RANDOM MATRICES: UNIVERSAL PROPERTIES OF EIGENVECTORS

TERENCE TAO AND VAN VU

Abstract. The four moment theorem asserts, roughly speaking, that the joint distribution of a small number of eigenvalues of a Wigner random matrix (when measured at the scale of the mean eigenvalue spacing) depends only on the first four moments of the entries of the matrix. In this paper, we extend the four moment theorem to also cover the coefficients of the eigenvectors of a Wigner random matrix. A similar result (with different hypotheses) has been proved recently by Knowles and Yin, using a different method. As an application, we prove some central limit theorems for these eigenvectors. In another application, we prove a universality result for the resolvent, up to the real axis. This implies universality of the inverse matrix.

1. Introduction

Consider a random Hermitian $n \times n$ matrix $M_n$ with $n$ real eigenvalues (counting multiplicity)

$$\lambda_1(M_n) \leq \ldots \leq \lambda_n(M_n).$$

By the spectral theorem, one can find an orthonormal basis

$$u_1(M_n), \ldots, u_n(M_n) \in S^{2n-1} \subset \mathbb{C}^n$$

of unit eigenvectors of this matrix:

$$M_n u_i(M_n) = \lambda_i(M_n) u_i(M_n),$$

where $S^{2n-1} := \{ z \in \mathbb{C}^n : |z| = 1 \}$ is the unit sphere of $\mathbb{C}^n$, and we view elements of $\mathbb{C}^n$ as column vectors. We write $u_{i,p}(M_n) = \langle u_i(M_n), e_p \rangle$ for the $p^{th}$ coefficient of $u_i(M_n)$ for each $1 \leq i, p \leq n$. If $M_n$ is not Hermitian, but is in fact real symmetric, then we can require the $u_i(M_n)$ to have real coefficients, thus taking values in the unit sphere $S^{n-1} := \{ x \in \mathbb{R}^n : |x| = 1 \}$ of $\mathbb{R}^n$.

Unfortunately, the eigenvectors $u_i(M_n)$ are not unique in either the Hermitian or real symmetric cases; even if one assumes that the spectrum of $M_n$ is simple, in the sense that

$$\lambda_1(M_n) < \ldots < \lambda_n(M_n),$$

T. Tao is supported by NSF Research Award DMS-0649473, the NSF Waterman award and a grant from the MacArthur Foundation.

V. Vu is supported by research grants DMS-0901216 and AFOSAR-FA-9550-09-1-0167.
one has the freedom to rotate each $u_i(M_n)$ by a unit\textsuperscript{1} phase $e^{\sqrt{-1} \theta} \in U(1)$. In the real symmetric case, in which we force the eigenvectors to have real coefficients, one only has the freedom to multiply each $u_i(M_n)$ by a sign $\pm \in O(1)$. However, one can eliminate this $U(1)$ phase ambiguity or $O(1)$ sign ambiguity (in the case of simple spectrum) by adopting a variety of viewpoints:

(i) One can consider the rank one projection operators

$$P_i(M_n) := u_i(M_n) u_i(M_n)^*$$

instead of the eigenvectors $u_i(M_n)$, thus the $pq$ coefficient $P_{i,p,q}(M_n)$ of $P_i(M_n)$ is given by the formula

$$P_{i,p,q}(M_n) = u_{i,p}(M_n) u_{i,q}(M_n).$$

(ii) One can consider the equivalence class $[u_i(M_n)] := \{ e^{\sqrt{-1} \theta} u_i(M_n) : \theta \in \mathbb{R} \} \in S^{2n-1}/U(1)$ (or $\{ \pm u_i(M_n) \} \in S^{n-1}/O(1)$, in the real symmetric case) of all eigenvectors with eigenvalue $\lambda_i$, instead of considering any one individual eigenvector.

(iii) One can perform the ad hoc normalization of requiring $u_{i,p}(M_n)$ to be positive real, where $p$ is the first index for which $u_{i,p}(M_n) \neq 0$ (generically we will have $p = 1$, and as we will see shortly, for Wigner matrices we will also have $p = 1$ with high probability).

(iv) One can perform the random normalization of replacing $u_i(M_n)$ with a randomly chosen rotation $e^{\sqrt{-1} \theta} u_i(M_n)$ (in the Hermitian case) or $\pm u_i(M_n)$ (in the real symmetric case) (each the random phase or sign being chosen independently of each other, and (if $M_n$ is itself random) of $M_n$).

Note that $P_i$, $[u_i(M_n)]$, the ad hoc normalized $u_i(M_n)$, and the randomly normalized $u_i(M_n)$ in viewpoints (i)-(iv) respectively will be uniquely defined as long as the spectrum is simple (indeed, it suffices to have $\lambda_{i-1}(M_n) < \lambda_i(M_n) < \lambda_{i+1}(M_n)$). In the proofs of our main results, we shall adopt viewpoint (i) (which is natural from the perspective of spectral theory). However, in order to express our results in explicit coordinates, we will adopt the more ad hoc viewpoint (iii) or the random viewpoint (iv) in the statements of our results.

Notations. We consider $n$ as an asymptotic parameter tending to infinity. We use $X \ll Y$, $Y \gg X$, $Y = \Omega(X)$, or $X = O(Y)$ to denote the bound $X \leq CY$ for all sufficiently large $n$ and for some constant $C$. Notations such as $X \ll_k Y$, $X = O_k(Y)$ mean that the hidden constant $C$ depend on another constant $k$. $X = o(Y)$ or $Y = \omega(X)$ means that $X/Y \to 0$ as $n \to \infty$; the rate of decay here will be allowed to depend on other parameters.

We will need some definitions that capture the intuition that a certain event $E$ occurs very frequently.

**Definition 1** (Frequent events). Let $E$ be an event depending on $n$.

\textsuperscript{1}We use $\sqrt{-1}$ to denote the imaginary unit, in order to free up the symbol $i$ as an index variable.
The goal of this paper is to understand the distribution of the eigenvectors $u_i(M_n)$ (as normalized using viewpoint (iii) or (iv), for sake of concreteness) of a class of random matrices known as Wigner random matrices.

We say that the Wigner matrix ensemble obeys condition $C_0$ if we have the exponential decay condition

$$P(|\xi_{ij}| \geq t^c) \leq e^{-t}$$

for all $1 \leq i, j \leq n$ and $t \geq C'$, and some constants $C, C'$ (independent of $i, j, n$). We say that the Wigner matrix ensemble obeys condition $C_1$ with constant $C_0$ if

$$E|\xi_{ij}|^{C_0} \leq C$$

for some constant $C$ (independent of $n$).

Of course, Condition $C_0$ implies Condition $C_1$ for any $C_0$, but not conversely.

An important example of a Wigner matrix ensemble is the Gaussian unitary ensemble (GUE), which in our notation corresponds to the case when $\xi_{ij}$ has the complex normal distribution $N(0, 1)_C$ with mean zero and variance one for $i > j$, and $\xi_{ij}$ has the real normal distribution $N(0, 1)_R$ with mean zero and variance one for $i = j$. This ensemble obeys Condition $C_0$, and hence also Condition $C_1$ for any $C_0$. Another important example is the Gaussian orthogonal ensemble (GOE), in which $\xi_{ij}$ has the real normal distribution $N(0, 1)_R$ for $i > j$ and the real normal distribution $N(0, 2)_R$ for $i = j$.

The GOE and GUE are continuous ensembles, and it is therefore not difficult to show that the spectrum of such matrices is almost surely simple. The GUE ensemble is invariant under conjugation by unitary matrices, which implies that each unit eigenvector $u_i(M_n)$ is uniformly distributed on the unit sphere $S^{2n-1}$ of $\mathbb{C}^n$, after quotienting out by the $U(1)$ action $u_i(M_n) \rightarrow e^{i\theta}u_i(M_n)$ (or equivalently, that $[u_i(M_n)]$ is uniformly distributed using the Haar measure on $S^{2n-1}/U(1)$). In particular, $P_i(M_n)$ is uniformly distributed on the space of rank one projections
in $\mathbb{C}^n$. Similarly, for GOE, the unit eigenvector $u_i(M_n)$ is uniformly distributed on the unit sphere of $\mathbb{R}^n$ (after quotienting out by the $O(1)$ action $u_i(M_n) \to \pm u_i(M_n)$). The same argument in fact shows that for GUE, the unitary matrix $(u_1(M_n), \ldots, u_n(M_n))$ is distributed according to Haar measure on the unitary group $U(n)$ (after quotienting out by the left-action of the diagonal unitary matrices $U(1)^n$), and for GOE, the orthogonal matrix $(u_1(M_n), \ldots, u_n(M_n))$ is distributed according to Haar measure on the orthogonal group $O(n)$ (after quotienting out by the left-action of the diagonal sign matrices $O(1)^n$). In particular, if one uses the random normalization (iv), then $(u_1(M_n), \ldots, u_n(M_n))$ will be distributed according to Haar measure on $U(n)$ (in the GUE case) or $O(n)$ (in the GOE case).

The distribution of coefficients of a matrix distributed using the Haar measure on the unitary or orthogonal groups has been studied by many authors [2], [16], [11], [23], [5], [4], [3], [12], [13]. It is known (see [2]) that each coefficient, after multiplication by $\sqrt{n}/n$, is asymptotically complex normal (in the unitary case) or real normal (in the orthogonal case). In fact the same is true for the joint distribution of multiple coefficients:

**Theorem 3.** [12], [13] Let $(u_1, \ldots, u_n)$ be distributed using Haar measure on $O(n)$. For $1 \leq i, p \leq n$, let $u_{i,p}$ be the $p^{th}$ coefficient of $u_{i,p}$. Let $\xi_{i,p} \equiv N(0,1)_{\mathbb{R}}$ for $1 \leq i, p \leq k$ be iid real gaussian random variables.

