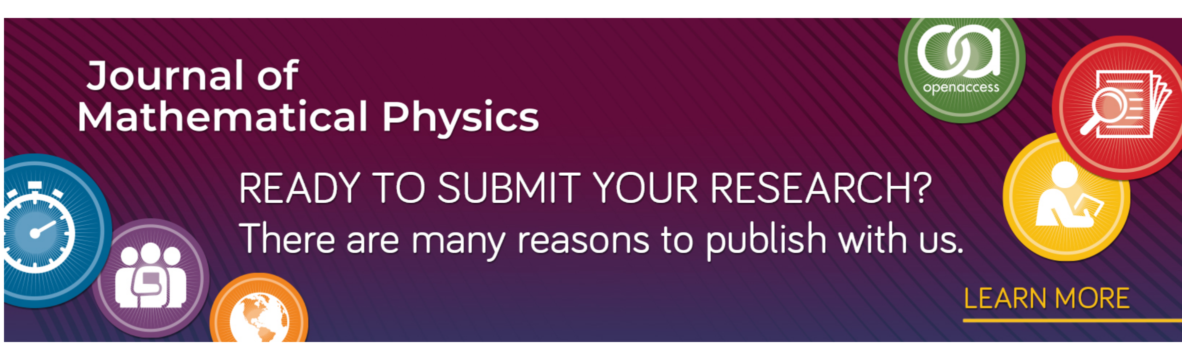
Journal of Mathematical Physics

Sparse general Wigner-type matrices: Local law and eigenvector delocalization

Cite as: J. Math. Phys. **60**, 023301 (2019); https://doi.org/10.1063/1.5053613
Submitted: 23 August 2018 . Accepted: 03 February 2019 . Published Online: 21 February 2019

Ioana Dumitriu, and Yizhe Zhu



Journal of Mathematical Physics ARTICLE scitation.org/journal/jmp

Sparse general Wigner-type matrices: Local law and eigenvector delocalization

Cite as: J. Math. Phys. 60, 023301 (2019); doi: 10.1063/1.5053613
Submitted: 23 August 2018 • Accepted: 3 February 2019 •
Published Online: 21 February 2019

Ioana Dumitriu[a] and Yizhe Zhu[b]

AFFILIATIONS

Department of Mathematics, University of Washington, Seattle, Washington 98195-0005, USA

[a]dumitriu@uw.edu
[b]yizhezhu@uw.edu

ABSTRACT

We prove a local law and eigenvector delocalization for general Wigner-type matrices. Our methods allow us to get the best possible interval length and optimal eigenvector delocalization in the dense case, and the first results of such kind for the sparse case down to $p = \frac{g(n)\log n}{n}$ with $g(n) \to \infty$. We specialize our results to the case of the stochastic block model, and we also obtain a local law for the case when the number of classes is unbounded.

I. INTRODUCTION

A. The stochastic block model

The Stochastic Block Model (SBM), first introduced by mathematical sociologists,[22] is a widely used random graph model for networks with communities. In the last decade, there has been considerable activity[1,2,8–11,25] in understanding the spectral properties of matrices associated with the SBM and to other generalized graph models, in particular, in connection to spectral clustering methods.

Stochastic Block Models represent a generalization of Erdős-Rényi graphs to allow for more heterogeneity. Roughly speaking, an SBM graph starts with a partitioning of the vertices into classes, followed by placing an Erdős-Rényi graph on each class (independent edges, each occurring with the same given probability depending on the class), and connecting vertices in two different blocks by independent edges, again with the same given probability which this time depends on the pair of classes. The random matrix associated with this graph is the adjacency matrix, which is a random block matrix whose entries have Bernoulli distributions, the parameters of which are dictated by the inter- and intra-block probabilities mentioned above.

Specifically, suppose for ease of numbering that $[n] = V_1 \cup V_2 \cup \cdots \cup V_d$ for some integer d, $|V_i| = N_i$ for $i = 1, \ldots, d$. Suppose that for any pair $(k, l) \in [d] \times [d]$ with $k \neq l$ there is a $p_{kl} \in [0, 1]$ such that for any $i \in V_k, j \in V_l$,

$$a_{ij} = \begin{cases} 1, & \text{with probability } p_{kl}, \\ 0, & \text{otherwise.} \end{cases}$$

Also, if $k = l$, there is p_k such that for any $i, j \in V_k$,

$$a_{ij} = \begin{cases} 0, & \text{if } i = j, \\ 1, & \text{with probability } p_k, \\ 0, & \text{otherwise.} \end{cases}$$

Each diagonal block is an adjacency matrix of a simple Erdős-Rényi graph, and off-diagonal blocks are adjacency matrix of bipartite graphs. While there is interest in studying the O(1) variance case [corresponding to all p_{ij}s and p_is being O(1) the "dense" case], special interest is given to the sparse case [when p_{ij}s and p_is are $o(1)$, and more specifically, when the average vertex degrees, given by np_{ij} as well as np_i, are growing very slowly with n or may even be large and constant].

The adjacency matrices of SBM graphs are themselves a particular form of general Wigner-type matrix, which have been shown to exhibit universal properties in Ref. 5 in the dense case. We detail the connection to the broader field of random matrix universality studies in Sec. I B.

B. Universality studies, general Wigner-type matrices, and related graph-based matrix models

At the same time with the increased interest in the spectra of SBM, the universality studies in random matrix theory pioneered by Refs. 30 and 13 had been gaining ground at tremendous pace. The pioneering work on Wigner matrices started in Refs. 30 and 13 has been now extended to cover generalized Wigner matrices,[18] Erdős-Rényi matrices (including sparse ones),[14–16] and general Wigner-type matrices.[5] All such studies start by proving a local law at "optimal" scale, that is, on intervals of length $(\log n)^{\alpha}/n$ or $n^{1-\varepsilon}$, which is necessary for the complicated machinery of either Ref. 30 or Ref. 17 to translate the local law into universality of eigenstatistics on "optimal-length" intervals.

In this paper, we prove a main theorem about (dense) generalized Wigner matrices and then apply it to cover sparse generalized Wigner matrices; finally, we show that our results translate to graph models like the SBM with bounded or unbounded number of blocks. We provide below a brief review of universality studies related to graph-based models.

After the original work on Wigner matrices, the first step in the direction of graph models, or graph-based matrices, came with Ref. 31, where the authors proved a local law for Erdős-Rényi graphs. Subsequently, Refs. 16 and 15 superseeded these results for the slightly denser cases and showed bulk universality in a $p \gg n^{-1/3}$ regime. The sparsity of the model is important here because it makes the problem more difficult. A more recent paper[24] refined the results of Refs. 15 and 16 and made them applicable for $p \gg n^{-1+\varepsilon}$ for any (fixed) $\varepsilon > 0$. Subsequently, in a departure from studying adjacency matrices, Ref. 23 proved bulk universality for the eigenvalue statistics of Laplacian matrices of Erdős-Rényi graphs, in the regime when $p \gg n^{-1+\varepsilon}$ for fixed $\varepsilon > 0$. Very recently, Ref. 21 proved a local law for the adjacency matrix of the Erdős-Rényi graph with $p \geq C \log n/n$ for some constant $C > 0$.

Finally, Ref. 3 examined a large class of sparse random graph-based matrices (two-atom and three-atom entry distributions), proved a local law up to intervals of length $1/n^{1-\varepsilon}$, and deduced (by the same means employed in Ref. 24) a bulk universality theorem. This is different from our results since our sparse matrices have entries that are not necessarily atomic but come from the product of a Bernoulli variable and a potentially continuous one (See Sec. III A); however, the[3] case does seem to cover the sparse SBM model for $p \gg n^{-1+\varepsilon}$. As an interesting aside, in the general case, there may not be an asymptotic global law (aka the limiting empirical spectral distribution); the cases we study here (SBM with bounded and unbounded number of classes) are specific enough that we can also prove the asymptotic global law. However, as it turns out, in the case of the SBM with an unbounded number of blocks, the prediction in the local law must still be made using the n-step approximation to the global law, not the global law itself, since convergence to the global law is not uniform.

Some of the methods used for examining the universality of these graph models rely on the work in Refs. 4 and 5, where a general (dense) Wigner-type matrix model is considered and universality is proved up to intervals of length $1/n^{1-\varepsilon}$. We will also appeal to Ref. 4 since it will help establish the existence of limiting distributions and the stability of their Stieltjes transforms.

We should mention the significant body of the literature that deals with global limits for the empirical spectral distributions of block matrices. Starting with the seminal work of Girko,[20] the topic was treated in Refs. 19 and 29 from a free probability perspective; more recently, Refs. 6 and 12 have examined the topic again for finitely many blocks (a claim in Ref. 12 that the method extends to a growing number of blocks is incorrect). The global law for stochastic block models with a growing number of blocks was derived in Ref. 34 via graphon theory.

