EIGENVALUES OUTSIDE THE BULK OF INHOMOGENEOUS
ERDŐS-RÉNYI RANDOM GRAPHS

BY ARIJIT CHAKRABARTY,
SUKRIT CHAKRABORTY
AND
RAJAT SUBHRA HAZRA

Indian Statistical Institute

In this article, an inhomogeneous Erdős-Rényi random graph on
\{1, \ldots, N\} is considered, where an edge is placed between vertices
i and j with probability \(\varepsilon Nf(i/N, j/N)\), for \(i \leq j\), the choice being
made independent for each pair. The function \(f\) is assumed to be
non-negative definite, symmetric, bounded and of finite rank \(k\). We
study the edge of the spectrum of the adjacency matrix of such an in-
homogeneous Erdős-Rényi random graph under the assumption that
\(N\varepsilon N \to \infty\) sufficiently fast. Although the bulk of the spectrum of the
adjacency matrix, scaled by \(\sqrt{N\varepsilon N}\), is compactly supported, the \(k\)-th
largest eigenvalue goes to infinity. It turns out that the largest eigen-
value after appropriate scaling and centering converges to a Gaussian
law, if the largest eigenvalue of \(f\) has multiplicity 1. If \(f\) has \(k\) dis-
tinct non-zero eigenvalues, then the joint distribution of the \(k\) largest
eigenvalues converge jointly to a multivariate Gaussian law. The first
order behaviour of the eigenvectors is derived as a byproduct of the
above results. The results compliment the homogeneous case derived
by [11].

1. Introduction. Given a graph on \(N\) vertices, say, \(\{1, \ldots, N\}\), let \(A_N\) denote the
adjacency matrix of the graph, whose \((i, j)\)-th entry is 1 if there is an edge between
vertices \(i\) and \(j\) and 0 otherwise. Important statistics of the graph are the eigenvalues and
eigenvectors of \(A_N\) which encode crucial information about the graph. The present article
considers the generalization of the most studied random graph, namely the Erdős–Rényi
random graph (ERRG). It is a graph on \(N\) vertices where an edge is present independently
with probability \(\varepsilon N\). The adjacency matrix of the ERRG is a symmetric matrix with
diagonal entries zero, and the entries above the diagonal are independent and identically
distributed Bernoulli random variables with parameter \(\varepsilon N\). We consider an inhomogeneous
extension of the ERRG where the presence of an edge between vertices \(i\) and \(j\) is given by
a Bernoulli random variable with parameter \(p_{i,j}\) and these \(\{p_{i,j} : 1 \leq i < j \leq N\}\) need not
be same. When \(p_{i,j}\) are same for all vertices \(i\) and \(j\) it shall be referred as (homogeneous)
ERRG.

The mathematical foundations of inhomogeneous ERRG where the connection proba-
bilities \(p_{i,j}\) come from a discretization of a symmetric, non-negative function \(f\) on \([0, 1]^2\)

MSC 2010 subject classifications: Primary 60B20, 05C80 ; secondary 46L54

Keywords and phrases: Adjacency matrices, Inhomogeneous Erdős-Rényi random graph, Largest eigen-
value, Scaling limit, Stochastic block model
was initiated in [5]. The said article considered edge probabilities given by

\[ p_{i,j} = \frac{1}{N} f \left( \frac{i}{N}, \frac{j}{N} \right). \]

In that case the average degree is bounded and the phase transition picture on the largest cluster size was studied in the same article (see also [4, 19] for results on inhomogeneous ERRG). The present article considers a similar set-up where the average degree is unbounded and studies the properties of eigenvalues of the adjacency matrix. The connection probabilities are given by

\[ p_{i,j} = \varepsilon f \left( \frac{i}{N}, \frac{j}{N} \right) \]

with the assumption that

\[ N \varepsilon \rightarrow \infty. \tag{1.1} \]