1. If $k = o(\sqrt{n})$, then $(\sqrt{n}u_{i,p})_{1 \leq i,p \leq k}$ and $(\xi_{i,p})_{1 \leq i,p \leq k}$ differ by $o(1)$ in variation norm. In other words, if $F: \mathbb{R}^{m^2} \to \mathbb{R}$ is any measurable function with $\|F\|_{L^\infty(\mathbb{C}^{k^2})} \leq 1$, then
   \[ |EF((\sqrt{n}u_{i,p})_{1 \leq i,p \leq k}) - EF((\xi_{i,p})_{1 \leq i,p \leq k})| \leq o(1) \]
   where the error is uniform in $F$.
2. If instead we make the weaker assumption that $k = o(n/\log n)$, then it is possible to couple together $u_{i,p}$ and $\xi_{i,p}$ such that $\sup_{1 \leq i,p \leq k} |\sqrt{n}u_{i,p} - \xi_{i,p}|$ converges to zero in probability.

Similar conclusions hold for the unitary group $U(n)$, replacing the real gaussian $N(0,1)_{\mathbb{R}}$ by the complex gaussian $N(0,1)_{\mathbb{C}}$.

In [12] it is also shown that the hypotheses $k = o(\sqrt{n})$, $k = o(n/\log n)$ in the above two results are best possible. Of course, by symmetry, one can replace the top left $k \times k$ minor $(u_{i,p})_{1 \leq i,p \leq k}$ of the orthogonal matrix $(u_1, \ldots, u_n)$ with any other $k \times k$ minor and obtain the same results.

As a corollary of Theorem 3, we obtain an asymptotic for the distribution of eigenvector coefficients of GUE or GOE (normalizing using viewpoint (iii)).

**Corollary 4 (Eigenvector coefficients of GOE and GUE).** Let $M_n$ be drawn from GOE, and let $(u_1(M_n), \ldots, u_n(M_n))$ be the eigenvectors (normalized using viewpoint (iii); this is almost surely well-defined since the spectrum is almost surely simple). For $1 \leq i, p \leq n$, let $\xi_{i,p}$ be independent random variables, with $\xi_{i,p} \equiv N(0,1)_{\mathbb{R}}$ for $p > 1$ and $\xi_{i,p} \equiv |N(0,1)_{\mathbb{R}}|$ for $p = 1$. Let $1 \leq k \leq n$, and let $1 \leq i_1 < \ldots < i_k \leq n$ and $1 \leq p_1 < \ldots < p_k \leq n$ be indices.
(i) If \( k = o(\sqrt{n}) \), then \((\sqrt{n}u_{ia, pb}(M_n))_{1 \leq a, b \leq k}\) and \((\xi_{ia, pb})_{1 \leq a, b \leq k}\) differ by \(o(1)\) in variation norm.

(ii) If instead we make the weaker assumption that \( k = o(n/\log n) \), then it is possible to couple together \( M_n \) and \( \xi_{i,p} \) such that \( \sup_{1 \leq a, b \leq k} |\sqrt{n}u_{ia, pb}(M_n) - \xi_{ia, pb}| \) converges to zero in probability.

Similar conclusions hold for GUE, replacing the real gaussian \( N(0, 1)_{\mathbb{R}} \) by the complex gaussian \( N(0, 1)_{\mathbb{C}} \). If one uses the random normalization (iv) instead of (iii), then the conclusions are the same, except that the absolute values are not present in the definition of \( \xi_{i,1} \).

The main objective of this paper is to develop analogues of Corollary 4 for the more general Wigner ensembles from Definition 2. In particular, we would like to consider ensembles \( M_n \) which are allowed to be discrete instead of continuous.

One immediate difficulty that arises in the discrete setting is that one can now have a non-zero probability that the spectrum is non-simple. However, we have the following gap theorem from [17], [18], [19]:

**Theorem 5 (Gap theorem).** Suppose that \( M_n \) is a Wigner random matrix obeying Condition C1 with a sufficiently large constant \( C_0 \), and let \( c_0 > 0 \) be independent of \( n \). Write \( A_n := \sqrt{n}M_n \) for the rescaled matrix. Then for any \( 1 \leq i < n \), one has

\[
P(|\lambda_{i+1}(A_n) - \lambda_i(A_n)| \leq n^{-c_0}) \leq n^{-c_1}
\]

for all sufficiently large \( n \), where \( c_1 > 0 \) depends only on \( c_0 \).

**Proof.** In the bulk case \( \varepsilon n < i < (1 - \varepsilon)n \) assuming Condition C0, see [17, Theorem 19]. For the extension to the edge case, see [18, Theorem 1.14]. For the relaxation of Condition C0 to Condition C1 (with a sufficiently large \( C_0 \)), see [19, Section 2]. See also [7] for some related level repulsion estimates (assuming some additional regularity hypotheses on \( M_n \)).

Of course, one has \( \lambda_i(A_n) = \sqrt{n}\lambda_i(M_n) \). From the above theorem and the union bound, we see that there is an absolute constant \( c > 0 \) such that if \( m = O(n^c) \) and \( 1 \leq i_1 < \ldots < i_m \leq n \), and \( M_n \) obeys Condition C1 with a sufficiently large \( C_0 \), then with probability \( 1 - O(n^{-c}) \), the eigenvalues \( \lambda_{i_1}(M_n), \ldots, \lambda_{i_m}(M_n) \) will all occur with multiplicity one, so that one can meaningfully normalise the eigenvectors \( u_{i_1}(M_n), \ldots, u_{i_m}(M_n) \) according to any of the viewpoints (i), (ii), (iii), (iv) mentioned previously, outside of an exceptional event of probability \( O(n^{-c}) \). On that exceptional event, we define the orthonormal eigenvector basis \( u_i(M_n) \) (and related objects such as the rank one projection \( P_i(M_n) \)) in some arbitrary (measurable) fashion.

1.1. **Main results.** We now turn to the distribution of the coefficients of the eigenvectors. As the ensembles \( M_n \) may be discrete, the eigenvector coefficients \( u_{i,j}(M_n) \) may be discrete also, and so one does not expect to have convergence in variation
distance any more. Instead, we will consider the weaker notion of vague convergence, which resembles the condition (1), but with $F$ now required to be compactly supported, and either continuous or smooth. We need the following definition:

**Definition 6** (Matching moments). Let $k, l \geq 1$. Two Wigner random matrices $M_n = (\xi_{ij})_{1 \leq i, j \leq n}$ and $M'_n = (\xi'_{ij})_{1 \leq i, j \leq n}$ are said to **match to order $k$$** off the diagonal, and **match to order $l$$** on the diagonal, if one has $\text{E} \Re(\xi_{ij})^a \Im(\xi_{ij})^b = \text{E} \Re(\xi'_{ij})^a \Im(\xi'_{ij})^b$ whenever $a, b \geq 0$ and $1 \leq i \leq j \leq n$ are integers such that $a + b \leq k$ (if $i < j$) or $a + b \leq l$ (if $i = j$).

We can now give our first main result, which partially extends Corollary 4 to other Wigner ensembles:

**Theorem 7** (Eigenvector coefficients of Wigner matrices). There are absolute constants $\delta, C_0 > 0$ such that the following statement holds. Let $M_n$ be drawn from a Wigner ensemble obeying Condition C1 with constant $C_0$ that matches GOE to order 4 off the diagonal and to order 2 on the diagonal (which, in particular, forces $M_n$ to be almost surely real symmetric), and let $(u_1(M_n), \ldots, u_n(M_n))$ be the eigenvectors (normalized using viewpoint (iii)). For $1 \leq i, p \leq n$, let $\xi_{i,p}$ be independent random variables, with $\xi_{i,p} \equiv N(0, 1)^\mathbb{R}$ for $p > 1$ and $\xi_{i,p} \equiv |N(0, 1)|$ for $p = 1$. Let $1 \leq k \leq n$, and let $1 \leq i_1 < \ldots < i_k \leq n$ and $1 \leq p_1 < \ldots < p_k \leq n$ be indices.

(i) **(Vague convergence)** If $k = O(n^\delta)$, then one has

\[
\text{E} F((\sqrt{n} u_{i_1, p_1})_{1 \leq a, b \leq k}) - \text{E} F((\xi_{i_a, p_b})_{1 \leq a, b \leq k}) = o(1)
\]

whenever $F : \mathbb{C}^{k^2} \rightarrow \mathbb{R}$ is a smooth function obeying the bounds

\[
|F(x)| = O(1)
\]

and

\[
|\nabla^j F(x)| = O(n^\delta)
\]

for all $x \in \mathbb{C}^{k^2}$ and $0 \leq j \leq 5$.

(ii) If instead we make the stronger assumption that $k = O(1)$, then it is possible to couple together $M_n$ and $\xi_{i_{a,b}}$ such that $\sup_{1 \leq a, b \leq k} |\sqrt{n} u_{i_a, p_b}(M_n) - \xi_{i_a, p_b}|$ converges to zero in probability. In particular, this implies that $(\sqrt{n} u_{i_a, p_b})_{1 \leq a, b \leq k}$ converges to $(\xi_{i_a, p_b})_{1 \leq a, b \leq k}$ in distribution.

The bounds are uniform in the choice of $i_1, \ldots, i_k, p_1, \ldots, p_k$, and in the choice of $F$, provided that the implied constants in the bounds on $F$ are kept fixed. Similar conclusions hold for GUE, replacing the real gaussian $N(0, 1)^\mathbb{R}$ by the complex gaussian $N(0, 1)^\mathbb{C}$.

If one uses the random normalization (iv) instead of (iii), then the conclusions are the same, except that the absolute values are not present in the definition of $\xi_{i,1}$.