C. This paper

The main difference in the results streaming from the seminal studies of Ref. 30, respectively,[13] is in the conditions imposed on the matrix entries: the former approach to universality is based on the "four moment match" condition but imposes relatively weak conditions on the tails, while the latter studies by imposing stronger conditions on the tails. In later studies, these stronger conditions have included bounded moments.[5,18,24] While the methods of Ref. 13 have been extended to increasingly more general matrix models, the methods in Ref. 30 have been used to focus on reaching the best (smallest) possible interval lengths for the classical, Wigner case via methods whose basis was set in Ref. 32.

This paper bridges the two approaches to obtain a local law and eigenvector delocalization for dense and sparse general Wigner-type matrices and for the SBM.

Our main result is a local law in the bulk down to interval length $\frac{CK^2 \log n}{n}$ for general Wigner-type matrices (see Sec. II) whose entries are compactly supported almost surely in $[-K, K]$, employing some of the ideas from Refs. 30 and 32. Our result is more refined than the one from Ref. 5, where the smallest interval length was $O(1/n^{1-\varepsilon})$ and bounded moments were assumed. With additional assumptions (either four-moment matching, as in the case of Ref. 30, or finite moments, as in Ref. 13 and subsequent studies), universality down to this smaller interval length should follow.

In addition to this main result, we also obtain the first local laws for sparse general Wigner-type matrices (see Sec. III A), down to interval length $\frac{CK^2\log n}{np}$. We specialize our results to sparse SBM with finite many blocks, where a limiting law exists. Finally, we extend these results to an unbounded number of blocks for the SBM, under certain conditions (see Sec. III B).

It should be said that our local laws for sparse general Wigner-type matrices are not sharp enough to yield universality, unless p is $\omega(1/n^{\varepsilon})$ for any $\varepsilon > 0$. This is an artifact of the use of the methods of Ref. 30 and is also observable in Ref. 31. It is to be expected that they can be refined (by us or by other researchers) in the near future to a point where universality can be deduced.

II. GENERAL WIGNER-TYPE MATRICES

Let $M_n := (\xi_{ij})_{1\le i,j\le n}$ be a random Hermitian matrix with variance profile $S_n = (s_{ij})_{1\le i,j\le n}$ such that ξ_{ij}, $1 \le i \le j \le n$ are independent with

$$\mathbb{E}\xi_{ij} = 0, \mathbb{E}|\xi_{ij}|^2 = s_{ij}$$

and compactly supported almost surely, i.e., $|\xi_{ij}| \le K$ for some $K = o\left(\sqrt{\frac{n}{\log n}}\right)$.

For the variance profile S_n, we assume

$$c \le s_{ij} \le 1$$

for some constant $c > 0$. Note this is equivalent to $c \le s_{ij} \le C$ by scaling. Define $W_n := \frac{M_n}{\sqrt{n}}$. The Stieltjes transform of the empirical spectral distribution of W_n is given by

$$s_n(z) := \frac{1}{n}\mathrm{tr}(W_n - zI)^{-1}.$$

We will show that $s_n(z)$ can be approximated by the solution of the following quadratic vector equation studied in Ref. 4:

$$m_n(z) = \frac{1}{n}\sum_{k=1}^{n} g_n^{(k)}(z), \tag{1}$$

$$-\frac{1}{g_n^{(k)}(z)} = z + \frac{1}{n}\sum_{l=1}^{n} s_{kl}\, g_n^{(l)}(z), \quad 1 \le k \le n. \tag{2}$$

From Theorem 2.1 in Ref. 4, Eq. (2) has a unique set of solutions $g_n^{(k)}(z) : \mathbb{H} \to \mathbb{H}, 1 \le k \le n$, which are analytic functions on the complex upper half plane $\mathbb{H} := \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$. The unique solution $m_n(z)$ in Eq. (1) is the Stieltjes transform of a probability measure ρ_n with $\mathrm{supp}(\rho_n) \subset [-2, 2]$ such that

$$\rho_n(x) := \lim_{\eta\downarrow 0}\frac{1}{\pi}\mathrm{Im}(m_n(x + i\eta)).$$

We use the following definition for bulk intervals of ρ_n.

Definition II.1. An interval I of a probability density function ρ on $\mathbb{R}$ is a bulk interval if there exists some fixed $\varepsilon > 0$ such that $\rho(x) \ge \varepsilon$, for any $x \in I$.

We obtain the following local law of M_n in the bulk.

Theorem II.2 (Local law in the bulk). *Let M_n be a general Wigner-type matrix and ρ_n be the probability measure corresponding to Eqs. (1) and (2). For any constant δ, C_1, there exists a constant $C_2 > 0$ such that with probability at least $1 - n^{-C_1}$, the following holds. For any bulk interval I of length $|I| \ge \frac{C_2K^2\log n}{n}$, the number of eigenvalues N_I of W_n in I obeys the concentration estimate*

$$\left|N_I - n\int_I \rho_n(x)dx\right| \le \delta n|I|. \tag{3}$$

As a consequence, we obtained an optimal upper bound for eigenvectors that corresponds to eigenvalues of W_n in the bulk interval.

Theorem II.3 (Optimal delocalization of eigenvectors in the bulk). *Let M_n be a general Wigner-type matrix. For any constant $C_1 > 0$ and any bulk interval I such that eigenvalue $\lambda_i(W_n) \in I$, with probability at least $1 - n^{-C_1}$, there is a constant C_2 such that the corresponding unit eigenvector $u_i(W_n)$ satisfies*

$$\|u_i(W_n)\|_\infty \le \frac{C_2 K \log^{1/2} n}{\sqrt{n}}.$$

Remark II.4. Theorems II.2 and II.3 also hold for the general Wigner-type matrix whose entries ξ_{ij}s are sub-gaussian with sub-gaussian norm bounded by K. As mentioned in Remark 4.2 in Ref. 27, the proof follows in the same way by using the inequality in Theorem 2.1 in Ref. 28 for sub-gaussian concentration instead of Lemma 1.2 in Ref. 32 for K-bounded entries.

We use standard methods from Ref. 30, adapted to fit the model considered here.

A. Proof of main results

1. Proof of Theorem II.2

For any $0 < \varepsilon < \frac{1}{2}$ and constant $C_1 > 0$, define a region

$$D_{n,\varepsilon} := \{z \in \mathbb{C} : \rho_n(\mathrm{Re}(z)) \ge \varepsilon, \mathrm{Im}(z) \ge \frac{C_3^2 K^2 \log n}{n\delta^6}\} \tag{4}$$

for some constant $C_3 > 0$ to be decided later.

Let $W_{n,k}$ be the matrix W_n with the kth row and column removed, and a_k be the kth row of W_n with the kth element removed.

Let $(W_n - zI)^{-1} := (q_{ij}^{(n)})_{1\le i,j\le n}$. From Schur's complement lemma (Theorem A.4 in Ref. 7), we have

$$q_{kk} = \frac{1}{-\frac{\xi_{kk}}{\sqrt{n}} - z - Y_k},$$

where

$$Y_k = a_k^*(W_{n,k} - zI)^{-1} a_k.$$

Let

$$f_n^{(k)}(z) := \frac{1}{-\frac{\xi_{kk}}{\sqrt{n}} - z - Y_k}, \tag{5}$$

then we can write $s_n(z)$ as

$$s_n(z) = \frac{1}{n}\mathrm{tr}(W_n - zI)^{-1} = \frac{1}{n}\sum_{k=1}^{n} f_n^{(k)}(z). \tag{6}$$

We first estimate Y_k to derive a perturbed version of (2). Let $(W_{n,k} - zI)^{-1} := (q_{ij}^{(n,k)})_{1\le i,j\le n-1}$, and $S_n^{(k)}$ be a diagonal matrix whose diagonal elements are the kth row of S_n with the kth entry removed. We have

$$\begin{aligned}\mathbb{E}[Y_k | W_{n,k}] &= \mathbb{E}[a_k^*(W_{n,k} - zI)^{-1} a_k | W_{n,k}]\\ &= \sum_{i=1}^{n-1} q_{ii}^{(n,k)} \mathbb{E}|a_{ki}|^2\\ &= \sum_{i=1}^{n-1} q_{ii}^{(n,k)} s_{ki}\\ &= \frac{1}{n}\mathrm{tr}[(W_{n,k} - zI)^{-1} S_n^{(k)}].\end{aligned} \tag{7}$$

The following 2 lemmas give estimates for Y_k, and the proofs are deferred for Secs II B 1 and II B 2.