Let \( \lambda_1(A_N) \geq \ldots \geq \lambda_N(A_N) \) be the eigenvalues of \( A_N \). It was shown in [8] (see also [22] for a graphon approach) that the empirical distribution of the centered adjacency matrix converges, after scaling with \( \sqrt{N \varepsilon_N} \), to a compactly supported measure \( \mu_f \). When \( f \equiv 1 \), the limiting law \( \mu_f \) turns out to be the semicircle law. Note that \( f \equiv 1 \) corresponds to the (homogeneous) ERRG (see also [9, 18] also for the homogeneous case). Quantitative estimates on the largest eigenvalue of the homogeneous case (when \( N \varepsilon_N \gg (\log N)^4 \)) were studied in [13, 21] and it follows from their work that the smallest and second largest eigenvalue converge to the edge of the support of semicircular law. The results were improved recently in [2] and the condition on sparsity can be extended to the case \( N \varepsilon_N \gg \log N \) (which is also the connectivity threshold). It was shown that inhomogeneous ERRG also has a similar behaviour. The largest eigenvalue of inhomogeneous ERRG when \( N \varepsilon_N \ll \log N \) was treated in [3]. Under the assumption that \( N^{\xi} \ll N \varepsilon_N \) for some \( \xi \in (2/3, 1] \), it was proved in [10, Theorem 2.7] that the second largest eigenvalue of the (homogeneous) ERRG after an appropriate centering scaling converge in distribution to the Tracy-Widom law. The results were recently improved in [15]. The properties of the largest eigenvector in the homogeneous case was studied in [10, 15, 18].

The scaling limit of the maximum eigenvalue of inhomogenous ERRG also turns out to be interesting. The fluctuations of the maximum eigenvalue in the homogeneous case were studied in [11]. It was proved that

\[ (\varepsilon_N(1 - \varepsilon_N))^{-1/2} (\lambda_1(A_N) - E[\lambda_1(A_N)]) \Rightarrow N(0, 2). \]

The above result was shown under the assumption that

\[ (\log N)^{\xi} \ll N \varepsilon_N \tag{1.2} \]

for some \( \xi > 8 \), which is a stronger assumption than (1.1).

It is well known that in the classical case of a (standard) Wigner matrix, the largest eigenvalue converges to the Tracy-Widom law. We note that there is a different scaling
between the edge and bulk of the spectrum in ERRG. As pointed out before that the bulk scales at $(N\varepsilon_N)^{-1/2}$ and the largest eigenvalue has the scaling $(N\varepsilon_N)^{-1}$. Letting

(1.3) \[ W_N = A_N - E(A_N), \]

where $E(A_N)$ is the entrywise expectation of $A_N$, it is easy to see that

\[ A_N = \varepsilon_N 11' + W_N, \]

where $1$ is the $N \times 1$ vector with each entry 1. Since empirical spectral distribution of $(N\varepsilon_N)^{-1/2}W_N$ converges to semi-circle law, the largest eigenvalue of the same converges to 2 almost surely. As $E[A_N]$ is a rank-one matrix, it turns out that the largest eigenvalue of $A_N$ scales like $N\varepsilon_N$, which is different from the bulk scaling.

To derive the fluctuations one needs to study in details what happens to the rank-one perturbations of a Wigner matrix. When $W_N$ is a symmetric random matrix with independent and identically distributed entries and the perturbation comes from a rank-one matrix then the fluctuation of the largest eigenvalue depends on the form of the deformation matrix (see [6, 7, 12]). For example, when

\[ M_N = \frac{W_N}{\sqrt{N}} + P_N \]

where $P_N = \theta 11'$ then $\lambda_1(M_N)$ has a Gaussian fluctuation. If $P_N$ is a diagonal matrix with single non-zero entry, the fluctuations depend on the distribution of the entries of $W_N$. The rank-one case was extended to finite rank case in the works of [1, 17]. We do not go into further discussion of the results there as they crucially use the fact that bulk behaviour (after scaling) in the limit is semicircular law, which is not generally the case here.