The bound $k = O(1)$ in (ii) can be extended by our method (with some effort) to $k = O(\log^{1/2} n)$, but this still falls far short of the analogous range in Corollary 4. It is reasonable to expect that these bounds are not best possible (particularly if one places some additional regularity hypotheses on the coefficients of $M_n$).
We will deduce Theorem 7 from Corollary 4 by establishing a four moment theorem for eigenvectors, which is the main technical result of this paper:

**Theorem 8 (Four Moment Theorem for eigenvectors).** There is an absolute constant $C_0 > 0$ such that for every sufficiently small constant $c_0 > 0$ there is a constant $\delta > 0$ such that the following holds. Let $M_n, M'_n$ be two Wigner random matrices satisfying $C_1$ with constant $C_0$ that match to order 4 off the diagonal and to order 2 on the diagonal. Let $k \leq n^\delta$, and let $G : \mathbb{R}^k \times \mathbb{C}^k \to \mathbb{R}$ be a smooth function obeying the derivative bounds

$$|\nabla^j G(x)| \leq n^{c_0}$$

for all $0 \leq j \leq 5$ and $x \in \mathbb{R}^k \times \mathbb{C}^k$. Then for any $1 \leq i_1, i_2, \ldots, i_k \leq n$ and $1 \leq p_1, \ldots, p_k, q_1, \ldots, q_k \leq n$, and for $n$ sufficiently large depending on $\varepsilon, \delta, c_0, C_0$ we have

$$|E_G(\Phi(\sqrt{n}M_n)) - E_G(\Phi(\sqrt{n}M'_n))| \leq n^{-c_0}$$

where for any matrix $M$ of size $n$, $\Phi(M) \in \mathbb{R}^k \times \mathbb{C}^k$ is the tuple

$$\Phi(M) := ((\lambda_{i_a}(M))_{1 \leq a \leq k}, (nP_{i_a, p_a, q_a}(M))_{1 \leq a \leq k}).$$

The bounds are uniform in the choice of $i_1, \ldots, i_k, p_1, \ldots, p_k, q_1, \ldots, q_k$.

The deduction of Theorem 7 from Corollary 4 and Theorem 8 is routine and is performed in Section 2.

Theorem 8 is an extension of the four moment theorem for eigenvalues established in [17], [18], [19]. Indeed, the latter theorem is essentially the special case of Theorem 8 in which $k = O(1)$, and $G$ only depends on the first $k$ components $((\sqrt{n}\lambda_{i_a}(M_n))_{1 \leq a \leq k})$ of $\Phi(M_n)$. Unsurprisingly, the proof of Theorem 8 will rely heavily on the machinery developed in [17], [18], [19].

Theorem 8 was announced at the AIM workshop “Random matrices” in December 2010. We have found out that recently a result in the same spirit has been proved by Knowles and Yin [14] using a somewhat different method. Knowles and Yin handled the $k = O(1)$ case with Condition $C_1$ replaced by the stronger Condition $C_0$. Furthermore, they needed control on $k + 5$ derivatives of $G$ rather than just 5 derivatives, and also a level repulsion hypothesis similar to (6), (7) below. On the other hand, their result holds for generalized Wigner matrices.

The need for four matching moments in Theorem 8 is believed to be necessary; see [20]. However, we conjecture that Corollary 4 continues to hold without the matching moment hypothesis.

We can use Theorem 8 (together with other tools) to obtain a four moment theorem for the resolvent (or Green’s function) coefficients

$$\left(\frac{1}{\sqrt{n}}M_n - zI\right)^{-1}_{pq} = \sum_{i=1}^{n} \frac{1}{\lambda_i(A_n) - nz} nP_{i, p, q}(A_n),$$

where $A_n$ is the rescaled matrix $A_n := \sqrt{n}M_n$ (so in particular $u_i(A_n) = u_i(M_n)$ and $P_i(A_n) = P_i(M_n)$).
More precisely, we have

**Theorem 9** (Four moment theorem for resolvents up to the real axis). Let $M_n, M'_n$ be two Wigner random matrices satisfying \( C_0 \) that match to order 4 off the diagonal and to order 2 on the diagonal for some sufficiently large $C_0$. Let $z = E + i\eta$ for some $E \in \mathbb{R}$ and some $\eta > 0$. We assume the level repulsion hypothesis that for any $c > 0$, one has with high probability that

\[
\inf_{1 \leq i \leq n} |\lambda_i(\sqrt{n}M_n) - nz| \geq n^{-c}
\]

and

\[
\inf_{1 \leq i \leq n} |\lambda_i(\sqrt{n}M'_n) - nz| \geq n^{-c}.
\]

Let $1 \leq p, q \leq n$. Then for any smooth function $G: \mathbb{C} \to \mathbb{C}$ obeying the bounds $\nabla^j G(x) = O(1)$ for all $x \in \mathbb{C}$ and $0 \leq j \leq 5$, one has

\[
\mathbb{E}G\left(\frac{1}{\sqrt{n}}M_n - zI\right)_{pq}^{-1} - \mathbb{E}G\left(\frac{1}{\sqrt{n}}M'_n - zI\right)_{pq}^{-1} = O(n^{-c_0})
\]

for some constant $c_0 > 0$ independent of $n$.

We isolate the $z = 0$ case of this theorem as a corollary:

**Corollary 10** (Four moment theorem for the inverse matrix). Under the conditions of Theorem 9, we have

\[
\mathbb{E}G\left(\frac{1}{\sqrt{n}}M_n\right)_{pq}^{-1} - \mathbb{E}G\left(\frac{1}{\sqrt{n}}M'_n\right)_{pq}^{-1} = O(n^{-c_0}).
\]

We prove Theorem 9 in Section 5; it is established by combining Theorem 8 with an eigenvalue rigidity result from [10] and a local semicircle law from [9]. This result generalizes a similar four moment theorem from ([9] ( [9, Theorem 2.3])). Its main strengths as compared against that result are that $\eta$ is allowed to go all the way to zero. In particular, one can take $z$ to be zero, thus giving control of the coefficients of the inverse matrix $M_n^{-1}$. On the other hand, the result in [9, Theorem 2.3] does not require the hypothesis (6), (7), and allows the entries in $M_n$ (or $M'_n$) to have different variances, and the bounds are slightly sharper.

**Remark 11.** The same proof allows one to control the joint distribution of $k$ coefficients of several resolvents with $k = O(n^\delta)$ for some sufficiently small $\delta > 0$, assuming a level repulsion estimate at each energy $z$.

The hypotheses (6), (7) are natural, as when these claims fail one would expect the resolvent to be unusually large. However, in practice such hypotheses are in fact automatic and can thus be omitted. For instance, we have

**Proposition 12** (Sufficient conditions for level repulsion). Let $M_n$ be a Wigner matrix whose entries $\xi_{ij}$ are identically distributed in the off-diagonal $1 \leq i < j \leq n$, and also identically distributed on the diagonal $1 \leq i = j \leq n$. Let $z \in \mathbb{C}$. 

(i) If the real and imaginary parts of the off-diagonal coefficients of $M_n$ are iid and supported on at least three points, and $M_n$ obeys Condition C1 for a sufficiently large $C_0$, then (6) holds.

(ii) If $z = 0$ and $M_n$ obeys Condition C1 for $C_0 = 4$, then (6) holds.

Proof. For (i), see [21]; an earlier result assuming smoothness and decay on the coefficients (but without the requirement of iid real and imaginary parts) was established in [7] (see also [15] for a more refined result). The claim (ii) was recently established (by a rather different method) in [22].

It is in fact likely that (6) and (7) in fact always hold whenever Condition C1 is satisfied for a sufficiently large $C_0$, but we do not attempt to establish this fact here.

As another application, we can use our new results to obtain central limit theorems concerning eigenvectors. Here is a sample result.

**Theorem 13.** Let $M_n$ be a random symmetric matrix obeying hypothesis C1 for a sufficiently large constant $C_0$, which matches the Gaussian Orthogonal Ensemble (GOE) to fourth order. Assume furthermore that the atom distributions of $M_n$ are symmetric (i.e. $\xi_{ij} = -\xi_{ij}$ for all $1 \leq i, j \leq n$), and identically distributed (or more precisely, that the $\xi_{ij}$ for $i > j$ are identically distributed, and the $\xi_{ij}$ for $i = j$ are also identically distributed). Let $i = i_n$ be an index (or more precisely, a sequence of indices) between 1 and $n$, and let $a = a_n \in S^{n-1}$ be a unit vector in $\mathbb{R}^n$ (or more precisely, a sequence of unit vectors). Assume either that

(a) $u_i(M_n)$ is normalized using the procedure (iii), and $a \cdot e_1 = o(1)$; or
(b) $u_i(M_n)$ is normalized using the procedure (iv).

Then $\sqrt{n} u_i(M_n) \cdot a$ tends to $N(0, 1)_{\mathbb{R}}$ in distribution as $n \to \infty$.

Note that if one normalizes $u_i(M_n)$ using procedure (iii), then $u_i(M_n) \cdot e_1 = u_{i,1}(M_n)$ will be non-negative (and thus cannot converge in distribution to $N(0, 1)_{\mathbb{R}}$), which helps explain the presence of the condition $a \cdot e_1$ in condition (a).

As an example to illustrate Theorem 13, we can take $a = a_n := \frac{1}{\sqrt{n}}(1, \ldots, 1) \in S^{n-1}$, and $i := \lfloor n/2 \rfloor$. Then Theorem 13 asserts that the sum of the entries of the middle eigenvector $u_{\lfloor n/2 \rfloor}(M_n)$ (using either normalization (iii) or normalization (iv)) is gaussian in the limit.

We prove Theorem 13 in Section 4 as a consequence of Corollary 4 and a general central limit theorem on averages of approximately independent symmetric random variables (Proposition 25) which may be of independent interest. It should be possible to extend this result to more general ensembles $M_n$ (and to obtain the analogous results for ensembles that match GUE rather than GOE), and to obtain central limit theorems for the joint distribution of several statistics of the form $\sqrt{n} u_i(M_n) \cdot a_n$, but we do not pursue this matter here.
Remark 14. The four moment theorem can be extended to handle the singular values of iid covariance matrices, instead of the eigenvalues of Wigner matrices; see [19]. It is possible to use that extension to establish an analogue of Theorem 8 for the singular values and singular vectors of such matrices, and an analogue of Theorem 9 for the inverses of such matrices (which, as is well known, can be expressed in terms of the singular value decomposition of the matrix). We omit the details.

2. Proof of Theorem 7

We establish the claim in the GOE case only, as the GUE case is similar. We shall also establish the claim just for the normalization (iii), as the normalization (iv) can be treated similarly (or deduced directly from the results for (iii)).

Let \( C_0 \) be as in Theorem 8, let \( \delta > 0 \) be sufficiently small, and let \( M_n, u_1(M_n), \ldots, u_n(M_n), \xi_{i,j}, k, r_1, \ldots, r_k, p_1, \ldots, p_k, \) and \( F \) be as in that theorem. Let \( M'_n \) be drawn from GOE, and write \( A_n := \sqrt{n}M_n \) and \( A'_n := \sqrt{n}M'_n \). We initially assume that \( k = O(n^3) \). By adding a dummy index and relabeling if necessary, we may assume without loss of generality that \( p_1 = 1 \).