Lemma II.5. Let $\Sigma_n^{(k)}$ be the diagonal matrix whose diagonal elements are the kth row of S_n. For any k, $1 \le k \le n$, and any fixed z with $\mathrm{Im}(z) \ge \frac{K^2 C_3^2 \log n}{n\delta^6}$,

$$\mathbb{E}[Y_k | W_{n,k}] = \frac{1}{n}\mathrm{tr}[(W_n - zI)^{-1}\Sigma_n^{(k)}] + O\left(\frac{1}{n\eta}\right),$$

where the constant in the $O\left(\frac{1}{n\eta}\right)$ term is independent of z.

A similar estimate holds for Y_k itself.

Lemma II.6. For any constant $C > 0$*, one can choose the constant* C_3 *defined in (4) sufficiently large such that for any* k*,* $1 \le k \le n$*,* $z \in D_{n,\varepsilon}$*, one has*

$$Y_k - \frac{1}{n}\mathrm{tr}[(W_n - zI)^{-1}\Sigma_n^{(k)}] = o(\delta^2) \tag{8}$$

with probability at least $1 - n^{-C-10}$.

With the help of Lemmas II.5 and II.6, note that, since $\frac{|\xi_{kk}|}{\sqrt{n}} = o(\delta^2)$,

$$\frac{1}{n}\mathrm{tr}[(W_n - zI)^{-1}\Sigma_n^{(k)}] = \frac{1}{n}\sum_{l=1}^{n} s_{kl} f_n^{(l)}, \tag{9}$$

and combining (5), (8), and (9), we have

$$f_n^{(k)}(z) + \frac{1}{\frac{1}{n}\sum_{l=1}^{n} s_{kl} f_n^{(l)}(z) + z + o(\delta^2)} = 0, \quad 1 \le k \le n \tag{10}$$

with probability at least $1 - n^{-C-9}$.

The next step involves using the stability analysis of quadratic vector equations provided in Ref. 4 to compare the solutions to (10) and (2). We have the following estimate.

Lemma II.7. For any constant $C > 0$*, one can choose* C_3 *in (4) sufficiently large such that*

$$\sup_{1\le k\le n} |f_n^{(k)}(z) - g_n^{(k)}(z)| = o(\delta^2), \tag{11}$$

for all $z \in D_{n,\varepsilon}$ *uniformly with probability at least* $1 - n^{-C-2}$.

With Lemma II.7, we have for any $C > 0$, there exists $C_3 > 0$ in (4) such that

$$|s_n(z) - m_n(z)| = \left|\frac{1}{n}\sum_{k=1}^{n} f_n^{(k)}(z) - \frac{1}{n}\sum_{k=1}^{n} g_n^{(k)}(z)\right| = o(\delta^2) \tag{12}$$

uniformly for all $z \in D_{n,\varepsilon}$ with probability at least $1 - n^{-C}$.

To complete the Proof of Theorem II.2, we need the following well-known connection between the Stieltjes transform and empirical spectral distribution, as shown, for example, in Lemma 64 in Ref. 30 and also Lemma 4.1 in Ref. 32.

Lemma II.8. Let M_n *be a general Wigner-type matrix. Let* $\varepsilon, \delta > 0$*, for any constant* $C_1 > 0$*, there exists a constant* $C > 0$ *such that suppose that one has the bound*

$$|s_n(z) - m_n(z)| \le \delta$$

with probability at least $1 - n^{-C}$ *uniformly for all* $z \in D_{n,\varepsilon}$*, then for any bulk interval* I *with* $|I| \ge \max\{2\eta, \frac{\eta}{\delta}\log\frac{1}{\delta}\}$ *where* $\eta = \frac{C_3^2 K^2 \log n}{n}$*, one has*

$$\left|N_I - n\int_I \rho_n(x)dx\right| \le \delta n|I|$$

with probability at least $1 - n^{-C_1}$.

From (12), for any constant $C_1 > 0$, we can choose C_3 in (4) large enough such that

$$|s_n(z) - m_n(z)| \le \delta$$

uniformly for all $z \in D_{n,\varepsilon}$ with probability $1 - n^{-C}$, where C is the constant in the assumption of Lemma II.8, and then Theorem II.2 follows from Lemma II.8.

2. Proof of Theorem II.3

The proof is based on Lemma 41 from Ref. 30 given below.

Lemma II.9. Let W_n be a $n \times n$ Hermitian matrix and $W_{n,k}$ be the submatrix of W_n with kth row and column removed, and let $u_i(W_n)$ be a unit eigenvector of W_n corresponding to $\lambda_i(W_n)$, and x_k be the kth coordinate of $u_i(W_n)$. Suppose that none of the eigenvalues of $W_{n,k}$ are equal to $\lambda_i(W_n)$. Let a_k be the kth row of W_n with of kth entry removed, then

$$|x_k|^2 = \frac{1}{1 + \sum_{j=1}^{n-1} (\lambda_j(W_{n,k}) - \lambda_i(W_n))^{-2} |u_j(W_{n,k})^* a_k|^2}, \tag{13}$$

where $u_j(W_{n,k})$ is a unit eigenvector corresponding to $\lambda_j(W_{n,k})$.

Another lemma we need is a weighted projection lemma for random vectors with different variances. It is a slight generalization of Lemma 1.2 in Ref. 32. Note that in the below

$$\mathbb{E}|u_j^* X|^2 = \mathrm{tr}(u_j u_j^* \Sigma),$$

and the proof follows verbatim, as in Ref. 32.

Lemma II.10. Let $X = (\xi_1, \ldots, \xi_n)$ be a K-bounded random vector in $\mathbb{C}^n$ such that $\mathrm{Var}(\xi_i) = \sigma_i^2$, $0 \le \sigma_i^2 \le 1$. Then there are constants $C, C' > 0$ such that the following holds. Let H be a subspace of dimension d with an orthonormal basis $\{u_1, \ldots, u_d\}$, and $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2)$. Then for any $1 \ge r_1, \ldots, r_d \ge 0$,

$$\mathbb{P}\left(\left| \sqrt{\sum_{j=1}^{d} r_j |u_j^* X|^2} - \sqrt{\sum_{j=1}^{d} r_j \mathrm{tr}(u_j u_j^* \Sigma)} \right| \ge t \right) \le C \exp(-C' \frac{t^2}{K^2}). \tag{14}$$

In particular, by squaring, it follows that

$$\mathbb{P}\left(\left| \sum_{j=1}^{d} r_j |u_j^* X|^2 - \sum_{j=1}^{d} r_j \mathrm{tr}(u_j u_j^* \Sigma) \right| \ge 2t \sqrt{\sum_{j=1}^{d} r_j \mathrm{tr}(u_j u_j^* \Sigma)} + t^2 \right) \le C \exp(-C' \frac{t^2}{K^2}). \tag{15}$$

Below we show how delocalization follows from Lemmas II.9, II.10, and Theorem II.2. For any $C_1 > 0$ and any $\lambda_i(W_n)$ in the bulk, by Theorem II.2, one can find an interval I centered at $\lambda_i(W_n)$ and $|I| = \frac{K^2 C_2 \log n}{n}$ for some sufficiently large C_2 such that $N_I \ge \delta_1 n|I|$ for some small $\delta_1 > 0$ with probability at least $1 - n^{-C_1-3}$. By Cauchy interlacing law, we can find a set $J \subset \{1, \ldots, n-1\}$ with $|J| \ge N_I/2$ such that $|\lambda_j(W_{n,k}) - \lambda_i(W_n)| \le |I|$ for all $j \in J$. Let X_k be the kth column of M_n with the kth entry removed. Note that from Lemma II.10, by taking $r_j = 1$, $j \in J$, and $t = C_3 K\sqrt{\log n}$ for some constant $C_3 \ge \frac{C_1+3}{C'}$ in (14), using assumption $s_{ij} \ge c$, we have

$$\begin{aligned}\sqrt{\sum_{j\in J} |u_j(W_{n,k})^* X_k|^2} &\ge \sqrt{\sum_{j\in J} \mathrm{tr}(u_j(W_{n,k}) u_j^*(W_{n,k}) \Sigma)} - C_3 K\sqrt{\log n} \\ &\ge \sqrt{c|J|} - C_3 K \sqrt{\log n} \\ &\ge (\sqrt{c} - \frac{C_3}{\sqrt{C_2 \delta_1/2}})\sqrt{|J|}\end{aligned} \tag{16}$$

with probability at least $1 - n^{-C_1-3}$. By choosing C_2 sufficiently large, (16) implies

$$\sum_{j\in J} |u_j(W_{n,k})^* X_k|^2 \ge C'|J|$$

for some constant $C' > 0$ with probability at least $1 - n^{-C_1-3}$. By (13),

$$
\begin{aligned}
|x_k|^2 &= \frac{1}{1+\sum_{j=1}^{n-1}(\lambda_j(W_{n,k})-\lambda_i(W_n))^{-2}|u_j(W_{n,k})^*\frac{X_k}{\sqrt{n}}|^2}\\
&\le \frac{1}{1+\sum_{j\in J}(\lambda_j(W_{n,k})-\lambda_i(W_n))^{-2}|u_j(W_{n,k})^*\frac{X_k}{\sqrt{n}}|^2}\\
&\le \frac{1}{1+n^{-1}|I|^{-2}\sum_{j\in J}|u_j(W_{n,k})^*X_k|^2}\\
&\le \frac{1}{1+n^{-1}|I|^{-2}C'|J|}\\
&\le \frac{2|I|}{C'\delta_1}\le \frac{K^2C_4^2\log n}{n}
\end{aligned}
$$

for some constant C_4 with probability at least $1-2n^{-C_1-3}$. Thus by taking a union bound, $\|u_i\|_\infty \le \frac{C_4K\sqrt{\log n}}{\sqrt{n}}$ with probability at least $1-n^{-C_1}$ for all $1\le i\le n$.

