\overline{s}$ is a positive constant. Let $\eta = N^{-1/7}\log N$. Then for some positive c and c_1, and for every $\varepsilon > 0$,*

$$(27) \quad \mathbb{P}\Big\{\sup_{z \in I_{\eta^*}} |m_H(z) - m_{\mu_\alpha \boxplus \mu_\beta}(z)| > \varepsilon + cs(A, B)\Big\} \leq \exp(-c_1(\log N)^2)$$

for all sufficiently large N.

PROOF. Proposition 4.1 is proved by combining Lemmas 4.2 and 4.3 below.

LEMMA 4.2. *Assume that a pair of probability measures (μ_α, μ_β) is smooth in a closed interval I. Let r and s be as defined in (18) and (19), respectively. Then for all sufficiently small r, s and η,*

$$|\mathbb{E}m_H(x + i\eta) - m_{\mu_\alpha \boxplus \mu_\beta}(x + i\eta)| < O(r + s).$$

Indeed, since $\mathbb{E}m_H = (\omega_A + \omega_B - z)^{-1}$ and $m_{\mu_\alpha \boxplus \mu_\beta} = (\omega_\alpha + \omega_\beta - z)^{-1}$, therefore,

$$\mathbb{E}m_H - m_{\mu_\alpha \boxplus \mu_\beta} = \frac{\omega_\alpha + \omega_\beta - \omega_A - \omega_B}{(\omega_A + \omega_B - z)(\omega_\alpha + \omega_\beta - z)}.$$

The denominator is bounded away from zero for small η by Proposition 3.1 [$\operatorname{Im}\omega_\alpha(x)$ and $\operatorname{Im}\omega_\beta(x)$ are bounded away from 0 by the assumption of Proposition 3.1, and ω_A and ω_B are close to ω_α and ω_β, respectively, by its conclusion]. The numerator can be estimated by Proposition 3.1 as $O(r+s)$.

In [28], the following result was proved (as Corollary 6).

LEMMA 4.3.   *For some positive c and c_1 which can depend on M, and for all $\delta > 0$,*

$$\mathbb{P}\Big\{\sup_{z\in I_{\eta^*}}|m_H(z) - \mathbb{E}m_H(z)| > \delta\Big\} \leq \exp\Big(-\frac{c\delta^2\eta^4}{\|B\|^2}N^2\Big),$$

provided that $N \geq c_1(\sqrt{-\log(\eta\delta)})/(\eta^2\delta)$.

Let us take $\delta = c\log N/(N\eta^2)$ in Lemma 4.3. Then $N \geq c_1(\sqrt{-\log(\eta\delta)})/(\eta^2\delta)$ provided that $\eta \geq N^{-1}\log N$. In particular, if $\eta = N^{-1/7}\log N$, then Lemma 4.3 implies that

$$\mathbb{P}\Big\{\sup_{z\in I_{\eta^*}}|m_H(z) - \mathbb{E}m_H(z)| > \frac{1}{N^{5/7}\log N}\Big\} \leq \exp(-c(\log N)^2).$$

In addition, the second part of Theorem 1.2 and the definition of r imply that if $\eta \gg N^{-1/7}$, then for every $\varepsilon > 0$ and all sufficiently large N, we have $|r(z)| < \varepsilon$. Hence, Lemma 4.2 implies that if $\eta \gg N^{-1/7}$, then

$$|\mathbb{E}m_H(z) - m_{\mu_\alpha\boxplus\mu_\beta}(z)| < \varepsilon + cs(A,B)$$

for all sufficiently large N. Together, these statements imply the claim of Proposition 4.1. $\square$

PROOF OF THEOREM 1.4.   The proof is similar to the proof of Corollary 4.2 in [23]. Let $\eta = cN^{-1/7}$, and c is sufficiently large, and let $\eta^* = M\eta$. Let

$$R(\lambda) := \frac{1}{\pi}\int_{x-\eta^*}^{x+\eta^*}\frac{\eta}{(x-\lambda)^2 + \eta^2}\,dx$$

$$= \frac{1}{\pi}\Big(\arctan\Big(\frac{x-\lambda}{\eta} + M\Big) - \arctan\Big(\frac{x-\lambda}{\eta} - M\Big)\Big).$$

Then $R = 1_{I^*} + T_1 + T_2 + T_3$, where 1_{I^*} is the indicator function of the interval $I^* = [x-\eta^*, x+\eta^*]$ and functions T_1, T_2 and T_3 satisfy the following properties:

$$|T_1| \leq c/\sqrt{M}, \qquad \operatorname{supp}(T_1) \subset I_1 = [x - 2\eta^*, x + 2\eta^*],$$
$$|T_2| \leq 1, \qquad \operatorname{supp}(T_2) \subset J_1 \cup J_2,$$

where J_1 and J_2 are intervals of length $\sqrt{M}\eta$ with midpoints at $x - \eta^*$ and $x + \eta^*$, respectively, and

$$|T_3| \leq \frac{C\eta\eta^*}{(\lambda - x)^2 + (\eta^*)^2}, \qquad \operatorname{supp}(T_3) \in I_1^c.$$