The adjacency matrix of the inhomogeneous ERRG does not fall directly into purview of the above results, since $W_N$, as in (1.3), is a symmetric matrix, with independent entries above the diagonal, but the entries have a variance profile, which also depends on the size of the graph. The inhomogeneity does not allow the use of local laws suitable for semicircle law in an obvious way. The present article aims at extending the results obtained in [11] for the case that $f$ is a constant to the case that $f$ is a non-negative, symmetric, bounded, Riemann integrable function on $[0, 1]^2$ which induces an integral operator of finite rank $k$, under the assumption that (1.2) holds. The case $k \geq 2$ turns out to be substantially difficult than the case $k = 1$ for the following reason. If $k = 1$, that is,

\[ E(A_N) = u_N u_N', \]

for some $N \times 1$ deterministic column vector $u_N$, then with high probability it holds that

\[ u_N' (\lambda I - W_N)^{-1} u_N = 1, \]

where $\lambda$ is the largest eigenvalue of $A_N$. The above equation facilitates the asymptotic study of $\lambda$. However, when $k \geq 2$, the above equation takes a complicated form. The observation which provides a way out of this is that $\lambda$ is also an eigenvalue of a $k \times k$
matrix with high probability; the same is recorded in Lemma 5.2 of Section 5. Besides, working with the eigenvalues of a \( k \times k \) matrix needs more linear algebraic work when \( k \geq 2 \). For example, the proof of Lemma 5.8, which is one of the major steps in the proof of a main result, becomes a tautology when \( k = 1 \).

The following results are obtained in the current paper. If the largest eigenvalue of the integral operator has multiplicity 1, then the largest eigenvalue of the adjacency matrix has a Gaussian fluctuation. More generally, it is shown that the eigenvalues which correspond to isolated eigenvalues, which will be defined later, of the induced integral operator jointly converge to a multivariate Gaussian law. Under the assumption that the function \( f \) is Lipschitz continuous, the leading order term in the expansion of the expected value of the isolated eigenvalues is obtained. Furthermore, under an additional assumption, the inner product of the eigenvector with the discretized eigenfunction of the integral operator corresponding to the other eigenvalues is shown to have a Gaussian fluctuation. Some important examples of such \( f \) include the rank-one case, and the stochastic block models. It remains an open question to see if the \((k+1)\)th eigenvalue follows a Tracy-Widom type scaling.

The mathematical set-up and the main results of the paper are stated in Section 2. Theorem 2.3 shows that of the \( k \) largest eigenvalues, the isolated ones, centred by their mean and appropriately scaled, converge to a multivariate normal distribution. Theorem 2.4 studies the first and second order of the expectation of the top \( k \) isolated eigenvalues. Theorems 2.5 and 2.6 study the behaviour of the eigenvectors corresponding to the top \( k \) isolated eigenvalues. Section 3 contains the special case when \( f \) is rank one and the example of stochastic block models. A few preparatory estimates are noted in Section 4, which are used later in the proofs of the main results, given in Section 5. The estimates in Section 4 are proved in Section 6.

2. The set-up and the results. Let \( f : [0,1] \times [0,1] \to [0,\infty) \) be a function which is symmetric, bounded, and Riemann integrable, that is,

\[
(2.1) \quad f(x,y) = f(y,x) , 0 \leq x,y \leq 1 ,
\]

and the set of discontinuities of \( f \) in \([0,1] \times [0,1]\) has Lebesgue measure zero. The integral operator \( I_f \) with kernel \( f \) is defined from \( L^2[0,1] \) to itself by

\[
(I_f(g))(x) = \int_0^1 f(x,y)g(y) \, dy , 0 \leq x \leq 1 .
\]

Besides the above, we assume that \( I_f \) is a non-negative definite operator and the range of \( I_f \) has a finite dimension.