B. Proof of auxiliary lemmas

We now prove the all the lemmas in the proof of Theorem II.2.

1. Proof of Lemma II.5

Let $\eta := \frac{C_3^2K^2\log n}{n\delta^6}$ and $z := x+\sqrt{-1}\cdot\eta$. By (7), it suffices to show for all $1\le k\le n$,

$$
\left|\mathrm{tr}[(W_n-zI)^{-1}\Sigma_n^{(k)}]-\mathrm{tr}[(W_{n,k}-zI)^{-1}S_n^{(k)}]\right|\le\frac{1}{\eta}. \tag{17}
$$

We will use the following result known as Lemma 1.1 in Chap. 1 of Ref. 20.

Lemma II.11. Let $\vec{c}=(c_1,\ldots,c_n)$ be a real column vector, and $M_n=(\xi_{ij})_{n\times n}$ be a Hermitian matrix, for any z with and $\mathrm{Im}z>0$, we have, for any $1\le k\le n$,

$$
\vec{c}^T(M_n-zI)^{-1}\vec{c}-\vec{c_k}^T(M_{n,k}-zI)^{-1}\vec{c_k}=\frac{c_k^2-\vec{\xi_k}^*R_k(2c_k\vec{c_k})+\vec{\xi_k}^*R_k\vec{c_k}\vec{c_k}^TR_k\vec{\xi_k}}{\xi_{kk}-z-\vec{\xi_k}^*R_k\vec{\xi_k}},
$$

where $R_k=(M_{n,k}-zI)^{-1}$, $\vec{c_k}$ is the vector $\vec{c}$ with the kth coordinate removed, and $\vec{\xi_k}$ is the kth column of M_n with the kth element removed.

We introduce a real random vector $\vec{c}=(c_1,\ldots,c_n)$ whose coordinates are mean zero, independent variables also independent of W_n with $\mathrm{Var}(c_i)=s_{ki}$ for $1\le i\le n$.
Apply Lemma II.11 to W_n and $\vec{c}$. We have $R_k=(W_{n,k}-zI)^{-1}$, and

$$
\vec{c}^T(W_n-zI)^{-1}\vec{c}-\vec{c_k}^T(W_{n,k}-zI)^{-1}\vec{c_k}=\frac{c_k^2-a_k^*R_k(2c_k\vec{c_k})+a_k^*R_k\vec{c_k}\vec{c_k}^TR_ka_k}{\frac{\xi_{kk}}{\sqrt{n}}-z-Y_k}.
$$

By taking the conditional expectation with respect to c, conditioned on W_n, we have

$$
\mathbb{E}[\vec{c}^T(W_n-zI)^{-1}\vec{c}-\vec{c_k}^T(W_{n,k}-zI)^{-1}\vec{c_k}\mid W_n]=\frac{s_{kk}+a_k^*R_kS_n^{(k)}R_ka_k}{\frac{\xi_{kk}}{\sqrt{n}}-z-Y_k}.
$$

Calculating the left-hand side yields

$$
\mathrm{tr}[(W_n-zI)^{-1}\Sigma_n^{(k)}]-\mathrm{tr}[(W_{n,k}-zI)^{-1}S_n^{(k)}]=\frac{s_{kk}+a_k^*R_kS_n^{(k)}R_ka_k}{\frac{\xi_{kk}}{\sqrt{n}}-z-Y_k}.
$$

Since we have

$$
\begin{aligned}
|s_{kk}+a_k^*R_kS_n^{(k)}R_ka_k| &\le 1+|a_k^*R_kS_n^{(k)}R_ka_k|\\
&\le 1+a_k^*((W_{n,k}-xI)^2+\eta^2I)^{-1}a_k,
\end{aligned}
$$

and

$$\mathrm{Im}\left(\frac{\xi_{kk}}{\sqrt{n}} - z - Y_k\right) = -\eta\left(1 + a_k^*((W_{n,k} - xI)^2 + \eta^2 I)^{-1} a_k\right),$$

(17) holds. This completes the proof of Lemma II.5.

2. Proof of Lemma II.6

We need a preliminary bound on the number of eigenvalues in a short interval. The following Lemma is similar to Proposition 66 in Ref. 30.

Lemma II.12. For any constant $C_1 > 0$, there exists a constant $C_2 > 0$ such that for any interval $I \subset \mathbb{R}$ with $|I| \geq \frac{C_2 K^2 \log n}{n}$, one has

$$N_I(W_n) = O(n|I|) \tag{18}$$

with probability at least $1 - n^{-C_1}$.

Proof. By the union bound, it suffices to show that the failure probability for (18) is less than $1 - n^{-C_1 - 1}$ for

$$|I| = \eta := \frac{C_2 K^2 \log n}{n}$$

for some sufficiently large C_2. By

$$\mathrm{Im}(s_n(x + \sqrt{-1}\eta)) = \frac{1}{n}\sum_{i=1}^{n} \frac{\eta}{\eta^2 + (\lambda_i(W_n) - x)^2}, \tag{19}$$

it suffices to show that the event

$$N_I \geq Cn\eta \tag{20}$$

and

$$\mathrm{Im}(s_n(x + \eta\sqrt{-1})) \geq C \tag{21}$$

fails with probability at least $1 - n^{-C_1 - 1}$ for some large absolute constant $C > 1$. Suppose we have (20), (21), by (19),

$$\frac{1}{n}\sum_{k=1}^{n}\left|\mathrm{Im}\left(\frac{1}{\frac{\xi_{kk}}{\sqrt{n}} - (x + \eta\sqrt{-1}) - Y_k}\right)\right| \geq C.$$

Using the bound $\left|\mathrm{Im}\left(\frac{1}{z}\right)\right| \leq \frac{1}{|\mathrm{Im}(z)|}$, it implies

$$\frac{1}{n}\sum_{k=1}^{n}\frac{1}{|\eta + \mathrm{Im}(Y_k)|} \geq C. \tag{22}$$

Note that

$$W_{n,k} = \sum_{j=1}^{n-1} \lambda_j(W_{n,k}) u_j^*(W_{n,k}) u_j(W_{n,k}),$$

where $u_j(W_{n,k})$, $1 \leq j \leq n-1$ are orthonormal basis of $W_{n,k}$, one has

$$Y_k = a_k^*(W_{n,k} - zI)^{-1} a_k = \sum_{j=1}^{n-1} \frac{|u_j^*(W_{n,k}) a_k|^2}{\lambda_j(W_{n,k}) - (x + \eta\sqrt{-1})},$$

and hence

$$\mathrm{Im} Y_k \geq \eta \sum_{j=1}^{n-1} \frac{|u_j^*(W_{n,k}) a_k|^2}{\eta^2 + \left(\lambda_j(W_{n,k} - x\right)^2}.$$

On the other hand, from (20), by Cauchy interlacing theorem, we can find an index set J with $|J| \geq \eta n$ such that $\lambda_j(W_{n,k}) \in I$ for all $j \in J$, then we have

$$\mathrm{Im}(Y_k) \geq \frac{1}{2\eta}\sum_{j \in J} |u_j^*(W_{n,k}) a_k|^2 = \frac{1}{2\eta}\|P_{H_k} a_k\|^2, \tag{23}$$

where P_{H_k} is the orthogonal projection onto a subspace H_k spanned by eigenvectors $u_j(W_{n,k}), j \in J$. From (22) and (23), we have

$$\frac{1}{n}\sum_{k=1}^{n}\frac{2\eta}{2\eta^2+\|P_{H_k}a_k\|^2} \geq C. \tag{24}$$

On the other hand, taking $r_j = 1, 1 \leq j \leq d$, $d = |J|$ and $t = C_4 K\sqrt{\log n}$ for some sufficiently large C_4 in (15), using assumption $s_{ij} \geq c$, we have that $\|P_{H_k}(a_k)\|^2 = \Omega(\eta)$ with probability at least $1 - O(n^{-C'C_4}) \geq 1 - n^{-C_1-5}$. Taking the union bound over all possible choice of J, we have (24) holds with probability at least $1 - n^{-C_1-1}$. The claim then follows by taking C sufficiently large.