Note that

$$\frac{\mathcal{N}_{\eta^*}(x)}{2\eta^* N} = \frac{1}{2\eta^*} \int 1_{I^*}(\lambda) \mu_{H_N}(d\lambda)$$

$$= \frac{1}{2\eta^*} \int R(\lambda) \mu_{H_N}(d\lambda) - \frac{1}{2\eta^*} \int (T_1 + T_2 + T_3) \mu_{H_N}(d\lambda).$$

The last integral can be estimated as follows:

$$\frac{1}{2\eta^*} \int |T_1 + T_2 + T_3| \mu_{H_N}(d\lambda) \leq \frac{c}{\sqrt{M}} \frac{\mathcal{N}_{I_1}}{2\eta^* N} + \frac{\mathcal{N}_{J_1} + \mathcal{N}_{J_2}}{2\eta^* N} + \frac{C\eta}{\eta^*} \rho_{\eta^*}(x),$$

where $\mathcal{N}_I$ denote the number of eigenvalues of H_N in interval I, and

$$\rho_{\eta^*}(x) := \frac{1}{\pi} \operatorname{Im} m_{H_N}(x + i\eta^*) = \frac{1}{\pi} \int \frac{\eta^*}{(x - \lambda)^2 + (\eta^*)^2} \mu_{H_N}(d\lambda).$$

Hence, by using the inequality $\mathcal{N}_\eta(x) \leq CN\eta\rho_\eta(x)$, one obtains

$$\frac{1}{2\eta^*} \int |T_1 + T_2 + T_3| \mu_{H_N}(d\lambda)$$

(28)

$$\leq \frac{c}{\sqrt{M}} (\rho_{2\eta^*}(x) + \rho_{\sqrt{M}\eta}(x - \eta^*) + \rho_{\sqrt{M}\eta}(x + \eta^*) + \rho_{\eta^*}(x)).$$

By the second path of Theorem 1.2, $\mathbb{E} m_{H_N}(x + i\eta^*) - m_{A_N}(\omega_B(x + i\eta^*)) = O(\frac{1}{N^2\eta^6}) = o(1)$. In addition, the assumption of smoothness and formula (20) imply that $m_{A_N}(\omega_B(x + i\eta^*))$ is bounded for every x in the interval I. Hence, the integral in (28) is bounded by $O(M^{-1/2})$.

The main term can be written as

$$\frac{1}{2\eta^*} \int_{I^*} \frac{1}{\pi} \operatorname{Im} m_{\mu_\alpha \boxplus \mu_\beta}(x + i\eta) \, dx$$

$$+ \frac{1}{2\eta^*} \int_{I^*} \frac{1}{\pi} \operatorname{Im}(m_{H_N}(x + i\eta) - m_{\mu_\alpha \boxplus \mu_\beta}(x + i\eta)) \, dx.$$

The first part converges to $\rho_{\mu_\alpha \boxplus \mu_\beta}(x)$ because the assumption that (μ_α, μ_β) is smooth at x implies that $\mu_\alpha \boxplus \mu_\beta$ has an analytic density in a neighborhood of x. For the second term, we can use the estimate in Proposition 4.1. The assumption that $s(A_N, B_N) = \max\{d_L(\mu_{A_N}, \mu_\alpha), d_L(\mu_{B_N}, \mu_\beta)\} \to 0$ implies that this term converges to 0 in probability as $N \to \infty$. $\quad\square$

5. Subordination and spikes. The proof of Theorem 1.5 is based on the following result.

PROPOSITION 5.1. *Let $H = A + UBU^*$ where A is a rank-one N-by-N Hermitian matrix with the nonzero eigenvalue θ, and B is an N-by-N Hermitian matrix with the empirical eigenvalue distribution μ_B. Then the expected Stieltjes transform of H satisfies the following equation for every $z \in \mathbb{C}^+$:*

$$(29) \quad \mathbb{E}m_H(z) = m_B(z) + \frac{1}{N}\frac{m'_B(z)}{m_B(z)}\left(\frac{1}{\theta m_B(z) + 1} - 1\right) + O_\eta\left(\frac{1}{N^2}\right).$$

Here $O_\eta(N^{-2})$ denotes a function $f(z)$ such that $N^2|f(z)| \le C(\operatorname{Im} z)^{-k}$ for some $k > 0$ and $C > 0$.

PROOF OF PROPOSITION 5.1. Note that

$$m_A(z) = -\frac{1}{z} + \frac{1}{N}\left(\frac{1}{\theta - z} + \frac{1}{z}\right).$$

From Theorem 1.2, we know that the following system holds for $\mathbb{E}m_H(z)$, $\omega_A(z)$, $\omega_B(z)$:

$$\mathbb{E}m_H(z) = -\frac{1}{\omega_B(z)} + \frac{1}{N}\left(\frac{1}{\theta - \omega_B(z)} + \frac{1}{\omega_B(z)}\right) + O_\eta(N^{-2}),$$

$$(30) \quad \mathbb{E}m_H(z) = m_B(\omega_A(z)) + O_\eta(N^{-2}) \quad \text{and}$$

$$z = \omega_A(z) + \omega_B(z) + \frac{1}{\mathbb{E}m_H(z)}.$$

This system can be considered as a perturbation of the system

$$\overline{m}_H(z) + \frac{1}{\overline{\omega}_B(z)} = 0,$$

$$(31) \quad \overline{m}_H(z) - m_B(\overline{\omega}_A(z)) = 0 \quad \text{and}$$

$$\overline{\omega}_A(z) + \overline{\omega}_B(z) + \frac{1}{\overline{m}_H(z)} = z.$$