The normalization given by viewpoint (iii) is unstable when \( u_{i,1} \) is close to zero, and so we will need to eliminate this possibility first. To this end, let \( F_0 : \mathbb{C}^k \to \mathbb{R} \) be a function of the form
\[
F_0(x_1, \ldots, x_k) := \prod_{a=1}^k (1 - \chi(n^{10\delta} x_a))
\]
where \( \chi : \mathbb{C} \to [0,1] \) is a smooth cutoff which equals one on the ball \( B(0,1) \) but vanishes outside of \( B(0,2) \). Observe from the union bound that
\[
\mathbb{E} F_0 \left( \left( |n \xi_{i,a}|^2 \right)_{1 \leq a \leq k} \right) = 1 - o(1)
\]
and thus by Corollary 4, we have an analogous estimate for the quantities \( nP_{i,a,1,1}(A'_n) = (\sqrt{n}u_{i,a,1}(A'_n))^2) \):
\[
\mathbb{E} F_0 \left( \left( nP_{i,a,1,1}(A'_n) \right)_{1 \leq a \leq k} \right) = 1 - o(1).
\]

Applying Theorem 8 (assuming that \( \delta \) is sufficiently small depending on \( c_0 \), so that the losses of \( n^{O(\delta)} \) coming from bounding the derivatives of \( F_0 \) can be absorbed into the \( n^{-c_0} \) factor) we conclude that
\[
\mathbb{E} F_0 \left( \left( nP_{i,a,1,1}(A_n) \right)_{1 \leq a \leq k} \right) = 1 - o(1),
\]
and thus we have
\[
\inf_{1 \leq a \leq k} |nP_{i,a,1,1}(A_n)| \geq n^{-10\delta}
\]
or equivalently that
\[
\inf_{1 \leq a \leq k} \sqrt{n}u_{i,a,1}(A_n) \geq n^{-5\delta}
\]
asymptotically almost surely.
Now we can prove part (i) of Theorem 7. From (8) one has
\[ \mathbf{E} \left| 1 - F_0 \left( (nP_{a,1,1} (A_n))_{1 \leq a \leq k} \right) \right| = o(1) \]
and thus (by the boundedness of \( F \))
\[ \mathbf{E} \left( 1 - F_0 \left( (nP_{a,1,1} (A_n))_{1 \leq a \leq k} \right) \right) \mathbf{F} \left( (\sqrt{n}u_{i_a,p_b} (A_n))_{1 \leq a,b \leq k} \right) = o(1). \]

A similar argument gives
\[ \mathbf{E} \left( 1 - F_0 \left( (|\xi_{i_a,1}|^2)_{1 \leq a \leq k} \right) \right) \mathbf{F} \left( (|\xi_{i_a,p_b}|)_{1 \leq a,b \leq k} \right) = o(1). \]

Thus, to prove (2), it suffices by the triangle inequality to show that
\[ \mathbf{E} \tilde{F} \left( (\sqrt{n}u_{i_a,p_b} (A_n))_{1 \leq a,b \leq k} \right) - \mathbf{E} \tilde{F} \left( (\xi_{i_a,p_b})_{1 \leq a,b \leq k} \right) = o(1) \]
where \( \tilde{F} : \mathbb{C}^k \to \mathbb{C} \) is the function
\[ \tilde{F} ((x_{a,b})_{1 \leq a,b \leq k}) = F_0 \left( (|x_{a,1}|^2)_{1 \leq a \leq k} \right) \mathbf{F} \left( (x_{a,b})_{1 \leq a,b \leq k} \right). \]

From Corollary 4 we have
\[ \mathbf{E} \tilde{F} \left( (\sqrt{n}u_{i_a,p_b} (A'_n))_{1 \leq a,b \leq k} \right) - \mathbf{E} \tilde{F} \left( (\xi_{i_a,p_b})_{1 \leq a,b \leq k} \right) = o(1) \]
so it suffices to show that
\[ \mathbf{E} \tilde{F} \left( (\sqrt{n}u_{i_a,p_b} (A'_n))_{1 \leq a,b \leq k} \right) - \tilde{F} \left( (\sqrt{n}u_{i_a,p_b} (A_n))_{1 \leq a,b \leq k} \right) = o(1). \]

But observe that in the support of \( \tilde{F} \), \( \sqrt{n}u_{i_a,1} \) is non-zero (and indeed, we have (9)), and the components \( \sqrt{n}u_{i_a,p_b} \) can then be recovered from the projection coefficients
\[ nP_{a,p_b} = nP_{a,p_b} u_{i_a,1} \]
by the formula
\[ \sqrt{n}u_{i_a,p_b} = \frac{nP_{a,p_b}}{\sqrt{nP_{a,1,1}}} \]

Thus we can write
\[ \tilde{F} \left( (\sqrt{n}u_{i_a,p_b} (A'_n))_{1 \leq a,b \leq k} \right) = G \left( (nP_{a,p_b} (A'_n))_{1 \leq a,b \leq k} \right) \]
where
\[ G ((x_{a,b})_{1 \leq a,b \leq k}) := \tilde{F} \left( \frac{x_{a,b}}{\sqrt{n}u_{i_a,1}}_{1 \leq a,b \leq k} \right) \]
with the understanding that the right-hand side vanishes when any of the \( x_{a,1} \) vanish. Similarly for \( A_n \). From construction we can easily verify that
\[ |\nabla^j G(x)| = O(n^{O(d)}) \]
for all \( 0 \leq j \leq 5 \). The claim (i) now follows from Theorem 8 (assuming \( \delta \) sufficiently small depending on \( c_0 \)).

Finally, we prove Claim (ii) of Theorem 7. Assume that \( k = O(1) \). Let \( \varepsilon = \varepsilon(n) = o(1) > 0 \) be a slowly decaying function of \( n \) to be chosen later. Observe that asymptotically almost surely, one has
\[ \sum_{a=1}^k \sum_{b=1}^k |\xi_{i_a,p_b}|^2 = O(1/\varepsilon^2), \]
and thus by (i), we conclude that asymptotically almost surely, we also have

$$\sum_{a=1}^{k} \sum_{b=1}^{k} |\sqrt{n}u_{i_a,p_b}|^2 = O(1/\varepsilon^2),$$

thus the vector \( \vec{u} := (\sqrt{n}u_{i_a,p_b})_{1 \leq a,b \leq k} \) asymptotically almost surely lies in a ball \( B \) of radius \( O(1/\varepsilon) \) in \( \mathbb{C}^k \). Since \( k = O(1) \), we may thus cover this ball by \( O(\varepsilon^{-O(1)}) \) disjoint half-open cubes \( Q_1, \ldots, Q_m \) of sidelength \( \varepsilon \); thus

$$\mathbf{P} \left( \vec{u} \in \bigcup_{i=1}^{m} Q_i \right) = 1 - o(1).$$

Write \( \vec{\xi} := (\xi_{i_a,p_b})_{1 \leq a,b \leq k} \). Applying Claim (i) (approximating the indicator functions \( 1_{Q_i} \) from above and below by smooth functions) and using the continuous nature of \( \vec{\xi} \), we see that

$$\sum_{i=1}^{m} |\mathbf{P}(\vec{u} \in Q_i) - \mathbf{P}(\vec{\xi} \in Q_i)| = o(1)$$

if \( \varepsilon \) is sufficiently slowly decaying in \( n \). From this, we see that we may couple \( \vec{u} \) and \( \vec{\xi} \) together in such a fashion that with probability \( 1 - o(1) \), \( \vec{u} \) and \( \vec{\xi} \) lie in the same cube \( Q_i \), which in particular implies that \( \vec{u} - \vec{\xi} = o(1) \). The claim follows.

**Remark 15.** By using the polynomial decay rate \( O(n^{-c}) \) in the bounds appearing in the proof of Claim (i), it is possible to relax the condition \( k = O(1) \) in Claim (ii) to \( k = o(\log^{1/2} n) \), as this allows us to absorb factors of the form \( O(\exp(O(k^2))) \) arising from the high dimension of the domain \( \mathbb{C}^{k^2} \) into the \( O(n^{-c}) \) factors. We omit the details.

## 3. Proof of four moment theorem

In this section we prove Theorem 8. We will follow closely the arguments from [17], [18], [19].

### 3.1. Heuristic discussion.**

Let us first review the general strategy from [17] for handling the eigenvalues. As in [17], we introduce the normalised matrices

$$A_n := \sqrt{n}M_n, A'_n = \sqrt{n}M'_n$$

(whose mean eigenvalue spacing is comparable to 1). For sake of exposition let us restrict attention to the case \( k = 1 \), thus we wish to show that the expectation \( \mathbf{E}G(\lambda_1(A_n)) \) of the random variable \( G(\lambda_1(A_n)) \) only changes by \( O(n^{-c}) \) if one replaces \( A_n \) with another random matrix \( A'_n \) with moments matching up to fourth order off the diagonal (and up to second order on the diagonal). To further simplify the exposition, let us suppose that the coefficients \( \zeta_{pq} \) of \( A_n \) (or \( A'_n \)) are real-valued rather than complex-valued.

At present, \( A'_n \) differs from \( A_n \) in all \( n^2 \) components. But suppose we make a much milder change to \( A_n \), namely replacing a single entry \( \sqrt{n}\zeta_{pq} \) of \( A_n \) with its counterpart \( \sqrt{n}\zeta'_{pq} \) for some \( 1 \leq p \leq q \leq n \). If \( p \neq q \), one also needs to replace the
companion entry $\sqrt{n}c_{pq} = \sqrt{n}c_{pq}'$ with $\sqrt{n}c_{qp}' = \sqrt{n}c_{pq}$, to maintain the Hermitian property. This creates another random matrix $\tilde{A}_n$ which differs from $A_n$ in at most two entries. Note that $\tilde{A}_n$ continues to obey Condition C1, and has matching moments with either $A_n$ or $A_n'$ up to fourth order off the diagonal, and up to second order on the diagonal.