Now we are ready to prove Lemma II.6. From Lemma II.5, it suffices to show

$$Y_k - \mathbb{E}[Y_k|W_{n,k}] = o(\delta^2), \quad 1 \leq k \leq n \tag{25}$$

with probability at least $1 - n^{-C-10}$. We can write

$$Y_k = \sum_{j=1}^{n-1}\frac{|u_j^*(W_{n,k})a_k|^2}{\lambda_j(W_{n,k}) - z},$$

where $\{u_j(W_{n,k})\}_{j=1}^{n-1}$ are orthonormal eigenvectors of $W_{n,k}$. Moreover,

$$\begin{aligned}\mathbb{E}[Y_k|W_{n,k}] &= \frac{1}{n}\mathrm{tr}[(W_{n,k} - zI)^{-1}S_n^{(k)}]\\ &= \frac{1}{n}\mathrm{tr}\left[\sum_{j=1}^{n-1}\frac{1}{\lambda_j(W_{n,k}) - z}u_j(W_{n,k})u_j^*(W_{n,k})S_n^{(k)}\right]\\ &= \frac{1}{n}\sum_{j=1}^{n-1}\frac{\mathrm{tr}[u_j(W_{n,k})u_j^*(W_{n,k})S_n^{(k)}]}{\lambda_j(W_{n,k}) - z}.\end{aligned}$$

Let $X_k = \sqrt{n}a_k$, and define

$$t_j := |u_j(W_{n,k})^* X_k|^2 - \mathrm{tr}[u_j(W_{n,k})u_j^*(W_{n,k})S_n^{(k)}].$$

It suffices to show that

$$\left|Y_k - \mathbb{E}[Y_k|W_{n,k}]\right| = \frac{1}{n}\left|\sum_{j=1}^{n-1}\frac{t_j}{\lambda_j(W_{n,k}) - x - \sqrt{-1}\eta}\right| = o(\delta^2)$$

with probability at least $1 - n^{-C-10}$. The remaining part of the proof goes through in the same way as in the Proof of Lemma 5.2 in Ref. 32 with Lemmas II.10 and II.12. Then Lemma II.6 follows. □

3. Proof of Lemma II.7

We define $g_n(z,x) := g_n^{(k)}(z)$ if $x \in [\frac{k-1}{n}, \frac{k}{n}), 1 \leq k \leq n$ and

$$S_n(x,y) := s_{ij} \quad \text{if } x \in [\frac{i-1}{n}, \frac{i}{n}), y \in [\frac{j-1}{n}, \frac{j}{n}). \tag{26}$$

Then (2) can be written as

$$m_n(z) = \int_0^1 g_n(z,x)dx, \tag{27}$$

$$-\frac{1}{g_n(z,x)} = z + \int_0^1 S_n(x,y)g_n(z,y)dy, \tag{28}$$

for all $x \in [0,1]$. Similarly, define $f_n(z,x) := f_n^{(k)}(z)$ if $x \in [\frac{k-1}{n}, \frac{k}{n}), 1 \leq k \leq n$. Then we can write (6) and (10) as

$$\begin{aligned}s_n(z) &= \int_0^1 f_n(z,x)dx,\\ -\frac{1}{f_n(z,x)} &= z + \int_0^1 S_n(x,y)f_n(z,y)dy + d_n(z,x),\end{aligned}$$

where for any fixed z from (10),

$$\|d_n(z)\|_\infty := \sup_{x\in[0,1]} |d_n(z,x)| = o(\delta^2) \tag{29}$$

with probability at least $1 - n^{-C-9}$ for any fixed $z \in D_{n,\varepsilon}$.

The following lemma follows from Theorem 2.12 in Ref. 4 which controls the stability of Eq. (28) in the bulk. Here we use the fact that $c \le s_{ij} \le 1$ to guarantee the assumptions of S_n in Theorem 2.12 in Ref. 4. Define

$$\Lambda(z) := \sup_{x\in[0,1]} |f_n(z,x) - g_n(z,x)|.$$

Lemma II.13. For any fixed $z \in D_{n,\varepsilon}$, there exist constants λ, $C_5 > 0$ depending on ε but independent of n such that for $z \in D_{n,\varepsilon}$,

$$\Lambda(z)\mathbf{1}\{\Lambda(z) \le \lambda\} \le C_5 \|d_n(z)\|_\infty. \tag{30}$$

Proof. Since the variance satisfies $c \le s_{ij} \le 1$, S_n satisfies condition A1-A3 in Chap. 1 of Ref. 4. Especially, it implies condition A3 with $L = 1$.

From the lower bound $\rho_n(z) \ge \varepsilon$, Lemma 5.4 (i) in Ref. 4 implies

$$\sup_{1\le k\le n} |g_n^{(k)}(z)| \le \frac{1}{\varepsilon} < \infty$$

for any z with $\mathrm{Re}(z) \in I$, $\mathrm{Im}(z) > 0$. Then the assumptions in Theorem 2.12 in Ref. 4 holds. (30) then follows from Theorem 2.12 in Ref. 4. □

Remark II.14. Lemma II.13 is a stability result for the solution of (28), which is deterministic and does not require moment assumptions on the random matrix M_n.

From Lemma II.13 and (29), we have for any fixed $z \in D_{n,\varepsilon}$,

$$\Lambda(z)\mathbf{1}\{\Lambda(z) \le \lambda\} = o(\delta^2) \tag{31}$$

with probability at least $1 - n^{-C-9}$.

We proceed with a continuity argument as in the proof of Theorem 3.2 in the bulk (Sec. 3.1 in Ref. 4) to show (31) holds uniformly for $z \in D_{n,\varepsilon}$ with probability at least $1 - n^{-C-2}$.

Now for any $0 < \varepsilon' < \frac{\lambda}{4}$, we consider a line segment

$$L = x + \sqrt{-1}\left[\frac{K^2 C_3^2 \log n}{n\delta^6}, n\right]$$

for some fixed x with $\rho_n(x) \ge \varepsilon$, $0 < \varepsilon < 1/2$, and let n be large enough such that $\frac{1}{n} < \varepsilon'$ and $\|d_n(z)\|_\infty \le \varepsilon'$. Let L_n consist of n^4 evenly spaced points on L. Then we have

$$\Lambda(z)\mathbf{1}\{\Lambda(z) \le \lambda\} \le \varepsilon' \tag{32}$$

for all $z \in L_n$ with probability at least $1 - n^{-C-5}$.

From Theorem 2.1 in Ref. 4, $g_n(z, x)$ is the Stieltjes transform of a probability measure; hence, the derivative of $g_n(z, x)$ is uniformly bounded by $\frac{1}{|\mathrm{Im}(z)|^2} \le n^2$ for $z \in D_{n,\varepsilon}$. Similarly, for $f_n(z, x)$, from (5), for $1 \le k \le n$,

$$\left|\frac{\partial f_n^{(k)}(z)}{\partial z}\right| = \left|\frac{1 + \frac{\partial Y_k}{\partial z}}{(\frac{\xi_{kk}}{\sqrt{n}} - z - Y_k)^2}\right|$$
$$\le \left|\frac{1 + a_k^*(W_{n,k} - zI)^{-2}a_k}{\frac{\xi_{kk}}{\sqrt{n}} - z - Y_k}\right| \frac{1}{|\frac{\xi_{kk}}{\sqrt{n}} - z - Y_k|}.$$

By Theorem A.6. in Ref. 7, for $z = x + \sqrt{-1}\eta$,

$$\left|\frac{1 + a_k^*(W_{n,k} - zI)^{-2}a_k}{(\frac{\xi_{kk}}{\sqrt{n}} - z - Y_k)}\right| \le \frac{1}{\eta},$$

and

$$\left|\frac{\xi_{kk}}{\sqrt{n}}-z-Y_k\right| \geq \left|\mathrm{Im}(\frac{\xi_{kk}}{\sqrt{n}}-z-Y_k)\right| = \eta(1+a_k^*((W_{n,k}-xI)^2+\eta^2 I)^{-1}a_k) \geq \eta.$$

Note that for $z \in D_{n,\varepsilon}$, $\eta \geq \frac{K^2C_3^2\log n}{n\delta^6} \geq \frac{1}{n}$, we get

$$\left|\frac{\partial f_n^{(k)}(z)}{\partial z}\right| \leq \frac{1}{\eta^2} \leq n^2, \quad 1 \leq k \leq n.$$

So both $f_n(z, x)$ and $g_n(z, x)$ are n^2-Lipschitz function in z for $z \in D_{n,\varepsilon}$. It follows that

$$|\Lambda(z') - \Lambda(z)| \leq 2n^2|z' - z|,$$

for any $z, z' \in L$. We first claim that

$$\Lambda(z)\mathbf{1}\{\Lambda(z) \leq \frac{\lambda\varepsilon}{2}\} \leq 2\varepsilon', \tag{33}$$

for all $z \in L$ with probability at least $1 - n^{-C-5}$.