Under the above assumptions \( I_f \) turns out to be a compact self-adjoint operator and from the spectral theory one obtains \( \theta_1 \geq \theta_2 \geq \ldots \geq \theta_k > 0 \) as the non-zero eigenvalues of \( I_f \) (where \( k \) is the dimension of the range of \( I_f \)), and eigenfunctions \( r_i \) corresponding to \( \theta_i \). Therefore, \( \{r_1, \ldots, r_k\} \) is an orthonormal set in \( L^2[0,1] \), and by assumption, each \( r_i \) is Riemann integrable (see Lemma 6.1 in Section 6). Also, for any \( g \in L^2[0,1] \) one has

\[
I_f(g) = \sum_{i=1}^k \theta_i \langle r_i, g \rangle_{L^2[0,1]} r_i .
\]
Note that this gives

\[
\int_0^1 \left( \sum_{i=1}^k \theta_i r_i(x)r_i(y) g(y) \right) \, dy = \int_0^1 f(x,y) g(y) \, dy \quad \text{for almost all } x \in [0,1].
\]

Since \( g \) is an arbitrary function in \( L^2[0,1] \) this immediately gives

\[
(2.2) \quad f(x,y) = \sum_{i=1}^k \theta_i r_i(x) r_i(y), \quad \text{for almost all } (x,y) \in [0,1] \times [0,1].
\]

Since the functions on both sides of the above equation are Riemann integrable, the corresponding Riemann sums are approximately equal, and hence there is no loss of generality in assuming that the above equality holds for every \( x \) and \( y \).

Let \( (\varepsilon_N : N \geq 1) \) be a real sequence satisfying

\[
0 < \varepsilon_N \leq \left[ \sup_{0 \leq x,y \leq 1} f(x,y) \right]^{-1}, \quad N \geq 1.
\]

We assume furthermore that \((1.2)\) holds for some \( \xi > 8 \), fixed once and for all, and that

\[
(2.3) \quad \lim_{N \to \infty} \varepsilon_N = \varepsilon_\infty,
\]

for some \( \varepsilon_\infty \geq 0 \). It’s worth emphasizing that we do not assume that \( \varepsilon_N \) necessarily goes to zero, although that may be the case.

For \( N \geq 1 \), let \( G_N \) be an inhomogeneous Erdős-Rényi graph where an edge is placed between vertices \( i \) and \( j \) with probability \( \varepsilon_N f(i/N,j/N) \), for \( i \leq j \), the choice being made independently for each pair in \( \{(i,j) : 1 \leq i \leq j \leq N\} \). Note that we allow self-loops. Let \( A_N \) be the adjacency matrix of \( G_N \). In other words, \( A_N \) is an \( N \times N \) symmetric matrix, where \( \{A_N(i,j) : 1 \leq i \leq j \leq N\} \) is a collection of independent random variable, and

\[
A_N(i,j) \sim \text{Bernoulli} \left( \varepsilon_N f \left( \frac{i}{N}, \frac{j}{N} \right) \right), \quad 1 \leq i \leq j \leq N.
\]

A few more notations are needed for stating the main results. For a moment, set \( \theta_0 = \infty \) and \( \theta_{k+1} = -\infty \), and define the set of indices \( i \) for which \( \theta_i \) is isolated as follows:

\[
\mathcal{I} = \{1 \leq i \leq k : \theta_{i-1} > \theta_i > \theta_{i+1} \}.
\]

For an \( N \times N \) real symmetric matrix \( M \), let \( \lambda_1(M) \geq \ldots \geq \lambda_N(M) \) denote its eigenvalues, as mentioned in Section 1. Finally, after the following definition, the main results will be stated.

**Definition.** A sequence of events \( E_N \) occurs with high probability, abbreviated as w.h.p., if

\[
P(E_N^c) = O \left( e^{-(\log N)^\eta} \right),
\]
for some $\eta > 1$. For random variables $Y_N, Z_N$,