Suppose that one could show that $\mathbf{E}G(\lambda_i(A_n))$ differed from $\mathbf{E}G(\lambda_i(\tilde{A}_n))$ by at most $n^{-2-\epsilon_0}$ when $p \neq q$ and by at most $n^{-1-\epsilon_0}$ when $p = q$. Then, by applying this swapping procedure once for each pair $1 \leq p \leq q \leq n$ and using the triangle inequality, one would obtain the desired bound $|\mathbf{E}G(\lambda_i(A_n)) - \mathbf{E}G(\lambda_i(A_n'))| = O(n^{-\epsilon_0})$.

Now let us see why we would expect $\mathbf{E}G(\lambda_i(A_n))$ to differ from $\mathbf{E}G(\lambda_i(\tilde{A}_n))$ by such a small amount. For sake of concreteness let us restrict attention to the off-diagonal case $p \neq q$, where we have four matching moments; the diagonal case $p = q$ is similar but one only assumes two matching moments, which is ultimately responsible for the $n^{-1-\epsilon_0}$ error rather than $n^{-2-\epsilon_0}$.

Let us freeze (or condition on) all the entries of $A_n$ except for the $pq$ and $qp$ entries. For any complex number $z$, let $A(z)$ denote the matrix which equals $A_n$ except at the $pq$, $qp$, entries, where it equals $z$ and $\overline{z}$ respectively. (Actually, with our hypotheses, we only need to consider real-valued $z$.) Thus it would suffice to show that

$$E \frac{d}{dz} F(z) = E \bigg( \frac{d}{dz} \bigg) F\bigg(\sqrt{n}\zeta_{pq}\bigg) + O(n^{-2-\epsilon_0})$$

for all (or at least most) choices of the frozen entries of $A_n$, where $F(z) := G(\lambda_i(A(z)))$. (A standard argument allows us to restrict attention to values of $z$ of size $O(n^{1/2+O(1/C_0)})$.)

Suppose we could show the derivative estimates

$$\frac{d}{dz} F(z) = O(n^{-\epsilon_0} + O(\epsilon_0))$$

for $l = 1, 2, 3, 4, 5$. Then by Taylor’s theorem with remainder, we would have

$$F(z) = F(0) + F'(0)z + \ldots + \frac{1}{4!} F^{(4)}(0)z^4 + O(n^{-5+O(\epsilon_0)} + O(1/C_0)) |z|^5$$

and so in particular (using the hypothesis $z = O(n^{1/2+O(1/C_0)})$)

$$F(\sqrt{n}\zeta_{pq}) = F(0) + F'(0)\sqrt{n}\zeta_{pq} + \ldots + \frac{1}{4!} F^{(4)}(0)\sqrt{n} c_{pq}^4 + O(n^{-5/2+O(\epsilon_0)} + O(1/C_0))$$

and similarly for $F(\sqrt{n}\zeta_{pq}')$. Since $n^{-5/2+O(\epsilon_0)} + O(1/C_0) = O(n^{-2-\epsilon_0})$ for $C_0$ and $n$ large enough and $\epsilon_0$ small enough, we thus obtain the claim (10) thanks to the hypothesis that the first four moments of $\zeta_{pq}$ and $\zeta_{pq}'$ match. (Note how this argument barely fails if only three moments are assumed to match, though it is possible that some refinement of this argument might still succeed by exploiting further cancellations in the fourth order term $\frac{1}{4!} F^{(4)}(0)\sqrt{n} c_{pq}^4$.)

We can use exactly this strategy for eigenvectors, replacing $\lambda_i$ by $P_{i,p,q}$ (say). A key point here is that the derivatives of $F$ is computed based on the basic relation
$A_n u_i = \lambda_i u_i$ and the chain rule. Thus, in order to bound the derivatives of $F$ (for the four moment theorem for eigenvalues), we needed to obtain estimates for the derivatives of both $\lambda_i$ and $u_i$. The same estimates can be used to bound the derivatives of $F$ in the proof of the four moment theorem for eigenvectors. In order to make $k$ as large as $n^{O(1)}$, one needs to follow the proof closely and make certain adjustments.

3.2. Formal proof. We now turn to the details. The first step is to truncate away the event that an eigenvalue gap is unexpectedly small. Define

$$Q_i(A_n) := \sum_{j \neq i} \frac{1}{|\lambda_j(A_n) - \lambda_i(A_n)|^2}.$$  

This quantity is usually bounded from above:

**Lemma 16 (Bound on $Q$).** If $c_0 > 0$, then one has

$$Q_i(A_n) \leq n^{c_0}$$

with high probability for all $1 \leq a \leq k$.

**Proof.** See [17, Lemma 49]. Strictly speaking, this lemma assumed Condition C0 and was restricted to the bulk of the spectrum, but this was solely because at the time of writing of that paper, the gap theorem (Theorem 5) was only established in that setting. Inserting Theorem 5 as a replacement for [17, Theorem 19] in the proof of [17, Lemma 49], we obtain the claim. \hfill $\square$

In view of this lemma (and its obvious counterpart for $M_n'$), we can (if $k = O(n^\delta)$ for a sufficiently small $\delta$) deduce the four moment theorem from the following truncated version:

**Theorem 17 (Truncated Four Moment Theorem for eigenvectors).** There is an absolute constant $C_0 > 0$ such that for every sufficiently small constant $c_1 > 0$ there is a constant $\delta > 0$ such that the following holds. Let $M_n, M'_n$ be two Wigner random matrices satisfying $C1$ with constant $C_0$ that match to order 4 off the diagonal and to order 2 on the diagonal. Let $k \leq n^{\delta}$, and let $G : \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{C}^k \to \mathbb{R}$ be a smooth function obeying the derivative bounds

$$|\nabla^j G(x)| \leq n^{c_1}$$

for all $0 \leq j \leq 5$ and $x \in \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{C}^k$, and which is supported on the region

$$\{(x_a)_{1 \leq a \leq k}, (y_a)_{1 \leq a \leq k}, (z_{ab})_{1 \leq a, b \leq k}) : |y_a| \leq n^{c_1} \text{ for all } 1 \leq a \leq k\}.$$  

Then for any $1 \leq i_1, i_2, \ldots, i_k \leq n$ and $1 \leq p_1, \ldots, p_k, q_1, \ldots, q_k \leq n$, and for $n$ sufficiently large depending on $\varepsilon, \delta, c_0$ (and the constant $C$ in Definition 2) we have

$$|\mathbf{E} G(\tilde{\Phi}(A_n)) - \mathbf{E} G(\tilde{\Phi}(A'_n))| \leq n^{-c_1}$$

where for any matrix $A_n, \tilde{\Phi}(A_n) \in \mathbb{R}^k \times \mathbb{C}^k$ is the tuple

$$\tilde{\Phi}(A_n) := ((\lambda_{i_a}(A_n))_{1 \leq a \leq k}, (Q_{i_a}(A_n))_{1 \leq a \leq k}(nP_{i_a}(A_n)_{p_a,q_a})_{1 \leq a \leq k}).$$
The reduction of Theorem 8 to Theorem 17 using Lemma 16 proceeds exactly as in [17, §3.3] and is omitted.

As in [17], we adopt the Lindeberg strategy of swapping each matrix entry of $M_n$ (and its transpose) one at a time. A key definition is that of a good configuration, which is a slightly modified version of the same concept from [17]. Let $\varepsilon_1, C_1$ be parameters (independent of $n$) to be selected later. For a given $n$, we fix $k, i_1, \ldots, i_k, G$ as in Theorem 17.

**Definition 18** (Good configuration). For a complex parameter $z$, let $A(z)$ be a (deterministic) family of $n \times n$ Hermitian matrices of the form

$$A(z) = A(0) + z e_p e_q^* + \overline{z} e_q e_p^*$$

where $e_p, e_q$ are unit vectors. We say that $A(z)$ is a good configuration if for every $1 \leq a \leq k$ and every $|z| \leq n^{1/2+\varepsilon_1}$ whose real and imaginary parts are multiples of $n^{-c_1}$, we have the following properties:

- **(Eigenvalue separation)** For any $1 \leq i \leq n$ with $|i - i_a| \geq n^{\varepsilon_1}$, we have
  \begin{equation}
  |\lambda_i(A(z)) - \lambda_{i_a}(A(z))| \geq n^{-\varepsilon_1} |i - i_a|.
  \end{equation}

- **(Delocalization)** There exists an orthonormal eigenfunction basis $u_1(A(z)), \ldots, u_n(A(z))$ such that
  \begin{equation}
  \sup_{1 \leq i,j \leq n} |u_{i,j}(A(z))| \ll n^{-1/2+\varepsilon_1}.
  \end{equation}

To show Theorem 17, it then suffices to establish the following two claims. The first claim, which we shall establish shortly, is an analogue of [17, Proposition 46], and involves a single (deterministic) good configuration $A(z)$:

**Proposition 19** (Replacement given a good configuration). Suppose that $C_1$ is sufficiently large, and let $\varepsilon_1 > 0$. Let $A(z)$ be a good configuration. Then one has

$$|\mathbb{E}G(\Phi(A(\zeta))) - \mathbb{E}G(\Phi(A(\zeta')))| = O(n^{-(r+1)/2+O(\varepsilon_1)+O(\delta)})$$

whenever $\zeta, \zeta'$ are random complex variables that match to order $r$ for some $r = 2, 3, 4$, and bounded almost surely by $O(n^{1/2+\varepsilon_1})$.

The second claim is an analogue of [17, Proposition 48]:

**Proposition 20** (Good configurations occur very frequently). Let $\varepsilon_1 > 0$ and $C_1 \geq 1$, and assume that $C_0 \geq 1$ is sufficiently large depending on $\varepsilon_1$. Let $A(0) = (\zeta_{ij})_{1 \leq i,j \leq n}$ be a random Hermitian matrix with independent upper-triangular entries and $|\zeta_{ij}| \leq n^{1/2+10/C_0}$ for all $1 \leq i,j \leq n$, with $\zeta_{pq} = \zeta_{qp} = 0$, but with $\zeta_{ij}$ having mean zero and variance $n$ for all other $ij$, and also being distributed continuously in the complex plane. Then $A(0)$ is a good configuration with overwhelming probability.