Since $0 < \varepsilon < 1/2$, if $z \in L_n$, (33) is true from (32). If $z \in L \setminus L_n$, choose some $z' \in L_n$ such that $|z - z'| \leq n^{-3}$. Suppose $\Lambda(z) \leq \frac{\lambda\varepsilon}{2}$, note that

$$|\Lambda(z') - \Lambda(z)| \leq 2n^2|z - z'| \leq \frac{2}{n}, \tag{34}$$

which implies

$$\Lambda(z') \leq \Lambda(z) + \frac{2}{n} \leq \frac{\lambda\varepsilon}{2} + \frac{2}{n} \leq \lambda$$

with probability at least $1 - n^{-C-5}$. From (32), $\Lambda(z') \leq \varepsilon'$ with probability at least $1 - n^{-C-5}$. From (34),

$$\Lambda(z) \leq \Lambda(z') + \frac{2}{n} < 2\varepsilon' \tag{35}$$

with probability at least $1 - n^{-C-5}$, therefore (33) holds.

In the next step, we show that the indicator function in (33) is identically equal to 1. From (32), we have $\Lambda(z) \notin (2\varepsilon', \lambda/2)$ with probability at least $1 - n^{-C-5}$.

Let E be the event that $\Lambda(z)\mathbf{1}\{\Lambda(z) \leq \frac{\lambda}{2}\} \leq 2\varepsilon'$ happens. Conditioning on E, since $\Lambda(z)$ is $2n^2$-Lipschitz in z, and L is simply connected, we have

$$\Lambda(L) := \{\Lambda(z) : z \in L\}$$

is simply connected. Therefore $\Lambda(L)$ is contained either in $[0, 2\varepsilon']$ or $[\frac{\lambda}{2}, \infty)$.

From (5), we have for $1 \leq k \leq n$,

$$|f_n^{(k)}(z)| = \frac{1}{\left|-\frac{\xi_{kk}}{\sqrt{n}} - z - Y_k\right|} \leq \frac{1}{|\mathrm{Im}(z)|},$$

and since $g_n^{(k)}(z)$ is a Stieltjes transform of a probability measure, for $1 \leq k \leq n$,

$$|g_n^{(k)}(z)| \leq \frac{1}{|\mathrm{Im}(z)|},$$

which implies

$$\Lambda(z) \leq \frac{2}{|\mathrm{Im}(z)|}.$$

Consider the point $z_n := x + \sqrt{-1} \cdot n \in L$, we have

$$\Lambda(z_n) \leq \frac{2}{\mathrm{Im}(z_n)} = \frac{2}{n} \leq 2\varepsilon',$$

which implies $\Lambda(z_n) \in [0, 2\varepsilon']$. Hence for all $z \in L$, $\Lambda(z) \leq 2\varepsilon'$ with probability at least $1 - n^{-C-5}$, and the indicator function in (33) is identically equal to 1.

Now we extend the estimate to all $z \in D_{n,\varepsilon}$. Consider n^3 lines segments

$$x_k + \sqrt{-1}\left[\frac{K^2C_3^2\log n}{n\delta^6}, n\right], \rho_n(x_k) \geq \varepsilon, 1 \leq k \leq n^3$$

such that the n^2-neighborhoods of points $\{x_k, 1 \leq k \leq n^3\}$ cover any bulk interval of ρ_n. By the $2n^2$-Lipschitz property of $\Lambda(z)$ again, we can show $\Lambda(z) \leq 4\varepsilon'$ for all z with $\rho_n(\text{Re}(z)) > \varepsilon$, $\frac{K^2C_3^2\log n}{n\delta^6} \leq \text{Im}z \leq n$, with probability at least $1 - n^{-C-2}$.

On the other hand, for all z with $\text{Im}(z) > n$,

$$\|f_n(z) - g_n(z)\|_\infty \leq \frac{2}{\text{Im}z} = O\left(\frac{1}{n}\right). \tag{36}$$

Combining these two cases, for all $z \in D_{n,\varepsilon}$ with probability at least $1 - n^{-C-2}$,

$$\|f_n(z) - g_n(z)\|_\infty = o(\delta^2).$$

This completes the Proof of Lemma II.7.

III. APPLICATIONS: SPARSE MATRICES

A. Sparse general Wigner-type matrices

Let M_n be a sparse general Wigner-type matrix with independent entries $M_{ij} = \delta_{ij}\xi_{ij}$ for $1 \leq i \leq j \leq n$. Here δ_{ij} are independent and identically distributed (i.i.d.) Bernoulli random variables which take value 1 with probability $p = \frac{g(n)\log n}{n}$, where $g(n)$ is any function for which $g(n) \to \infty$ as $n \to \infty$, and ξ_{ij} are independent random variables such that

$$\mathbb{E}\xi_{ij} = 0, \mathbb{E}|\xi_{ij}|^2 = s_{ij}, c \leq s_{ij} \leq 1,$$

and in addition, $|\xi_{ij}| \leq K$ almost surely for $K = o(\sqrt{g(n)})$.

We can regard this model as the sparsification of a general Wigner-type matrix by uniform sampling. Similar models were considered in Refs. 26 and 33.

Considering the empirical spectral distribution of $W_n := \frac{M_n}{\sqrt{np}}$, we specify a local law for this model.

Corollary III.1. Let M_n be a sparse general Wigner-type matrix, let ρ_n be the probability measure corresponding to Eqs. (1) and (2). For any constants δ, $C_1 > 0$, there exists a constant $C_2 > 0$ such that with probability at least $1 - n^{-C_1}$, the following holds. For any bulk interval I of length $|I| \geq \frac{C_2K^2\log n}{np}$, the number of eigenvalues N_I of $W_n := \frac{M_n}{\sqrt{np}}$ in I obeys the concentration estimate

$$\left|N_I - n\int_I \rho_n(x)dx\right| \leq \delta n|I|. \tag{37}$$

Proof. Define

$$H_n := \frac{M_n}{\sqrt{p}} = (h_{ij})_{1\leq i,j\leq n}.$$

Then $\mathbb{E}h_{ij} = 0, \mathbb{E}|h_{ij}|^2 = s_{ij}$, and $|h_{ij}| \leq \frac{K}{\sqrt{p}} = o\left(\sqrt{\frac{n}{\log n}}\right)$, and (37) follows as a corollary of Theorem II.2 for H_n. ◻

The infinity norm of eigenvectors in the bulk can be estimated in a similar way.

Corollary III.2. Let M_n be a sparse general Wigner-type matrix and $W_n = \frac{M_n}{\sqrt{np}}$. For any constant $C_1 > 0$ and any bulk interval I such that eigenvalue $\lambda_i(W_n) \in I$, with probability at least $1 - n^{-C_1}$, there is a constant C_2 such that the corresponding unit eigenvector $u_i(W_n)$ satisfies

$$\|u_i(W_n)\|_\infty \leq \frac{C_2K\log^{1/2}n}{\sqrt{np}}.$$

B. Sparse stochastic block models

1. Finite number of classes

Our analysis of sparse random matrices applies to the adjacency matrices of sparse stochastic block models.

Consider the adjacency matrix $A_n = (a_{ij})_{1\le i,j\le n}$ of an SBM graph, where A_n is a random real symmetric block matrix with d^2 blocks. Recall that we partition all indices $[n]$ into d sets,

$$[n] = V_1 \cup V_2 \cup \cdots \cup V_d \tag{38}$$

such that $|V_i| = N_i$. We assume $a_{ii} = 0, 1 \le i \le n$ and $a_{ij}, i \ne j$ are Bernoulli random variables such that if a_{ij} is in the (k, l)th block, $a_{ij} = 1$ with probability p_{kl} and $a_{ij} = 0$ with probability $1 - p_{kl}$.

Let $\sigma_{kl}^2 := p_{kl}(1 - p_{kl})$. Define $p := \max_{kl} p_{kl}$ and $\sigma^2 = p(1 - p)$. Assume

$$p = \frac{g(n)\log n}{n},$$

where $\sup_n p < 1$ and $g(n) \to \infty$ as $n \to \infty$. We also assume that

$$\frac{N_i}{n} = \alpha_i + o\left(\frac{1}{g(n)}\right), \tag{39}$$

$$\frac{\sigma_{kl}^2}{\sigma^2} = c_{kl} + o\left(\frac{1}{g(n)}\right), \tag{40}$$

where $\alpha_i > 0, 1 \le i \le d$ and $c_{kl} \ge c > 0, 1 \le k, l \le d$ for some constant c. The quadratic vector equation becomes

$$m(z) = \sum_{k=1}^{d} \alpha_k g_k(z), \tag{41}$$

$$-\frac{1}{g_k(z)} = z + \sum_{l=1}^{d} \alpha_l c_{kl} g_l(z). \tag{42}$$

We state the following local law for sparse SBM.