The solution of the unperturbed system is $\overline{m}_H(z) = m_B(z)$, $\overline{\omega}_A(z) = z$ and $\overline{\omega}_B(z) = -1/m_B(z)$. We compute the derivative of the unperturbed system (31) with respect to $(\overline{m}_H, \overline{\omega}_A, \overline{\omega}_B)$ at the solution and find

$$J = \begin{pmatrix} 1 & 0 & -\dfrac{1}{\overline{\omega}_B^2(z)} \\ 1 & -m'_B(\overline{\omega}_A(z)) & 0 \\ -\dfrac{1}{\overline{m}_H^2(z)} & 1 & 1 \end{pmatrix}$$

(32)
$$= \begin{pmatrix} 1 & 0 & -m_B^2(z) \\ 1 & -m_B'(z) & 0 \\ -\dfrac{1}{\overline{m}_B^2(z)} & 1 & 1 \end{pmatrix}.$$

From (30), the perturbation of the system is

$$\left(\frac{1}{N} \left(\frac{1}{\theta - \overline{\omega}_B(z)} + \frac{1}{\overline{\omega}_B(z)} \right) + O_\eta(N^{-2}), O_\eta(N^{-2}), 0 \right).$$

Note that

$$\frac{1}{N} \left(\frac{1}{\theta - \overline{\omega}_B(z)} + \frac{1}{\overline{\omega}_B(z)} \right) = \frac{m_B(z)}{N} \left(\frac{1}{\theta m_B(z) + 1} - 1 \right).$$

Hence, the linearized system is

$$J \begin{pmatrix} \Delta \overline{m}_H \\ \Delta \overline{\omega}_A \\ \Delta \overline{\omega}_B \end{pmatrix} = \begin{pmatrix} \dfrac{m_B(z)}{N} \left(\dfrac{1}{\theta m_B(z) + 1} - 1 \right) + O_\eta(N^{-2}) \\ O_\eta(N^{-2}) \\ 0 \end{pmatrix},$$

where $\Delta \overline{m}_H$, $\Delta \overline{\omega}_A$ and $\Delta \overline{\omega}_B$ denote the first-order changes in the solution caused by perturbation. By using this linearization and the formula (32) for the derivative J, we can easily compute the linear approximation for the solution of the perturbed system. In particular,

$$\mathbb{E} m_H(z) = m_B(z) + \frac{1}{N} \frac{m_B'(z)}{m_B(z)} \left(\frac{1}{\theta m_B(z) + 1} - 1 \right) + O_\eta \left(\frac{1}{N^2} \right).$$

The contribution of the higher order terms is $O_\eta(N^{-2})$. $\square$

PROOF OF THEOREM 1.5.   Proof of this theorem is similar to the proof of Theorem 2.1 in [17] and for this reason we will be concise. Let us start with the case when $\rho_\mu(\theta_0) > L$. Then the first step is to show that for large N there are no eigenvalues of H_N in $S_\varepsilon := (L+\varepsilon, \rho_\mu(\theta_0)-\varepsilon) \cup (\rho_\mu(\theta_0)+\varepsilon, \infty)$. In order to do this, we note that for all sufficiently large N, the first correction term in formula (29),

$$\mathcal{L}_N(z) := \frac{m_{B_N}'}{m_{B_N}} \left(\frac{1}{\theta m_{B_N} + 1} - 1 \right),$$

is the Stieltjes transform of a distribution Λ_{B_N} with a compact support which must be outside of S_ε. Verification of this fact can be done as in the proof of Proposition 4.5 in [17].

Next, one can use the Stieltjes inversion formula, which holds for distributions by the results of Tillmann in [42]. Applying it to formula (29), one finds that for every $\varphi \in C_c^\infty(\mathbb{R})$,

$$\mathbb{E}[N^{-1}\operatorname{Tr}(\varphi(H_N))]$$
$$= \int \varphi \, d\mu_{B_N} + \frac{1}{N}\Lambda_{B_N}(\varphi) - \frac{1}{\pi}\lim_{\eta \to 0^+}\operatorname{Im}\int_{\mathbb{R}}\varphi(x)f(x+i\eta)\,dx,$$

where $f(x)$ denotes the error term in (29), $f(x) = O_\eta(N^{-2})$. The last term is $O(N^{-2})$ (see Section 6 in [25] or the Appendix in [16]) and, therefore, we find that

$$\mathbb{E}[N^{-1}\operatorname{Tr}(\varphi(H_N))] = \int \varphi \, d\mu_{B_N} + \frac{1}{N}\Lambda_{B_N}(\varphi) + O(N^{-2}).$$