$$Y_N = O_{hp}(Z_N),$$

means there exists a deterministic finite constant $C$ such that

$$|Y_N| \leq C|Z_N| \text{ w.h.p.},$$

and

$$Y_N = o_{hp}(Z_N),$$

means that for all $\delta > 0$,

$$|Y_N| \leq \delta|Z_N| \text{ w.h.p.}$$

We shall say

$$Y_N = O_p(Z_N),$$

to mean that

$$\lim_{x \to \infty} \sup_{N \geq 1} P(|Y_N| > x|Z_N|) = 0,$$

and

$$Y_N = o_p(Z_N),$$

to mean that for all $\delta > 0$,

$$\lim_{N \to \infty} P(|Y_N| > \delta|Z_N|) = 0.$$

The reader may note that if $Z_N \neq 0$ a.s., then “$Y_N = O_p(Z_N)$” and “$Y_N = o_p(Z_N)$” are equivalent to “$(Z_N^{-1}Y_N : N \geq 1)$ is stochastically tight” and “$Z_N^{-1}Y_N \xrightarrow{p} 0$”, respectively. Besides, “$Y_N = O_{hp}(Z_N)$” is a much stronger statement than “$Y_N = O_p(Z_N)$”, and so is “$Y_N = o_{hp}(Z_N)$” than “$Y_N = o_p(Z_N)$”.

In the rest of the paper, the subscript ‘$N$’ is dropped from notations like $A_N$, $W_N$, $\varepsilon_N$ etc. and the ones that will be introduced. The first result is about the first order behaviour of $\lambda_i(A)$.

**Theorem 2.1.** For every $1 \leq i \leq k$,

$$\lambda_i(A) = N\varepsilon_i \theta_i (1 + o_{hp}(1)).$$

An immediate consequence of the above is that for all $1 \leq i \leq k$, $\lambda_i(A)$ is non-zero w.h.p. and hence dividing by the same is allowed, as done in the next result. Define

$$e_i = \begin{bmatrix} N^{-1/2}r_i(1/N) \\ N^{-1/2}r_i(2/N) \\ \vdots \\ N^{-1/2}r_i(1) \end{bmatrix}, \quad 1 \leq i \leq k. \tag{2.4}$$

The second main result studies the asymptotic behaviour of $\lambda_i(A)$, for $i \in \mathcal{I}$, after appropriate centering and scaling.
Theorem 2.2. For every $i \in \mathcal{I}$, as $N \to \infty$,

$$\lambda_i(A) = E(\lambda_i(A)) + \frac{N \theta_i \varepsilon}{\lambda_i(A)} e_i' W e_i + o_p(\sqrt{\varepsilon}),$$

where $W$ is as defined in (1.3).

The next result is the corollary of the previous two.

Theorem 2.3. Assuming (1.2) and (2.3), if $\mathcal{I}$ is a non-empty set, then as $N \to \infty$,

$$(\varepsilon^{-1/2} (\lambda_i(A) - E[\lambda_i(A)]) : i \in \mathcal{I}) \Rightarrow (G_i : i \in \mathcal{I}),$$

where the right hand side is a multivariate normal random vector in $\mathbb{R}^{|\mathcal{I}|}$, with mean zero and

$$\text{Cov}(G_i, G_j) = 2 \int_0^1 \int_0^1 r_i(x)r_i(y)r_j(x)r_j(y)f(x,y) \left[1 - \varepsilon \infty f(x,y)\right] dx dy,$$

for all $i, j \in \mathcal{I}$.

It may be checked that the Lindeberg-Lévy central limit theorem implies that as $N \to \infty$,

$$(\varepsilon^{-1/2} e_i' W e_i : i \in \mathcal{I}) \Rightarrow (G_i : i \in \mathcal{I}),$$

where the right hand side is as in Theorem 2.3. Therefore, the latter would follow from Theorems 2.1 and 2.2.

Remark 2.1. If $f > 0$ a.e. on $[0,1] \times [0,1]$, then the Krein-Rutman theorem (see Lemma 6.2) implies that $1 \in \mathcal{I}$, and that $r_1 > 0$ a.e. Thus, in this case, if $\varepsilon \infty = 0$, then

$$\text{Var}(G_1) = 2 \int_0^1 \int_0^1 r_1(x)^2 r_1(y)^2 f(x,y) dx dy > 0.$$ 

Remark 2.2. That the claim of Theorem 2.3 may not hold if $i \notin \mathcal{I}$ is evident from the following example. Suppose that $\varepsilon \infty = 0$ and

$$f(x,y) = 1 \left(x \vee y < \frac{1}{2}\right) + 1 \left(x \wedge y > \frac{1}{2}\right) , 0 \leq x, y \leq 1.$$ 