---

2Note that now that $k$ is as large as $O(n^\delta)$, some factors of $O(n^{O(\delta)})$ may be lost in the bounds, but this can be absorbed by the $n^{-\varepsilon_1}$ gain in the conclusion of Theorem 17.
Proof. This is almost identical to the arguments\(^3\) in [17, §5]. Those arguments already give (14) with overwhelming probability. To obtain the delocalization of eigenvectors, one can use [18, Proposition 1.12] (see also the proof of [17, Corollary 63] to deal with the fact that the \(pq\) and \(qp\) entries are not random). \(\square\)

With these two propositions, we can establish Theorem 17 by first using Condition C1 to truncate to the case when \(M_n, M'_n\) have entries of size \(O(n^{10/C_0})\) (say), and perturbing them to be continuous, and then replacing the entries of \(M_n\) with \(M'_n\) one at a time just as in [17, §3.3]. Because the entries of \(M_n\) and \(M'_n\) match to order 4 off the diagonal and to order 2 on the diagonal, each of the \(O(n^2)\) off-diagonal replacements costs an error of \(O(n^{-5/2+O(\varepsilon_1)+O(\delta)})\), while each of the \(O(n)\) diagonal replacements costs an error of \(O(n^{-5/2+O(\varepsilon_1)+O(\delta)})\), and so the net error is acceptable if \(\varepsilon_1, \delta\) are sufficiently small (and if \(C_0\) is large enough that Proposition 20 applies).

It remains to establish Proposition 19. As in [17], we will need derivative bounds on various spectral statistics of \(A(z)\). For each \(1 \leq i \leq n\), we introduce the resolvent-type matrices

\[
R_i(A) = \sum_{j \neq i} \frac{1}{\lambda_j(A) - \lambda_i(A)} P_j(A).
\]

**Lemma 21.** Let \(1 \leq i \leq n\), and let \(A_0\) be a Hermitian matrix which has a simple eigenvalue at \(\lambda_i(A_0)\). Then \(\lambda_i(A), P_i(A), R_i(A),\) and \(Q_i(A)\) depend smoothly on \(A\) in a neighborhood of \(A_0\).

**Proof.** See [17, Lemma 54]. \(\square\)

Now we turn to more quantitative bounds on derivatives. If \(f(z)\) is a scalar, vector, or matrix-valued function depending smoothly (but not holomorphically) on a complex parameter \(z\), we define the derivatives \(\nabla^m f\) of \(f\) to be the vector (or tensor)-valued quantity

\[
\nabla^m f(z) := \left( \frac{\partial^m f}{\partial \text{Re}(z)^l \partial \text{Im}(z)^{m-l}} \right)_{l=0}^m.
\]

We have a crude bound:

**Lemma 22** (Crude bound). Let \(A = A(z)\) be an \(n \times n\) Hermitian matrix varying (real)-linearly in \(z\) (thus \(\nabla^k A = 0\) for \(k \geq 2\)), with

\[
\|\nabla A\|_{op} \leq V
\]

for some \(V > 0\). Let \(1 \leq i \leq n\). At some fixed value of \(z\), suppose we have the spectral gap condition

\[
|\lambda_j(A(z)) - \lambda_i(A(z))| \geq r
\]

\(\square\)

\(\text{\(3\)The hypotheses in [17, §5] did not have the loss of } n^{10/C_0} \text{ in the bound for } \zeta_{ij}, \text{ but such bounds only cause losses of } O(n^{O(1/C_0)}) \text{ in the final bounds, which can be absorbed into the } n^{\varepsilon_1} \text{ factors in the definition of a good configuration if } C_0 \text{ is sufficiently large depending on } \varepsilon_1.\)
for all \( j \neq i \) and some \( r > 0 \) (in particular, \( \lambda_i(A(z)) \) is a simple eigenvalue). Then for all \( k \geq 1 \) we have (at this fixed choice of \( z \))

\[
|\nabla^k \lambda_i(A(z))| \ll_k V^k r^{1-k}
\]

and

\[
\|\nabla^k P_i(A(z))\|_{\text{op}} \ll_k V^k r^{-k}
\]

and

\[
\|\nabla^k R_i(A(z))\|_{\text{op}} \ll_k V^k r^{-k-1}.
\]

and

\[
|\nabla^k Q_i| \ll_k n V^k r^{-k-2}.
\]

**Proof.** See [17, Lemma 56].

In practice, this crude bound is insufficient, and we will need the following more advanced bound:

**Proposition 23** (Better bound). Let \( A = A(z) \) be an \( n \times n \) matrix depending on a complex parameter \( z \) of the form

\[
A(z) = A(0) + z e_p e_q^* + \Sigma e_q e_p^*
\]

for some vectors \( e_p, e_q \). We abbreviate \( \lambda_i = \lambda_i(A(z)) \), \( P_i = P_i(A(z)) \), etc.

Let \( 1 \leq i \leq n \). At some fixed value of \( z \), suppose that \( \lambda_i = \lambda_i(A(z)) \) is a simple eigenvalue, and that we have a partition

\[
I = P_i + \sum_{\alpha \in J} P_\alpha
\]

where \( J \) is a finite index set, and \( P_\alpha \) are orthogonal projections to invariant spaces on \( A \) (i.e. to spans of eigenvectors not corresponding to \( \lambda_i \)). Suppose that on the range of each \( P_\alpha \), the eigenvalues of \( A - \lambda_i \) have magnitude at least \( r_\alpha \) for some \( r_\alpha > 0 \). Suppose also that we have the incompressibility bounds

\[
\|P_\alpha e_p\|, \|P_\alpha e_q\| \leq w d_\alpha^{1/2}
\]
for all $\alpha \in J \cup \{i\}$ and some $w > 0$ and $d_\alpha \geq 1$, with $d_i := 1$. Then at this value of $z$, and for all $k \geq 1$, we have the bounds

(21) \[ |\nabla^k \lambda_i| \ll_k \left( \sum_{\alpha \in J} \frac{d_\alpha}{r_\alpha} \right)^{k-1} w^{2k} \]

(22) \[ \|P_\alpha(\nabla^k P_i)P_i\|_F \ll_k \left( \sum_{\alpha \in J} \frac{d_\alpha}{r_\alpha} \right)^{k-1} w^{2k} \]

(23) \[ \|P_\alpha(\nabla^k P_i)P_i\|_F = \|P_i(\nabla^k P_\alpha)P_\alpha\|_F \ll_k \left( \sum_{\alpha \in J} \frac{d_\alpha}{r_\alpha} \right)^{k-1} \frac{d^{1/2}_\beta}{r_\alpha r_\beta} \left( \sum_{\alpha \in J} \frac{d_\alpha}{r_\alpha} \right)^{k-2} w^{2k} \]

(24) \[ \|P_\alpha(\nabla^k P_\beta)P_\beta\|_F \ll_k \left( \sum_{\alpha \in J} \frac{d_\alpha}{r_\alpha} \right)^{k+2} w^{2k} \]

(25) \[ |\nabla^k Q_i| \ll_k \left( \sum_{\alpha \in J} \frac{d_\alpha}{r_\alpha} \right)^{k+2} w^{2k} \]

for all $k \geq 0$ and all $\alpha, \beta \in J$ at this value of $z$. Here $\|T\|_F := (\text{trace } TT^*)^{1/2}$ is the Frobenius norm of $T$.

Proof. See [17, Corollary 58], which is a special case of [17, Lemma 57]. The bounds (22), (23), (24) do not appear explicitly in [17, Corollary 58], but come directly from the equations [17, (73), (74), (75)] in [17, Lemma 57] after plugging in the parameters indicated in the proof of [17, Corollary 58]. \hfill \square

We can now prove Proposition 19 and thus Theorem 8. This will be a repetition of the material in [17, Section 4.3]; we sketch the main points here.

Fix $k \geq 1$, $r = 2, 3, 4$ and $\varepsilon_1 > 0$, and suppose that $C_1$ is sufficiently large. We assume $A(0), e_p, e_q, i_1, \ldots, i_k, G, F, \zeta, \zeta'$ are as in the proposition.

We may of course assume that $F(z_0) \neq 0$ for at least one $z_0$ with $|z_0| \leq n^{1/2+\varepsilon_1}$, since the claim is vacuous otherwise.

Using Taylor expansion and the chain rule exactly as in [17, Section 4.3], it suffices to show that

**Lemma 24.** Suppose that $F(z_0) \neq 0$ for at least one $z_0$ with $|z_0| \leq n^{1/2+\varepsilon_1}$. Then for all $z$ with $|z| \leq n^{1/2+\varepsilon_1}$, and all $1 \leq j \leq k$, we have

(26) \[ |\nabla^m \lambda_{ij}(A(z))| \ll n^{O(\varepsilon_1)} n^{-m} \]

and

(27) \[ n|\nabla^m P_{ij,p_j,q_j}(A(z))| \ll n^{O(\varepsilon_1)} n^{-m} \]

and

(28) \[ |\nabla^m Q_{ij}(A(z))| \ll n^{O(\varepsilon_1)} n^{-m} \]

for all $z$ with $|z| \leq n^{1/2+\varepsilon_1}$ and all $0 \leq m \leq 10$. 


The bounds (26), (28) were already proven in [17, Lemma 59]; the arguments in the proof of that lemma also show that
\[ Q_{ij}(A(z)) \ll n^{\varepsilon_1} \]
uniformly for all \( z \) with \( |z| \leq n^{1/2+\varepsilon_1} \).