In particular, if the support of φ is in S_ε, then the first and the second terms are zero and $\mathbb{E}[N^{-1}\operatorname{Tr}(\varphi(H_N))] = O(N^{-2})$. If in addition φ is nonnegative, then by the Markov inequality

$$\mathbb{P}\left[N^{-1}\operatorname{Tr}(\varphi(H_N)) > \frac{1}{2N}\right] < \frac{\mathbb{E}[N^{-1}\operatorname{Tr}(\varphi(H_N))]}{2N} = O\left(\frac{1}{N}\right).$$

By using a sequence of functions φ that approximate the indicator function of S_ε, it follows that

$$\mathbb{P}[\text{there is an eigenvalue of } H_N \text{ in } S_\varepsilon] < \frac{c}{N}.$$

The next step is to show that for sufficiently large N, there is exactly one eigenvalue to the right of $\rho_\mu(\theta_0) - \varepsilon$. This can be done similarly to the corresponding result (Theorem 4.5) in [17]. Namely, note that $\rho_\mu(\theta)$ is an increasing function for $\theta > \theta_0 - \varepsilon$. [This follows from the fact that $m_\mu(x)$ is a decreasing function for $x > L$.] Hence, we can find an interval $[\alpha, \beta]$ in $(\theta_0 - \varepsilon, \theta_0)$ that will map to an interval $[a, b]$ in $(L, \rho_\mu(\theta_0))$, with $a := \rho_\mu(\alpha)$ and $b := \rho_\mu(\beta)$. The claim is that if $\lambda_1(H_N)$ and $\lambda_2(H_N)$ are the largest and the second largest eigenvalues of H_N, then

$$\mathbb{P}[\lambda_2(H_N) < a \text{ and } \lambda_1(H_N) > b] \to 1$$

as $N \to \infty$.

From interlacing inequalities for matrices, we immediately obtain that $\lambda_2(H_N) < a$ for sufficiently large N. In order to prove that $\lambda_1(H_N) > b$, we consider matrix $cA_N + B_N$. By using Weyl's inequalities and the uniform bound on norms of B_N we obtain that the largest eigenvalue $\lambda_1(cA_N + B_N) \geq c\theta - \delta$ for some positive δ. On the other hand $\rho_\mu(c\beta) \sim c\beta$ for large c. We conclude that

$$\lambda_1(\overline{c}A_N + B_N) > \rho_\mu(\overline{c}\beta)$$

for a sufficiently large $\overline{c}$. In addition, Weyl's inequalities imply that $\lambda_1(c_1 A_N + B_N) - \lambda_1(c_2 A_N + B_N) \leq |c_1 - c_2|\theta_0$. Hence, if c changes slowly, then the first eigenvalue of $cA_N + B_N$ changes slowly. By what we proved above, there are no eigenvalues of $cA_N + B_N$ in the interval $(L + \varepsilon, \rho_\mu(c\theta_0) - \varepsilon)$ with large probability. Since $\rho_\mu(c\theta_0)$ is an increasing function of c for $c \geq 1$, hence the length of this interval is always $\geq \rho_\mu(\theta_0) - L - 2\varepsilon > \varepsilon' > 0$. By changing c along a finite sequence $\overline{c} = c_1 > c_2 > \cdots > c_l = 1$ with $|c_i - c_{i+1}| \leq \varepsilon'/\theta_0$, we can ensure that $\lambda_1(c_i A_N + B_N) > \rho_\mu(c_i \beta)$ for all i with large probability. Hence, as N grows, the probability that $\lambda_1(H_N) \geq \rho_\mu(\beta) > \rho_\mu(\theta) - \varepsilon$ approaches 1. Together with the fact that with high probability the interval $(L + \varepsilon, \rho_\mu(\theta) - \varepsilon) \cup (\rho_\mu(\theta) + \varepsilon, \infty)$ contains no eigenvalues, this implies that λ_1 converges in probability to $\rho_\mu(\theta)$ as $N \to \infty$.

Next, consider the case when $\rho_\mu(\theta_0) \leq L$. Then we conclude (by the argument at the start of the proof) that for every fixed $\varepsilon > 0$ there are no eigenvalues of H_N in $S_\varepsilon := (L + \varepsilon, \infty)$ with high probability for large N. On the other hand, by Weyl's inequalities $\lambda_1(H_N) \geq \lambda_1(B_N)$. Since $\lambda_1(B_N) \to L$ in probability, we conclude that $\lambda_1(H_N) \to L$ in probability. $\square$

APPENDIX A: A DERIVATION OF FORMULA (13)

Let $G(z) \equiv G_H(z) = (A + B - z)^{-1}$, where $B = U\widetilde{B}U^*$ and U is a uniformly distributed unitary matrix. Let $B_t = e^{iXt} B e^{-iXt}$ where X is Hermitian and let $G_t = (A + B_t - z)^{-1}$. Then $\mathbb{E}_U(dG_t/dt) = 0$ for every Hermitian matrix X. Let us for clarity omit the subscript U in the expectations below and treat A as fixed. It is easy to compute that $\partial G_{uv}/\partial B_{xy} = -G_{ux}G_{yv}$ and that $dB_t/dt = i[X, B]$. By using the chain rule, we calculate dG_t/dt and infer that

$$\mathbb{E}((G_H)_{ua}(BG_H)_{bv}) = \mathbb{E}((G_H B)_{ua}(G_H)_{bv}).$$

Setting $u = a$ and summing over all a gives the identity

$$\mathbb{E}(m_H B G_H) = \mathbb{E}(f_B G_H).$$

It follows that

$$\mathbb{E}(m_H G_H) = \mathbb{E}(m_H G_A - m_H G_A B G_H)$$
$$= \mathbb{E}(m_H G_A - G_A f_B G_H),$$

where we used the identity $G_H(z) = G_A(z) - G_A(z)BG_H(z)$ in the first line.