Then, Theorem 2.3 itself implies that there exists $\beta_N \in \mathbb{R}$ such that

$$\varepsilon^{-1/2} (\lambda_1(A) - \beta) \Rightarrow G_1 \vee G_2,$$

where $G_1$ and $G_2$ are i.i.d. from normal with mean 0 and variance 2, and hence there doesn’t exist a centering and a scaling by which $\lambda_1(A)$ converges weakly to a non-degenerate normal distribution.
For the remaining results in this section, \( f \) will be assumed to be a Lipschitz function. The next main result of the paper studies asymptotics of \( \mathbb{E}(\lambda_i(A)) \) for \( i \in I \).

**Theorem 2.4.** Assume that \( f \) is Lipschitz continuous, that is, there exists \( K < \infty \) such that

\[
|f(x, y) - f(x', y')| \leq K \left( |x - x'| + |y - y'| \right).
\]

Then, for all \( i \in I \),

\[
\mathbb{E} \left[ \lambda_i(A) \right] = \lambda_i(B) + O \left( \sqrt{\varepsilon} + (N\varepsilon)^{-1} \right),
\]

where \( B \) is a \( k \times k \) symmetric deterministic matrix, depending on \( N \), defined by

\[
B(j, l) = \sqrt{\theta_j \theta_l} N \varepsilon e'_j e_l + \theta_i^{-2} \sqrt{\theta_j \theta_l} (N\varepsilon)^{-1} \mathbb{E} \left[ e'_j W^2 e_l \right], \quad 1 \leq j, l \leq k,
\]

and \( e_j \) and \( W \) are as defined in (2.4) and (1.3), respectively.

The next result studies the asymptotic behaviour of the normalized eigenvector corresponding to \( \lambda_i(A) \), again for isolated vertices \( i \). It is shown that the same is asymptotically aligned with \( e_i \), and hence it is asymptotically orthogonal to \( e_j \). Upper bounds on rates of convergence are obtained.

**Theorem 2.5.** As in Theorem 2.4, let \( f \) be a Lipschitz continuous function. Then, for a fixed \( i \in I \),

\[
\lim_{N \to \infty} P \left( \lambda_i(A) \text{ is an eigenvalue of multiplicity } 1 \right) = 1.
\]

If \( v \) is the eigenvector, with \( L^2 \)-norm 1, of \( A \) corresponding to \( \lambda_i(A) \), then

\[
e'_i v = 1 + O_p \left( (N\varepsilon)^{-1} \right),
\]

that is, \( N\varepsilon(1 - e'_i v) \) is stochastically tight. When \( k \geq 2 \), it holds that

\[
e'_j v = O_p \left( (N\varepsilon)^{-1} \right), \quad j \in \{1, \ldots, k\} \setminus \{i\}.
\]

The last main result of this paper studies finer fluctuations of (2.10) under an additional condition.

**Theorem 2.6.** Continue assuming \( f \) to be Lipschitz continuous, and let \( k \geq 2 \) and \( i \in I \). Furthermore, assume that

\[
N^{-2/3} \ll \varepsilon \ll 1.
\]

If \( v \) is as in Theorem 2.5, then, for all \( j \in \{1, \ldots, k\} \setminus \{i\} \),

\[
e'_j v = \frac{1}{\theta_i - \theta_j} \left[ \theta_i \frac{1}{\lambda_i(A)} e'_i W e_j + (N\varepsilon)^{-2} \frac{1}{\theta_i} \mathbb{E} \left[ e'_i W^2 e_j \right] \right] + o_p \left( \frac{1}{N\sqrt{\varepsilon}} \right).
\]
Remark 2.3. An immediate consequence of Theorem 2.6 is that under (2.11), there exists a deterministic sequence \((z_N : N \geq 1)\) given by

\[
z = \frac{1}{(N\varepsilon)^2 \theta_i (\theta_i - \theta_j)} \mathbb{E} (e'_i W^2 e_j),
\]

such that as \(N \to \infty\),

\[
N\sqrt{\varepsilon} (e'_j v - z)
\]

converges weakly to a normal distribution with mean zero, for all \(i \in I\) and \(j \in \{1, \ldots, k\} \setminus \{i\}\). Furthermore, the convergence holds jointly for all \(i\) and \(j\) (satisfying the above), and with (2.5), to a multivariate normal distribution in \(\mathbb{R}^{|I|}\) with mean zero, whose covariance matrix is not hard to calculate.