Corollary III.3. Let A_n be the adjacency matrix of a stochastic block model with the assumptions above, let ρ be the probability measure corresponding to Eq. (41). For any constant $\delta, C_1 > 0$, there exists a constant $C_2 > 0$ such that with probability at least $1 - n^{-C_1}$, the following holds. For any bulk interval I of length $|I| \ge \frac{C_2 \log n}{np}$, the number of eigenvalues N_I of $\frac{A_n}{\sqrt{n}\sigma}$ in I obeys the concentration estimate

$$\left| N_I - n\int_I \rho(x)dx \right| \le \delta n |I|.$$

Proof. We have the following well-known Cauchy Interlacing Lemma, appearing, for example, as Lemma 36 from Ref. 30.

Lemma III.4. Let A, B be symmetric matrices with the same size, and B has rank 1. Then for any interval I, we have

$$|N_I(A + B) - N_I(B)| \le 1, \tag{43}$$

where $N_I(M)$ is the number of eigenvalues of M in I.

Let $\tilde{A}_n$ be the matrix whose off diagonal entries are equal to A_n and

$$\tilde{a}_{ii} = p_{kk} \tag{44}$$

if (i, i) is in the kth block.

From Lemma III.4, since rank $\mathbb{E}(\tilde{A}_n) = d$, we have

$$|N_I(A_n) - N_I(A_n - \mathbb{E}(\tilde{A}_n))| \le d = o(n|I|).$$

Therefore it suffices to prove the local law for

$$W_n = \frac{A_n - \mathbb{E}\tilde{A}_n}{\sqrt{n}\sigma}.$$

Let $\frac{A_n-\mathbb{E}\tilde{A}_n}{\sigma}=(\xi_{ij})_{1\le i,j\le n}$. By Schur's complement, we can write the Stieltjes transform of the empirical measure $s_n(z)$ in the following way:

$$s_n(z)=\frac{1}{n}\sum_{k=1}^{n}\frac{1}{-\frac{\xi_{kk}}{\sqrt{n}}-z-Y_k}.$$

We do the following partition of $s_n(z)$ into d parts:

$$s_n(z):=\sum_{l=1}^{d}\frac{N_l}{n}f_n^{(l)}(z),$$

where

$$f_n^{(l)}(z):=\frac{1}{N_l}\sum_{k\in V_l}\frac{1}{-\frac{\xi_{kk}}{\sqrt{n}}-z-Y_k}. \tag{45}$$

The kth diagonal element in $\frac{A_n-\mathbb{E}\tilde{A_n}}{\sigma}$ is $\frac{-p_{kk}}{\sqrt{n}\sigma}=o(1)$. Similar with (10), we have

$$-\frac{1}{f_n^{(l)}(z)}=\sum_{m=1}^{d}\frac{N_m}{n}c_{ml}f_n^{(m)}(z)+z+o(1),\quad 1\le l\le d \tag{46}$$

for any $z\in D_{n,\varepsilon}$ with probability at least $1-n^{-C-9}$. Using the assumptions (39), (40) and the fact that $|f_n^{(l)}|\le\frac{1}{\eta}$, we have

$$-\frac{1}{f_n^{(l)}(z)}=z+\sum_{m=1}^{d}\alpha_m c_{ml}f_n^{(m)}(z)+o(1),\quad 1\le l\le d \tag{47}$$

for any fixed $z\in D_{n,\varepsilon}$ with probability at least $1-n^{-C-9}$.

Since d is fixed and all coefficients c_{kl}, $1\le k,l\le d$ in (42) are positive and bounded, from Theorem 2.10 in Ref. 4,

$$\sup_{1\le i\le d}|g_i(z)|<\infty,\forall z\in\mathbb{H}.$$

Theorem 2.12(i) in Ref. 4 implies Lemma II.13 holds with $\Lambda(z):=\sup_{1\le i\le d}|f_n^{(i)}(z)-g_i(z)|$ for any fixed $z\in D_{n,\varepsilon}$. Similar to the Proof of Lemma II.7, we have

$$\begin{aligned}|s_n(z)-m(z)|&=|\sum_{l=1}^{d}\frac{N_l}{n}f_n^{(l)}(z)-\sum_{l=1}^{d}\alpha_l g_l(z)|\\&\le|\sum_{l=1}^{d}\frac{N_l}{n}f_n^{(l)}(z)-\sum_{l=1}^{d}\alpha_l f_n^{(l)}(z)|+|\sum_{l=1}^{d}\alpha_l f_n^{(l)}(z)-\sum_{l=1}^{d}\alpha_l g_l(z)|\\&\le\sum_{l=1}^{d}|(\frac{N_l}{n}-\alpha_l)f_n^{(l)}(z)|+\sum_{l=1}^{d}\alpha_l|f_n^{(l)}(z)-g_l(z)|=o(1)\end{aligned}$$

uniformly for all $z\in D_{n,\varepsilon}$ with probability at least $1-n^{-C}$. Hence the local law for $\frac{A_n}{\sqrt{n}\sigma}$ is proved. □

We have the corresponding infinity norm bound for eigenvectors in the bulk.

Corollary III.5. Let A_n be an adjacency matrix of a stochastic block model. For any bulk interval I such that eigenvalue $\lambda_i(\frac{A_n}{\sqrt{n}\sigma})\in I$ and any constant $C_1>0$, with probability at least $1-n^{-C_1}$, the corresponding unit eigenvector $u_i(\frac{A_n}{\sqrt{n}\sigma})$ satisfies

$$\left\|u_i\left(\frac{A_n}{\sqrt{n}\sigma}\right)\right\|_\infty\le\frac{C_2\sqrt{\log n}}{\sqrt{np}}$$

for some constant $C_2>0$.

Proof. Let $W_n:=\frac{A_n}{\sqrt{n}\sigma}$. For any $\lambda_i(W_n)$ in the bulk, by Corollary III.3, one can find an interval I centered at $\lambda_i(W_n)$ and $|I|=\frac{C_2\log n}{np}$ such that $N_I\ge\delta_1 n|I|$ for some small $\delta_1>0$ with probability at least $1-n^{-C_1-3}$. We can find a set $J\subset\{1,\ldots,n-1\}$ with $|J|\ge N_I/2$ such that $|\lambda_j(W_{n-1})-\lambda_i(W_n)|\le|I|$ for all $j\in J$. Let X_k be the kth column of $\frac{A_n}{\sigma}$ with the kth entry removed, then $X_k=\sqrt{n}a_k$.

Since X_k is not centered, we need to show

$$\sum_{j\in J}|u_j(W_{n,k})^*X_k|^2 = \|\pi_H(X_k)\|^2 = \Omega(|J|) \tag{48}$$

with probability at least $1-n^{-C_1-3}$, where H is the subspace spanned by all orthonormal eigenvectors associated with eigenvalues $\lambda_j(W_{n,k})$, $j\in J$ and $\dim(H)=|J|$.

Let $H_1 = H\cap H_2$, where H_2 is the subspace orthogonal to the vector $\mathbb{E}a_k$. The dimension of H_1 is at least $|J|-1$. Let $b_k = a_k - \mathbb{E}a_k$, then the entries of b_k are centered with the same variances as a_k. By Lemma II.10, we have

$$\|\pi_{H_1}(b_k)\|^2 = \Omega\left(\frac{|J|}{n}\right)$$

with probability at least $1-n^{-C_1-3}$. Moreover,

$$\|\pi_H(a_k)\| = \|\pi_H(b_k+\mathbb{E}a_k)\| \geq \|\pi_{H_1}(b_k+\mathbb{E}a_k)\| = \|\pi_{H_1}(b_k)\|,$$

which implies (48) holds. The rest of the proof follows from the Proof of Theorem II.3. □

2. Unbounded number of classes

For the Stochastic Block Models, if we allow the number of classes $d\to\infty$ as $n\to\infty$, a local law can be proved under the following assumptions:

$$d = o\left(\frac{n}{g(n)}\right), \tag{49}$$

$$\sum_{i=1}^{d}\left|\frac{\sigma_{kl}^2}{\sigma^2} - c_{kl}\right| = o\left(\frac{1}{g(n)}\right). \tag{50}$$

We will compare the Stieltjes transform of the empirical spectral distribution to the measure whose Stieltjes transform satisfies the following equations:

$$m_n(z) = \sum_{i=1}^{d}\frac{N_i}{n}g_{n,i}(z), \tag{51}$$

$$-\frac{1}{g_{n,i}(z)} = z + \sum_{i=1}^{d}\frac{N_i}{n}c_{ij}g_{n,j}(z). \tag{52}$$

We have the following local law for SBM with unbounded number of blocks.