This can be written in the following equivalent form:

$$\mathbb{E}m_H \mathbb{E}G_H = (\mathbb{E}m_H)G_A - (\mathbb{E}f_B)G_A \mathbb{E}G_H$$
$$- \mathbb{E}[(m_H - \mathbb{E}m_H)G_H] - G_A \mathbb{E}[(f_B - \mathbb{E}f_B)G_H]$$
$$= (\mathbb{E}m_H)G_A - (\mathbb{E}f_B)G_A \mathbb{E}G_H + \mathbb{E}\Delta_A,$$

where

$$\Delta_A = -(m_H - \mathbb{E}m_H)G_H - G_A(f_B - \mathbb{E}f_B)G_H.$$

This expression can be further rewritten (after we multiply it by $A - z$ and rearrange terms) as

$$\mathbb{E}m_H \left(A - \left(z - \frac{\mathbb{E}f_B}{\mathbb{E}m_H} \right) \right) \mathbb{E}G_H = \mathbb{E}m_H + (A - z)\mathbb{E}\Delta_A.$$

Let $z' := z - \mathbb{E}f_B/\mathbb{E}m_H$. Then

$$\mathbb{E}m_H \mathbb{E}G_H = G_A(z')\mathbb{E}m_H + (A - z)G_A(z')\mathbb{E}\Delta_A.$$

Divide the resulting expression by $\mathbb{E}m_H$. Then we obtain

$$\mathbb{E}G_H(z) = G_A(z') + \frac{1}{\mathbb{E}m_H}((A - z)G_A(z')\mathbb{E}\Delta_A)$$

$$= G_A(z') + R_A.$$

APPENDIX B: SOME HELPFUL LEMMAS ABOUT EXPECTED RESOLVENT

The following result is from [7].

LEMMA B.1.   *Suppose that U is a uniformly distributed random unitary matrix. Then $\mathbb{E}[(A + UBU^*)^{-1}]$ belongs to the algebra generated by the matrix A. In particular, if A is diagonal, then $\mathbb{E}[(A + UBU^*)^{-1}]$ is diagonal.*

PROOF.   If V is an arbitrary unitary matrix that commutes with A, then

$$V\mathbb{E}[(A + UBU^*)^{-1}]V^* = \mathbb{E}[(VAV^* + VUB(VU)^*)^{-1}]$$

$$= \mathbb{E}[(A + UBU^*)^{-1}].$$

Hence, $\mathbb{E}[(A + UBU^*)^{-1}]$ commutes with V. Since von Neumann algebras are generated by their unitaries, we conclude that $\mathbb{E}[(A + UBU^*)^{-1}]$ belongs to the bicommutant of A. By the basic theorem about von Neumann algebras, this bicommutant coincides with the algebra generated by A.   □

Similarly, one can prove the following result.

LEMMA B.2.   *Suppose that U is a uniformly distributed random unitary matrix. Then*

$$\mathbb{E}[UBU^*] = \left(\frac{1}{N} \operatorname{Tr}(B) \right) I_N.$$

LEMMA B.3. *Let A_j, $j = 1, \ldots, m$, be a family of normal (finite-dimensional) operators. Suppose that the eigenvalues of all A_j are contained in a closed disc $D \subset \mathbb{C}$, and let $H = \sum p_j A_j$ be a convex combination of A_j. Then all eigenvalues of H are contained in D.*

PROOF. By subtracting a multiple of the identity operator from all A_j, we can reduce the problem to the case when disc D has its center at 0. Assume that this is indeed the case. Let R be the radius of D. Since the operators are normal, their norms are equal to the maximum of the absolute values of eigenvalues. Hence, $\|A_j\| \leq R$. Hence, $\|H\| \leq \sum p_j \|A_j\| \leq R$. It follows that all eigenvalues of H have absolute value $\leq R$. □

APPENDIX C: ESTIMATES OF THE RESOLVENT ENTRIES, THE STIELTJES TRANSFORM AND RELATED QUANTITIES

In this section, we assume that $G(z) = (A + UBU^* - z)^{-1}$, where A and B are N-by-N Hermitian matrices and U is a random Haar-distributed unitary matrix.