3. Examples and special cases.

The rank one case. Let us consider the special case of \(k = 1\), that is,

\[
f(x, y) = \theta r(x) r(y),
\]

for some \(\theta > 0\), and a bounded Riemann integrable \(r : [0, 1] \to [0, \infty)\) satisfying

\[
\int_0^1 r(x)^2 \, dx = 1.
\]

In this case, Theorem 2.3 implies that

\[
\varepsilon^{-1/2} (\lambda_1(A) - \mathbb{E} (\lambda_1(A))) \Rightarrow G_1,
\]

as \(N \to \infty\), where

\[
G_1 \sim N (0, \sigma^2),
\]

with

\[
\sigma^2 = 2\theta \left( \int_0^1 r(x)^3 \, dx \right)^2 - 2\theta^2 \varepsilon_{\infty} \left( \int_0^1 r(x)^4 \, dx \right)^2.
\]

If \(r\) is Lipschitz and \(\varepsilon_{\infty} = 0\), then the claim of Theorem 2.4 boils down to

\[
(3.1) \quad \mathbb{E} [\lambda_1(A)] = \theta N \varepsilon e'_1 e_1 + (N\varepsilon\theta)^{-1} \mathbb{E} (e'_1 W^2 e_1) + O \left( \sqrt{\varepsilon} + (N\varepsilon)^{-1} \right),
\]

where

\[
e_1 = N^{-1/2} [r(1/N) r(2/N) \ldots r(1)].
\]

Lipschitz continuity of \(r\) implies that

\[
e'_1 e_1 = 1 + O \left( N^{-1} \right),
\]

and hence (3.1) becomes

\[
(3.2) \quad \mathbb{E} [\lambda_1(A)] = \theta N \varepsilon + (N\varepsilon\theta)^{-1} \mathbb{E} (e'_1 W^2 e_1) + O \left( \sqrt{\varepsilon} + (N\varepsilon)^{-1} \right).
\]
Clearly,

\[ E(e_1^t W^2 e_1) = \frac{1}{N} \sum_{i=1}^{N} r \left( \frac{i}{N} \right)^2 E[W^2(i,i)] \]

\[ = \frac{1}{N} \sum_{i=1}^{N} r \left( \frac{i}{N} \right)^2 \sum_{1 \leq j \leq N, j \neq i} \varepsilon f \left( \frac{i}{N}, \frac{j}{N} \right) \left( 1 - \varepsilon f \left( \frac{i}{N}, \frac{j}{N} \right) \right) \]

\[ = \theta \varepsilon N^{-1} \sum_{1 \leq i \neq j \leq N} r \left( \frac{i}{N} \right)^3 r \left( \frac{j}{N} \right) + O(1) \]

\[ = N \theta \varepsilon \int_{0}^{1} r(x)^3 \, dx \int_{0}^{1} r(y) \, dy + O(\varepsilon). \]

In conjunction with (3.2) this yields

\[ E[\lambda_1(A)] = \theta N \varepsilon + \int_{0}^{1} r(x)^3 \, dx \int_{0}^{1} r(y) \, dy + O(\varepsilon^2) + (N \varepsilon)^{-1} \).

### Stochastic block model

Another important example is the stochastic block model, defined as follows. Suppose that

\[ f(x, y) = \sum_{i,j=1}^{k} p(i, j) 1_{B_i}(x) 1_{B_j}(y), 0 \leq x, y \leq 1, \]

where \( p \) is a \( k \times k \) symmetric positive definite matrix, and \( B_1, \ldots, B_k \) are disjoint Borel subsets of \([0,1]\) whose boundaries are sets of measure zero, that is, their indicators are Riemann integrable. We show below how to compute the eigenvalues and eigenfunctions of \( I_f \), the integral operator associated with \( f \).