Now we prove (27). Arguing inductively as in the proof of [17, Lemma 59], it suffices to establish the claim for \( z \) in the ball \( B(z_0, n^{-1-2\varepsilon_1}) \). Let us first establish this for \( z \) whose real and imaginary parts are a multiple of \( n^{-C_1} \). At this value of \( z \), we can apply Proposition 23 exactly as in [17, Section 4.3] to obtain the bounds
\[\| P_{ij}(\nabla^m P_{ij})P_{ij} \|_F \ll n^{-m+O(\varepsilon_1)} \left( \sum_{0 \leq \alpha \leq \log n} \frac{2^\alpha}{r_\alpha} \right)^m \]
(30)
\[\| P_{ij}^{(\alpha)}(\nabla^m P_{ij})P_{ij} \|_F = \| P_{ij}(\nabla^m P_{ij})P_{ij} \|_{P_{ij},\alpha} \| \ll \frac{2^{\alpha/2}}{r_\alpha} n^{-m+O(\varepsilon_1)} \left( \sum_{0 \leq \alpha \leq \log n} \frac{2^\alpha}{r_\alpha} \right)^{m-1} \]
(31)
for all \( 0 \leq m \leq 10, 1 \leq j \leq k \), and \( 0 \leq \alpha, \beta \leq \log n \), where \( r_\alpha \) is the minimal value of \( |\lambda_i - \lambda_j| \) for \( |i - j| \geq 2^\alpha \), and \( P_{ij}^{(\alpha)} \) is the spectral projection to those eigenvalues with \( 2^{\alpha} \leq |i - j| < 2^{\alpha+1} \).

In [17, Section 4.3] it is shown that
\[ \sum_{0 \leq \alpha \leq \log n} \frac{2^\alpha}{r_\alpha} \ll n^{O(\varepsilon_1)} ; \]
in particular
\[ r_\alpha \gg n^{-O(\varepsilon_1)} 2^\alpha . \]
We conclude that
\[ \| P_{ij}^{(\alpha)}(\nabla^m P_{ij})P_{ij}^{(\beta)} \|_F \ll 2^{-\alpha/2} 2^{-\beta/2} n^{-m+O(\varepsilon_1)} \]
(32)
for all \( -1 \leq \alpha, \beta \leq \log n \), where we adopt the convention that \( P_{ij}^{(-1)} := P_{ij} \). Now, we expand
\[ (\nabla^m P_{ij})p_{ij} \varepsilon_{ij} = e_{p_{ij}}(\nabla^m P_{ij})e_{q_{ij}} = \sum_{-1 \leq \alpha, \beta \leq \log n} (e_{p_{ij}} P_{ij}^{(\alpha)})(P_{ij}^{(\alpha)}(\nabla^m P_{ij})P_{ij}^{(\beta)})(P_{ij}^{(\beta)} e_{q_{ij}}) . \]
Applying the triangle and Cauchy-Schwarz inequalities we conclude that
\[ |(\nabla^m P_{ij})p_{ij} \varepsilon_{ij}| \leq \sum_{-1 \leq \alpha, \beta \leq \log n} \| e_{p_{ij}} P_{ij}^{(\alpha)} \| \| P_{ij}^{(\alpha)}(\nabla^m P_{ij})P_{ij}^{(\beta)} \|_F \| P_{ij}^{(\beta)} e_{q_{ij}} \| . \]
From the hypothesis (15) and Pythagoras’ theorem we have
\[ \| e_{p_{ij}} P_{ij}^{(\alpha)} \| \ll n^{\varepsilon_1} 2^{\alpha/2} n^{-1/2} \]
and similarly
\[ \| P_{ij}^{(\beta)} e_{q_{ij}} \| \ll n^{\varepsilon_1} 2^{\beta/2} n^{-1/2} \]
so from (32) we obtain (27) as required, at least when the real and imaginary parts of \( z \) are multiples of \( n^{-C_1} \). The general case can then be handled (for \( C_1 \) large enough) by an appeal to Lemma 22 by arguing exactly as in [17, Section 4.3].

4. Proof of Theorem 13

We now prove Theorem 13. The main tool is the following general central limit theorem:

**Proposition 25** (Central limit theorem). Suppose one has a random unit vector \( u = (u_1, \ldots, u_n) \in \mathbb{R}^n \) such that

1. **(Exchangeability)** The distribution of \( u \) is symmetric with respect to the permutation group \( S_n \).
2. **(For any distinct \( i_1, \ldots, i_k \) with \( k \) fixed, \( \sqrt{n}u_{i_1}, \ldots, \sqrt{n}u_{i_k} \) converges jointly in distribution to \( k \) iid copies of \( N(0, 1) \) as \( n \to \infty \).**
3. **(Symmetry)** The distribution of \( u \) is symmetric with respect to the reflection group \( \{-1, +1\}^n \).

Then for any unit vector \( a \in \mathbb{R}^n \) (depending on \( n \), of course), \( \sqrt{n}a \cdot u \) converges in distribution to \( N(0, 1) \).

Indeed, part (b) of Theorem 13 follows immediately from this proposition and Corollary 4, noting from the symmetries of \( M_n \) that the eigenvector \( u_i(M_n) \) is both symmetric and exchangeable. Part (a) also follows after first reducing to the case \( a \cdot e_1 = 0 \) (noting from Corollary 4 that \( u_i(M_n) \cdot e_1 \) converges in distribution to \( |N(0, 1)_\mathbb{R}| \), so that \( (u_i(M_n) \cdot e_1)(a \cdot e_1) \) converges in distribution to 0 if \( a \cdot e_1 = o(1) \)).

We now prove Proposition 25. Let \( A = A_n \) be a quantity growing slowly to infinity that we will choose later. We truncate \( u_i = u_{i,\leq} + u_{i,>} \)

where \( u_{i,\leq} := u_i |u_i| \leq A/\sqrt{n} \) and \( u_{i,>} := u_i |u_i| > A/\sqrt{n} \). We then split \( u = u_{\leq} + u_\geq \) correspondingly. It will suffice to show that, for a suitable choice of \( A \),

1. \( a \cdot u_{\leq} \) converges in distribution to \( N(0, 1)_\mathbb{R} \), and
2. \( a \cdot u_\geq \) converges in probability to zero.

We first consider the second claim. Here we use the second moment method. It suffices to show that

\[ n \mathbb{E}[a \cdot u_\geq]^2 = o(1). \]

Set \( a = (a_1, \ldots, a_n) \). The left-hand side can be expanded as

\[ n \sum_{i=1}^n a_i \sum_{j=1}^n \mathbb{E}[a_i u_{i,\geq} a_j u_{j,\geq}]. \]

Using the symmetry hypothesis (iv), we see that if \( i \neq j \), then \( u_{i,\geq} u_{j,\geq} \) has a symmetric distribution and thus has mean zero. Thus only the diagonal terms
contribute. Using hypothesis (i) and the unit normalization of \( a \), the second moment becomes

\[ nE u_{1,>}^2. \]

On the other hand, from (i) we have

\[ nE u_{1}^2 = 1 \]

while from (ii) we have

\[ nE u_{1,>}^2 1_{|u_1| \leq K} = E G^2 1_{|G| \leq K} + o(1) \]

for any fixed \( K \), where \( G \equiv N(0, 1) \), and thus (since \( A = A_n \) goes to infinity)

\[
\liminf_{n \to \infty} nE u_{1,>}^2 1_{|u_1| \leq A} \geq \sup_K E G^2 1_{|G| \leq K} = 1.
\]

We conclude that

\[ nE u_{1,>}^2 = 1 - o(1) \]

and

\[ nE u_{1,>}^2 = o(1) \]

and Claim 2 follows.

Now we turn to Claim 1. Here we use the moment method. By Carleman’s theorem (see e.g. [1]), it suffices to show that for each fixed positive integer \( k \),

\[ n^{k/2} E (a \cdot u_\leq)^k = E G^k + o(1). \]

The left-hand side expands as

\[
n^{k/2} \sum_{1 \leq i_1, \ldots, i_k \leq n} a_{i_1} \cdots a_{i_k} E u_{i_1,\leq} \cdots u_{i_k,\leq}.
\]

By the symmetry hypothesis (iii), the expectation vanishes unless each index \( i \) appears an even number of times. Using hypothesis (ii) (and (iii)), we see that

\[ n^{k/2} E u_{i_1,\leq} \cdots u_{i_k,\leq} = E G_{i_1} \cdots G_{i_k} + o(1) \]

where \( G_1, \ldots, G_n \) are iid copies of \( N(0, 1) \), uniformly in \( i_1, \ldots, i_k \), if \( A_n \) grows sufficiently slowly to infinity. Observe that

\[
\sum_{1 \leq i_1, \ldots, i_k \leq n} a_{i_1} \cdots a_{i_k} E G_{i_1} \cdots G_{i_k} = E \left( \sum_{i=1}^n a_i G_i \right)^k = E G^k
\]

since \( \sum_{i=1}^n a_i G_i \equiv G \), and that (as before) the summands vanish unless each index \( i \) appears an even number of times. Thus it suffices to show that

\[ \sum_{*} a_{i_1} \cdots a_{i_k} = O(1) \]

where the sum \( * \) is over all \( k \)-tuples \( 1 \leq i_1, \ldots, i_k \leq n \) in which each index \( i \) appears an even number of times, and the implied constants in the \( O() \) notation are allowed to depend on \( k \).

Suppose there are \( l \) distinct indices appearing, then the contribution of this case is at most

\[ O \left( \sum_{j_1=1}^n \cdots \sum_{j_l=1}^n a_{j_1}^2 \cdots a_{j_l}^2 \right). \]
(since we have \( a_m^r \leq a_r^2 \), whenever \( m \) is a positive even number). But as \( a \) is a unit vector, this sums to \( O(1) \) as required. This concludes the proof of Proposition 25 and hence Theorem 13.

**Remark 26.** An extension of the above argument shows that, under the hypotheses of Proposition 25, if \( a_1, \ldots, a^l \) are a bounded number of orthonormal vectors in \( \mathbb{R}^n \), then \( \sqrt{n}a^1 \cdot u, \ldots, \sqrt{n}a^l \cdot u \) converges jointly in distribution to \( l \) iid copies of \( N(0, 1) \).

5. **Proof of Theorem 9**

We now prove Theorem 9. Let \( M_n, M'_n, C, z, E, \eta, p, q \) be as in that theorem. Let \( c_1 > 0 \) be a small constant to be chosen later, and let \( c_2 > 0 \) be an even smaller constant (depending on \( c_1 \)) to be chosen later. Write \( A_n := \sqrt{n}M_n \) and \( A'_n := \sqrt{n}M'_n \).