LEMMA C.1. *Let $z = E + i\eta$ where $\eta > 0$. Then, for a numeric $c > 0$ and every $\delta > 0$:*

(i)

$$\mathbb{P}\{|G_{ij}(z) - \mathbb{E}G_{ij}(z)| > \delta\} \leq \exp\left(-\frac{c\delta^2\eta^4}{\|B\|^2}N\right) \quad and$$

(33)

$$\mathrm{Var}(G_{ij}(z)) \leq \frac{\|B\|^2}{c\eta^4 N};$$

(ii) *Let $h := N^{-1}\mathrm{Tr}(FG)$, where F does not depend on U. Then*

$$\mathbb{P}\{|h(z) - \mathbb{E}h(z)| > \delta\} \leq \exp\left(-\frac{c\delta^2\eta^4}{\|F\|^2\|B\|^2}N^2\right) \quad and$$

(34)

$$\mathrm{Var}(h(z)) \leq \frac{\|F\|^2\|B\|^2}{c\eta^4 N^2}.$$

REMARK. By applying the second part of the lemma to $h = I$, $A - z$ and $(A - z)^{-1}$, we can compute probabilities of deviations and variances for $m(z) := N^{-1}\mathrm{Tr}\,G(z)$, $f_B(z) := N^{-1}\mathrm{Tr}(BG(z)) = 1 - N^{-1}\mathrm{Tr}((A - z)G(z))$ and $h_A(z) = N^{-1}\mathrm{Tr}((A - z)^{-1}G(z))$, respectively. In particular,

$$\mathbb{P}\{|m(z) - \mathbb{E}m(z)| > \delta\} \leq \exp\left(-\frac{c\delta^2\eta^4}{\|B\|^2}N^2\right) \quad and$$

(35)
$$\mathrm{Var}(m(z)) \leq \frac{\|B\|^2}{c\eta^4 N^2};$$

$$\mathbb{P}\{|f_B(z) - \mathbb{E}f_B(z)| > \delta\} \leq \exp\left[-\frac{c\delta^2\eta^4}{\|A-z\|^2\|B\|^2}N^2\right] \quad \text{and}$$

(36)
$$\mathrm{Var}(f_B(z)) \leq \frac{\|A-z\|^2\|B\|^2}{c\eta^4 N^2};$$

$$\mathbb{P}\{|h_A(z) - \mathbb{E}h_A(z)| > \delta\} \leq \exp\left[-\frac{c\delta^2\eta^6}{\|B\|^2}N^2\right] \quad \text{and}$$

(37)
$$\mathrm{Var}(h_A(z)) \leq \frac{\|B\|^2}{c\eta^6 N^2}.$$

PROOF OF LEMMA C.1. (i) In a small neighborhood of identity matrix, all unitary matrices can be written as $U = e^{iX}$ where X is Hermitian. Then G_H can be thought of as a function of X and we can compute its derivative as follows [let $\widetilde{B}$ denote UBU^*, $B(X) = e^{iX}\widetilde{B}e^{-iX}$ and $G_H(z,X) = (A + B(X) - z)^{-1}$]:

$$|d_X G_H(z,X)| = \left|\sum_{x,y} \frac{\partial G_H(z)}{\partial \widetilde{B}_{xy}} d_X B_{xy}(X)\right|$$

$$= \left|\sum_{x,y} \frac{\partial G_H(z)}{\partial \widetilde{B}_{xy}}[X,\widetilde{B}]_{xy}\right|$$

$$= \left|\sum_{x,y} \left[\frac{\partial G_H(z)}{\partial \widetilde{B}_{xy}}, \widetilde{B}\right]X_{xy}\right|,$$

where we used the fact that $d_X B(X)|_{X=0} = [X,\widetilde{B}] \equiv X\widetilde{B} - \widetilde{B}X$.

We compute

$$\frac{\partial G_{ij}}{\partial \widetilde{B}_{xy}} = -G_{ix}G_{yj}.$$

Therefore,

$$\left\|\frac{\partial G_{ij}}{\partial \widetilde{B}_{xy}}\right\|_2 = \sqrt{\sum_{x,y}|G_{xi}|^2|G_{yj}|^2} = \sqrt{\|Ge_i\|^2\|Ge_j\|^2} \leq \|G\|^2 \leq \frac{1}{\eta^2},$$

where $\|M\|_2 := \mathrm{Tr}(M^*M)$ is the Frobenius norm of matrix M.

If $\|X\|_2 = 1$, then it follows that

$$|d_X G_{ij}(z,X)| \leq \left\|\left[\frac{\partial G_{ij}(z)}{\partial \widetilde{B}_{xy}}, \widetilde{B}\right]\right\|_2$$

$$\leq 2\left\|\frac{\partial G_{ij}(z)}{\partial \widetilde{B}_{xy}}\right\|_2 \|B\|$$

$$\leq \frac{2\|B\|}{\eta^2}.$$

In the second line, we used the fact that $\|AB\|_2 \leq \|A\|_2\|B\|$. (See Exercise 20 on page 313 in Section 5.6 of [26].)

Next, we note that the Ricci's curvature of $SU(N)$ is $(N/2)I$ with respect to the metric induced by $\|\cdot\|_2$ norm on X. By Gromov's theorem, if $g\colon(\mathcal{SU}(N),\|ds\|_2)\to\mathbb{R}$ is an L-Lipschitz function and if $\mathbb{E}g=0$, then $P\{|g|>\delta\}\leq\exp(-cN\delta^2/L^2)$ for every $\delta>0$ and some numeric $c>0$. For details of the argument, the reader can consult Section 4.4.2 in [1], especially Theorem 4.4.7. We apply this theorem to a complex-valued function but the proof is the same except for some minor changes.