Corollary III.6. Let A_n be an adjacency matrix of SBM with assumptions (49) and (50). Let ρ_n be the probability measure corresponding to Eqs. (51) and (52). For any constants $\delta, C_1>0$, there exists a constant C_2 such that with probability at least $1-n^{-C_1}$ the following holds. For any bulk interval I of length $|I|\geq\frac{C_2\log n}{np}$, the number of eigenvalues N_I of $\frac{A_n}{\sqrt{n}\sigma}$ in I obeys the concentration estimate

$$\left|N_I - n\int_I\rho_n(x)dx\right| \leq \delta n|I|. \tag{53}$$

Proof. Since $d = o\left(\frac{n}{g(n)}\right)$, recall the definition of $\tilde{A}_n$ from (44), by Cauchy interlacing law,

$$|N_I(A_n) - N_I(A_n - \mathbb{E}(\tilde{A}_n))| \leq d = o(n|I|).$$

It suffices to prove the statement for the centered matrix $W_n := \frac{A_n - \mathbb{E}\tilde{A}_n}{\sqrt{n}\sigma}$. The proof then follows from Corollary III.3 with assumption (50). □

Remark III.7. Different from Corollary III.3, in Corollary III.6, we are not comparing the empirical spectral distribution to a limiting spectral distribution ρ independent of n. If we assume $\frac{N_i}{n}\to\alpha_i, \alpha_1\geq\alpha_2\cdots\geq\cdots$, and $\sum_{i=1}^{\infty}\alpha_i = 1$, one can show that ρ_n converge to some ρ (see Sec. 7 in Ref. 34 for further details). But we do not have a local law comparing N_I with $n\int_I\rho(x)dx$. In fact, let S_n be the symmetric function on $[0,1]^2$ representing the variance profile as in (26) and S be its point-wise limit, there is no upper bound for rate of convergence on $\sup_{x,y}|S_n(x,y) - S(x,y)|$.

Remark III.8. With the same argument in the Proof of Corollary III.5, the infinity norm bound for eigenvectors in Corollary III.5 still holds for the SBM with unbounded number of classes.

ACKNOWLEDGMENTS

The authors would like to thank PCMI Summer Session 2017 on Random Matrices, during which a part of this work was performed. This work was supported by NSF Grant No. DMS-1712630.

REFERENCES

[1] E. Abbe and C. Sandon, "Community detection in general stochastic block models: Fundamental limits and efficient algorithms for recovery," in *2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (FOCS) (IEEE, 2015), pp. 670–688.
[2] E. Abbe, A. S. Bandeira, and G. Hall, "Exact recovery in the stochastic block model," IEEE Trans. Inf. Theory **62**(1), 471–487 (2016).
[3] B. Adlam and Z. Che, "Spectral statistics of sparse random graphs with a general degree distribution," e-print arXiv:1509.03368 (2015).
[4] O. H. Ajanki, L. Erdős, and T. Krüger, "Quadratic vector equations on complex upper half-plane," Mem. Amer. Math. Soc. (to appear); preprint arXiv:1506.05095 (2015).
[5] O. H. Ajanki, L. Erdős, and T. Krüger, "Universality for general Wigner-type matrices," Probab. Theory Relat. Fields **169**(3-4), 667–727 (2015).
[6] K. Avrachenkov, L. Cottatellucci, and A. Kadavankandy, "Spectral properties of random matrices for stochastic block model," in *2015 13th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)* (IEEE, 2015), pp. 537–544.
[7] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices* (Springer, 2010), Vol. 20.
[8] F. Benaych-Georges, C. Bordenave, and A. Knowles, "Spectral radii of sparse random matrices," preprint arXiv:1704.02945 (2017).
[9] F. Benaych-Georges, C. Bordenave, and A. Knowles, "Largest eigenvalues of sparse inhomogeneous Erdős-Rényi graphs," Ann. Prob. (to appear); preprint arXiv:1704.02953 (2017).
[10] G. Brito, I. Dumitriu, S. Ganguly, C. Hoffman, and L. V. Tran, "Recovery and rigidity in a regular stochastic block model," in *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms* (Society for Industrial and Applied Mathematics, 2016), pp. 1589–1601.
[11] A. Coja-Oghlan, "Graph partitioning via adaptive spectral techniques," Combinatorics, Probab. Comput. **19**(02), 227–284 (2010).
[12] X. Ding, "Spectral analysis of large block random matrices with rectangular blocks," Lith. Math. J. **54**(2), 115–126 (2014).
[13] L. Erdős, B. Schlein, and H.-T. Yau, "Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices," Ann. Probab. **37**(3), 815–852 (2009).
[14] L. Erdős, H.-T. Yau, and J. Yin, "Universality for generalized Wigner matrices with Bernoulli distribution," J. Combinatorics **2**(1), 15–81 (2011).
[15] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin, "Spectral statistics of Erdős-Rényi graphs II: Eigenvalue spacing and the extreme eigenvalues," Commun. Math. Phys. **314**(3), 587–640 (2012).
[16] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin, "Spectral statistics of Erdős-Rényi graphs I: Local semicircle law," Ann. Probab. **41**(3B), 2279–2375 (2013).
[17] L. Erdős, S. Péché, J. A. Ramírez, B. Schlein, and H.-T. Yau, "Bulk universality for Wigner matrices," Commun. Pure Appl. Math. **63**(7), 895–925 (2010).
[18] L. Erdős, H.-T. Yau, and J. Yin, "Bulk universality for generalized Wigner matrices," Probab. Theory Relat. Fields **154**(1-2), 341–407 (2012).
[19] R. R. Far, T. Oraby, W. Bryc, and R. Speicher, "Spectra of large block matrices," preprint arXiv:cs/0610045 (2006).
[20] V. L. Girko, *Theory of Stochastic Canonical Equations* (Springer Science & Business Media, 2001), Vol. 2.
[21] Y. He, A. Knowles, and M. Marcozzi, "Local law and complete eigenvector delocalization for supercritical Erdős-Rényi graphs," Ann. Prob. (to appear); preprint arXiv:1808.09437 (2018).
[22] P. W. Holland, K. B. Laskey, and S. Leinhardt, "Stochastic blockmodels: First steps," Soc. Networks **5**(2), 109–137 (1983).
[23] J. Huang and B. Landon, "Spectral statistics of sparse Erdős-Rényi graph Laplacians," Ann. Inst. H. Poincare Probab. Statist. (to appear); e-print arXiv:1510.06390v1 (2015).
[24] J. Huang, B. Landon, and H.-T. Yau, "Bulk universality of sparse random matrices," J. Math. Phys. **56**(12), 123301 (2015).
[25] F. Krzakala, C. Moore, E. Mossel, N. Joe, S. Allan, L. Zdeborová, and P. Zhang, "Spectral redemption in clustering sparse networks," Proc. Natl. Acad. Sci. U. S. A. **110**(52), 20935–20940 (2013).
[26] K. Luh and V. Vu, "Sparse random matrices have simple spectrum," preprint arXiv:1802.03662 (2018).
[27] S. O'Rourke, V. Vu, and K. Wang, "Eigenvectors of random matrices: A survey," J. Comb. Theory, Ser. A **144**, 361–442 (2016).
[28] M. Rudelson, R. Vershynin, and R. Vershynin, "Hanson-Wright inequality and sub-Gaussian concentration," Electron. Commun. Probab. **18**(82), 1–9 (2013).
[29] D. Shlyakhtenko, "Gaussian random band matrices and operator-valued free probability theory," Banach Cent. Publ. **43**(1), 359–368 (1998).
[30] T. Tao and V. Vu, "Random matrices: Universality of local eigenvalue statistics," Acta Math. **206**(1), 127–204 (2011).
[31] L. V. Tran, V. H. Vu, and K. Wang, "Sparse random graphs: Eigenvalues and eigenvectors," Random Struct. Algorithms **42**(1), 110–134 (2013).
[32] V. Vu and K. Wang, "Random weighted projections, random quadratic forms and random eigenvectors," Random Struct. Algorithms **47**(4), 792–821 (2015).
[33] P. M. Wood, "Universality and the circular law for sparse random matrices," Ann. Appl. Probab. **22**(3), 1266–1300 (2012).
[34] Y. Zhu, "A graphon approach to limiting spectral distributions of Wigner-type matrices," preprint arXiv:1806.11246 (2018).