Let us call an expression depending on \( M_n \) (or \( A_n \)) stable if it only changes by \( O(n^{-c}) \) for some \( c > 0 \) if \( M_n \) is replaced by \( M'_n \). Our task is thus to show that

\[
\mathbb{E} G \left( \left( \frac{1}{\sqrt{n}} M_n - z I \right)^{-1} \right) \text{ is stable.}
\]

We first subtract off the “global” portion of the resolvent \( \left( \frac{1}{\sqrt{n}} M_n - z I \right)^{-1}_{pq} \). Set \( z_0 := E + in^{-1+c_2/2} \). Applying the local semicircle law from \([9, \text{Theorem } 2.1]\), one has

\[
\left( \frac{1}{\sqrt{n}} M_n - z_0 I \right)_{pq}^{-1} = m_{sc}(z_0)\delta_{pq} + O(n^{-c})
\]

with overwhelming probability for some \( c > 0 \), where \( \delta_{pq} \) is the Kronecker delta function and \( m_{sc}(z_0) \) is semicircular Stieltjes transform

\[
m_{sc}(z_0) := \frac{-z_0 + \sqrt{z_0^2 - 4}}{2}.
\]

This type of result already gives the claim when \( \eta \geq n^{-1+c_2/2} \), so we may assume that \( \eta < n^{-1+c_2/2} \). After shifting \( G \) by \( m_{sc}(z_0)\delta_{pq} \), and using the regularity bounds on \( G \), it thus suffices to show that the expression

\[
\mathbb{E} G \left( \left( \frac{1}{\sqrt{n}} M_n - z I \right)_{pq}^{-1} - \left( \frac{1}{\sqrt{n}} M_n - z_0 I \right)_{pq}^{-1} \right)
\]

is stable.

From (5) we have

\[
\left( \frac{1}{\sqrt{n}} M_n - z I \right)_{pq}^{-1} = \sum_{i=1}^{n} F(\lambda_i(A_n))nP_{i,p,q}(A_n)
\]

---

4Strictly speaking, the statement of \([9, \text{Theorem } 2.1]\) only controls the diagonal component \( p = q \) of the resolvent. However, an inspection of the proof of that theorem (see in particular \([9, (3.13)]\) and \([9, \text{Proposition } 3.3]\)) reveals that the off-diagonal components \( p \neq q \) were also controlled by the argument.
where
\[ F(x) := \frac{1}{x - nz} - \frac{1}{x - n z_0}. \]
Note that \( F \) decays quadratically in \( x \) rather than linearly, due to the subtraction of the comparison term \( \frac{1}{x - n z_0} \); this will allow us to easily neglect the contribution of the spectrum that is far from \( E \) in the rest of the argument.

We now invoke the eigenvalue rigidity result from [10, Theorem 2.2]. Among other things, this theorem gives an interval \([i_-, i_+] \subset [1, n]\) of length \( i_+ - i_- = O(n^{c_2}) \) (depending only on \( n, E, \) and \( c_2 \)) which contains most of the eigenvalues close to \( E \), in the sense that we have
\[
\inf_{i \in [1, n] \setminus [i_-, i_+]} |\lambda_i(A_n) - E| \geq n^{c_2-o(1)}
\]
with overwhelming probability. In fact, we have the stronger assertion that
\[
|\lambda_i(A_n) - E| > n^{-o(1)}(n^{c_2} + \text{dist}(i, [i_-, i_+]))
\]
for all \( i \in [1, n] \setminus [i_-, i_+] \) with overwhelming probability. In particular, this implies that
\[
|F(\lambda_i(A_n))| \leq \frac{n^{c_2/2+o(1)}}{(n^{c_2} + \text{dist}(i, [i_-, i_+]))^2},
\]
and thus
\[
(33) \sum_{i \in [1, n] \setminus [i_-, i_+]} |F(\lambda_i(A_n))| \leq n^{-c_2/2+o(1)}
\]
with overwhelming probability.

Also, by the delocalization of eigenvalues (established in the bulk in [7, Theorem 4.8] (see also the earlier result [6, Theorem 1.2] handling the smooth case), and up to the edge in [18, Proposition 1.12]), we have
\[
|u_{i,p}(A_n)| = O(n^{-1/2+o(1)})
\]
with overwhelming probability for all \( 1 \leq i, p \leq n \), and hence
\[
nP_{i,p,q}(A_n) = O(n^{o(1)})
\]
with overwhelming probability for all \( 1 \leq i, p, q \leq n \). As such, we see that with overwhelming probability, we have
\[
\left( \frac{1}{\sqrt{n}} M_n - z I \right)_{pq}^{-1} - \left( \frac{1}{\sqrt{n}} M_n - z_0 I \right)_{pq}^{-1} = \sum_{i \in [i_-, i_+]} F(\lambda_i(A_n))nP_{i,p,q}(A_n) + O(n^{-c_2/2+o(1)})
\]
with overwhelming probability. Using the regularity bounds on \( G \), it thus suffices to show that the quantity
\[
\mathbb{E}G( \sum_{i \in [i_-, i_+]} F(\lambda_i(A_n))nP_{i,p,q}(A_n) )
\]
is stable.
We let $\chi : \mathbb{C} \to [0,1]$ be a cutoff function supported on $B(0,1)$ that equals 1 on $B(0,1/2)$. Define the truncated version

$$\tilde{F}(x) := F(x)(1 - \chi(n^{c_1}(x-z)))$$

of $F$. From the hypothesis (6), we see that for each $i \in [i_-,i_+]$, one has

$$F(\lambda_i(A_n)) = \tilde{F}(\lambda_i(A_n))$$

with probability at least $1 - O(n^{-10c_2})$ (say), if $c_2$ is sufficiently small depending on $c_1$ and on the implied constant in the high probability event (6). In particular, by the union bound, we have

$$G(\sum_{i \in [i_-,i_+]} F(\lambda_i(A_n))nP_{i,p,q}(A_n)) = G(\sum_{i \in [i_-,i_+]} \tilde{F}(\lambda_i(A_n))nP_{i,p,q}(A_n))$$

with high probability.

Similarly for $A'_n$ using (7). It thus suffices to show that the quantity

$$E_G(\sum_{i \in [i_-,i_+]} \tilde{F}(\lambda_i(A_n))nP_{i,p,q}(A_n))$$

is stable. But this follows directly from Theorem 8 (if $c_1, c_2$ are small enough). The proof of Theorem 9 is now complete.

Remark 27. The exponential decay hypothesis (Condition C0) in Theorem 9 is needed only to be able to use the eigenvalue rigidity result from [10] and the local semicircle law from [9]. It is quite likely that in both cases, one can relax Condition C0 to Condition C1 (conceding some factors of $n^{O(1/C_0)}$ in the process), which would then allow one to achieve a similar relaxation in Theorem 9.

References

[1] Z. D. Bai and J. Silverstein, Spectral analysis of large dimensional random matrices, Mathematics Monograph Series 2, Science Press, Beijing 2006.
[2] E. Borel, Introduction géométrique à quelques théories physiques, Gauthier-Villars, Paris, 1906.
[3] B. Collins, Intégrales Matricielles et Probabilités Non-commutatives. Thèse de Doctorat of Université Paris 6, 2003.
[4] P. Diaconis, M. Eaton, L. Lauritzen, Finite deFinetti theorems in linear models and multivariate analysis, Scand. J. Statist. 19 (1992), 289–315.
[5] P. Diaconis, D. Freedman, A dozen de Finetti-style results in search of a theory. Ann. Inst. Henri Poincaré, 23 (1987), 397-423.
[6] L. Erdős, B. Schlein and H.-T. Yau, Local semicircle law and complete delocalization for Wigner random matrices, Comm. Math. Phys. 287 (2009), no. 2, 641655.
[7] L. Erdős, B. Schlein and H.-T. Yau, Wegner estimate and level repulsion for Wigner random matrices, Int. Math. Res. Notices 2010 (2010), 436–479.
[8] L. Erdős, B. Schlein and H.-T. Yau, Universality of Random Matrices and Local Relaxation Flow, arXiv:0907.5605
[9] L. Erdős, H-T Yau, and J. Yin, Bulk universality for generalized Wigner matrices, arXiv:1001.3453
[10] L. Erdős, H-T Yau, and J. Yin, Rigidity of Eigenvalues of Generalized Wigner Matrices, arXiv:1007.4652
[11] L. Gallardo, Au sujet du contenu probabiliste d’un lemme d’Henri Poincaré, Annales de l’université de clermont, Vol. 69, (1983), 192–197.
[12] T. Jiang (2006). How Many Entries of A Typical Orthogonal Matrix Can Be Approximated By Independent Normals?, *Ann. Probab.* **34**, 1497–1529.

[13] T. Jiang, The Entries of Haar-invariant Matrices from the Classical Compact Groups, *Journal of Theoretical Probability*, **23**, 1227–1243.

[14] A. Knowles, J. Yin, Eigenvector Distribution of Wigner Matrices, *arXiv:1102.0057*

[15] A. Maltsev, B. Schlein, Average Density of States for Hermitian Wigner Matrices, *arXiv:1011.5594*

[16] A. J. Stam, Limit theorems for uniform distributions on high dimensional Euclidean spaces. *J. Appl. Prob.*, 19 (1982), 221–228.

[17] T. Tao and V. Vu, Random matrices: Universality of local eigenvalue statistics, *arXiv:0906.0510.*

[18] T. Tao and V. Vu, Random matrices: Universality of local eigenvalue statistics up to the edge, *Communications in Mathematical Physics*, **298** (2010), 549–572.

[19] T. Tao and V. Vu, Random covariance matrices: Universality of local statistics of eigenvalues, *arXiv:0912.0966.*

[20] T. Tao and V. Vu, Random matrices: Localization of the eigenvalues and the necessity of four moments, *arXiv:1005.2901.*

[21] T. Tao and V. Vu, The Wigner-Dyson-Mehta bulk universality conjecture for Wigner matrices, *arXiv:1101.5707.*

[22] R. Vershynin, *Invertibility of symmetric random matrices, arXiv:1102.0900.*

[23] M. Yor, Inégalités de martingales continus arrêtés à un temps quelconques I. Springer Lecture Notes in Math., No. 1118, 1985.

**Department of Mathematics, UCLA, Los Angeles CA 90095-1555**

_E-mail address: tao@math.ucla.edu_

**Department of Mathematics, Rutgers, New Jersey, NJ 08854**

_E-mail address: vanvu@math.rutgers.edu_