For variance, we note that for every positive random variable X, it is true that $\mathbb{E}X = \int_0^\infty (1 - \mathcal{F}_X(t))\,dt$, where $\mathcal{F}_X(t)$ is cumulative distribution function of X. We can apply this to the random variables $(\mathrm{Im}(G_{ij} - \mathbb{E}G_{ij}))^2$ and $(\mathrm{Re}(G_{ij} - \mathbb{E}G_{ij}))^2$, and find that the expectation of both expression is smaller than $\frac{\|B\|^2}{c\eta^4 N}$. Hence,

$$\mathrm{Var}(G_{ij}(z)) \equiv \mathbb{E}((G_{ij} - \mathbb{E}G_{ij})(\overline{G_{ij} - \mathbb{E}G_{ij}})) \leq \frac{\|B\|^2}{c\eta^4 N}$$

with a possibly different constant.

(ii) The proof is similar and boils down to showing that if $h(z,X) := N^{-1}\,\mathrm{Tr}(FG_H(z,X))$ and if $\|X\|_2 = 1$, then

$$|d_X h_A(z,X)| = \left|\frac{1}{N}\sum_{x,y}([GFG,\widetilde{B}])_{yx}X_{xy}\right|$$

$$\leq \frac{2}{\sqrt{N}}\|GFG\|\|B\|$$

$$\leq \frac{2\|F\|\|B\|}{\eta^2\sqrt{N}}. \qquad\qquad \square$$

LEMMA C.2.   *Assume that* $\max\{\|A\|,\|B\|\}\leq K$, $\mathrm{Im}\,z=\eta>0$ *and* $|z|\leq R$. *We have* $(\mathbb{E}m_H(z))^{-1}\leq c'/\eta$, *where* c' *depends only on* K *and* R.

PROOF.   We have

$$\mathrm{Im}\,\mathbb{E}\left[\frac{1}{N}\,\mathrm{Tr}\,G_H(x+i\eta)\right] = \mathbb{E}\left[\frac{\eta}{N}\,\mathrm{Tr}[((H-xI_N)^2+\eta^2 I_N)^{-1}]\right].$$

Since all eigenvalues of $(H - xI_N)^2 + \eta^2 I_N$ are $\leq ((K+R)^2 + R^2)$, hence all eigenvalues of $((H - xI_N)^2 + \eta^2 I_N)^{-1}$ are $\geq ((K+R)^2 + R^2)^{-1}$ and, therefore,

$$\mathbb{E}\left[\frac{\eta}{N}\operatorname{Tr}[(H - xI_N)^2 + \eta^2 I_N]^{-1}\right] \geq c\eta,$$

which implies the claim of the lemma. $\square$

REFERENCES

[1] ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2010). *An Introduction to Random Matrices. Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge. MR2760897

[2] BAI, Z. and YAO, J.-F. (2008). Central limit theorems for eigenvalues in a spiked population model. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 447–474. MR2451053

[3] BAIK, J., BEN AROUS, G. and PÉCHÉ, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** 1643–1697. MR2165575

[4] BAIK, J. and SILVERSTEIN, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97** 1382–1408. MR2279680

[5] BELINSCHI, S. T. (2008). The Lebesgue decomposition of the free additive convolution of two probability distributions. *Probab. Theory Related Fields* **142** 125–150. MR2413268

[6] BELINSCHI, S. T. and BERCOVICI, H. (2007). A new approach to subordination results in free probability. *J. Anal. Math.* **101** 357–365. MR2346550

[7] BELINSCHI, S. T., BERCOVICI, H., CAPITAINE, M. and FÉVRIER, M. (2012). Outliers in the spectrum of large deformed unitarily invariant models. Available at arXiv:1207.5443.

[8] BENAYCH-GEORGES, F., GUIONNET, A. and MAIDA, M. (2011). Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices. *Electron. J. Probab.* **16** 1621–1662. MR2835249

[9] BENAYCH-GEORGES, F., GUIONNET, A. and MAIDA, M. (2012). Large deviations of the extreme eigenvalues of random deformations of matrices. *Probab. Theory Related Fields* **154** 703–751. MR3000560

[10] BENAYCH-GEORGES, F. and NADAKUDITI, R. R. (2011). The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.* **227** 494–521. MR2782201

[11] BENAYCH-GEORGES, F. and NADAKUDITI, R. R. (2012). The singular values and vectors of low rank perturbations of large rectangular random matrices. *J. Multivariate Anal.* **111** 120–135. MR2944410

[12] BIANE, P. (1998). Processes with free increments. *Math. Z.* **227** 143–174. MR1605393

[13] BLOEMENDAL, A. and VIRÁG, B. (2011). Limits of spiked random matrices II. Available at arXiv:1109.3704.

[14] BLOEMENDAL, A. and VIRÁG, B. (2013). Limits of spiked random matrices I. *Probab. Theory Related Fields* **156** 795–825. MR3078286

[15] CAPITAINE, M. (2013). Additive/multiplicative free subordination property and limiting eigenvectors of spiked additive deformations of Wigner matrices and spiked sample covariance matrices. *J. Theoret. Probab.* **26** 595–648. MR3090543

[16] CAPITAINE, M. and DONATI-MARTIN, C. (2007). Strong asymptotic freeness for Wigner and Wishart matrices. *Indiana Univ. Math. J.* **56** 767–803. MR2317545

[17] CAPITAINE, M., DONATI-MARTIN, C. and FÉRAL, D. (2009). The largest eigenvalues of finite rank deformation of large Wigner matrices: Convergence and nonuniversality of the fluctuations. *Ann. Probab.* **37** 1–47. MR2489158

[18] CAPITAINE, M., DONATI-MARTIN, C., FÉRAL, D. and FÉVRIER, M. (2011). Free convolution with a semicircular distribution and eigenvalues of spiked deformations of Wigner matrices. *Electron. J. Probab.* **16** 1750–1792. MR2835253

[19] CHISTYAKOV, G. P. and GÖTZE, F. (2011). The arithmetic of distributions in free probability theory. *Cent. Eur. J. Math.* **9** 997–1050. MR2824443

[20] ERDŐS, L. (2011). Universality of Wigner random matrices: A survey of recent results. *Uspekhi Mat. Nauk* **66** 67–198. MR2859190

[21] ERDŐS, L. and KNOWLES, A. (2011). Quantum diffusion and eigenfunction delocalization in a random band matrix model. *Comm. Math. Phys.* **303** 509–554. MR2782623

[22] ERDŐS, L., KNOWLES, A., YAU, H.-T. and YIN, J. (2013). The local semicircle law for a general class of random matrices. *Electron. J. Probab.* **18** no. 59, 58. MR3068390

[23] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2009). Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.* **37** 815–852. MR2537522

[24] FÉRAL, D. and PÉCHÉ, S. (2007). The largest eigenvalue of rank one deformation of large Wigner matrices. *Comm. Math. Phys.* **272** 185–228. MR2291807

[25] HAAGERUP, U. and THORBJØRNSEN, S. (2005). A new application of random matrices. *Ann. of Math. (2)* **162** 711–775. MR2183281

[26] HORN, R. A. and JOHNSON, C. R. (1985). *Matrix Analysis.* Cambridge Univ. Press, Cambridge. MR0832183

[27] KANTOROVIČ, L. V. (1948). Functional analysis and applied mathematics. *Uspehi Matem. Nauk (N.S.)* **3** 89–185. MR0027947

[28] KARGIN, V. (2012). A concentration inequality and a local law for the sum of two random matrices. *Probab. Theory Related Fields* **154** 677–702. MR3000559

[29] KARGIN, V. (2013). An inequality for the distance between densities of free convolutions. *Ann. Probab.* **41** 3241–3260. MR3127881

[30] KNOWLES, A. and YIN, J. (2014). The outliers of a deformed Wigner matrix. *Ann. Probab.* **42** 1980–2031. MR3262497

[31] KNOWLES, A. and YIN, J. (2013). The isotropic semicircle law and deformation of Wigner matrices. *Comm. Pure Appl. Math.* **66** 1663–1750. MR3103909

[32] MAÏDA, M. (2007). Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles. *Electron. J. Probab.* **12** 1131–1150 (electronic). MR2336602

[33] MALE, C. (2012). The norm of polynomials in large random and deterministic matrices. *Probab. Theory Related Fields* **154** 477–532. With an appendix by Dimitri Shlyakhtenko. MR3000553

[34] NICA, A. and SPEICHER, R. (2006). *Lectures on the Combinatorics of Free Probability. London Mathematical Society Lecture Note Series* **335**. Cambridge Univ. Press, Cambridge. MR2266879

[35] PASTUR, L. and VASILCHUK, V. (2000). On the law of addition of random matrices. *Comm. Math. Phys.* **214** 249–286. MR1796022

[36] PÉCHÉ, S. (2006). The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields* **134** 127–173. MR2221787

[37] PENG, M. (2012). Eigenvalues of deformed random matrices. Available at arXiv:1205.0572.

[38] Pizzo, A., Renfrew, D. and Soshnikov, A. (2013). On finite rank deformations of Wigner matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 64–94. MR3060148

[39] Renfrew, D. and Soshnikov, A. (2013). On finite rank deformations of Wigner matrices II: Delocalized perturbations. *Random Matrices Theory Appl.* **2** 1250015, 36. MR3039820

[40] Shiryaev, A. N. (1996). *Probability*, 2nd ed. *Graduate Texts in Mathematics* **95**. Springer, New York. MR1368405

[41] Speicher, R. (1993). Free convolution and the random sum of matrices. *Publ. Res. Inst. Math. Sci.* **29** 731–744. MR1245015

[42] Tillmann, H.-G. (1953). Randverteilungen analytischer Funktionen und Distributionen. *Math. Z.* **59** 61–83. MR0057345

[43] Voiculescu, D. (1991). Limit laws for random matrices and free products. *Invent. Math.* **104** 201–220. MR1094052

[44] Voiculescu, D. (1993). The analogues of entropy and of Fisher's information measure in free probability theory. I. *Comm. Math. Phys.* **155** 71–92. MR1228526

Statistical Laboratory
Center for Mathematical Sciences
Wilberforce Road
Cambridge CB3 OWB
United Kingdom
E-mail: vladislav.kargin@gmail.com