Let \( \beta_i \) denote the Lebesgue measure of \( B_i \), which we assume without loss of generality to be strictly positive. Rewrite

\[ f(x, y) = \sum_{i,j=1}^{k} \tilde{p}(i, j)s_i(x)s_j(y), \]

where

\[ \tilde{p}(i, j) = p(i, j)\sqrt{\beta_i \beta_j}, 1 \leq i, j \leq k, \]

and

\[ s_i = \beta_i^{-1/2} 1_{B_i}, 1 \leq i \leq k. \]

Thus, \( \{s_1, \ldots, s_k\} \) is an orthonormal set in \( L^2[0,1] \). Let

\[ \tilde{p} = U' D U, \]

be a spectral decomposition of \( \tilde{p} \), where \( U \) is a \( k \times k \) orthogonal matrix, and

\[ D = \text{Diag}(\theta_1, \ldots, \theta_k), \]
for some $\theta_1 \geq \ldots \geq \theta_k > 0$.

Define functions $r_1, \ldots, r_k$ by

$$
\begin{bmatrix}
  r_1(x) \\
  \vdots \\
  r_k(x)
\end{bmatrix} = U \begin{bmatrix}
  s_1(x) \\
  \vdots \\
  s_k(x)
\end{bmatrix}, \quad x \in [0, 1].
$$

It is easy to see that $r_1, \ldots, r_k$ are orthonormal in $L^2[0, 1]$, and for $0 \leq x, y \leq 1$,

$$
\begin{align*}
  f(x, y) &= [s_1(x) \ldots s_k(x)] \tilde{p} [s_1(x) \ldots s_k(x)]' \\
  &= [r_1(x) \ldots r_k(x)] U \tilde{p} U' [r_1(x) \ldots r_k(x)]' \\
  &= \sum_{i=1}^{k} \theta_i r_i(x) r_i(y).
\end{align*}
$$

Thus, $\theta_1, \ldots, \theta_k$ are the eigenvalues of $I_f$, and $r_1, \ldots, r_k$ are the corresponding eigenfunctions.

**4. Estimates.** In this section, we’ll record a few estimates that will subsequently be used in the proof. Since their proofs are routine, they are being postponed to Section 6 which is the Appendix. Let $W$ be as defined in (1.3).

**Lemma 4.1.** There exist constants $C_1, C_2 > 0$ such that

$$
P \left( \|W\| \geq 2\sqrt{MN\varepsilon} + C_1(N\varepsilon)^{1/4}(\log N)^{\xi/4} \right) \leq e^{-C_2(\log N)^{\xi/4}},
$$

where $M = \sup_{0 \leq x, y \leq 1} f(x, y)$. Consequently,

$$
\|W\| = O_h p \left( \sqrt{N\varepsilon} \right).
$$

The notations $e_1$ and $e_2$ introduced in the next lemma and used in the subsequent lemmas should not be confused with $e_j$ defined in (2.4). Continuing to suppress ‘$N$’ in the subscript, let

$$
L = \lfloor \log N \rfloor,
$$

where $\lfloor x \rfloor$ is the largest integer less than or equal to $x$.

**Lemma 4.2.** There exists $0 < C_1 < \infty$ such that if $e_1$ and $e_2$ are $N \times 1$ vectors with each entry in $[-1/\sqrt{N}, 1/\sqrt{N}]$, then

$$
|E (e_1' W^n e_2) | \leq (C_1 N\varepsilon)^{n/2}, 2 \leq n \leq L.
$$

**Lemma 4.3.** There exists $\eta_1 > 1$ such that for $e_1, e_2$ as in Lemma 4.2, it holds that

$$
\max_{2 \leq n \leq L} P \left( |e_1' W^n e_2 - E (e_1' W^n e_2) | > N^{(n-1)/2} \varepsilon^{n/2}(\log N)^{n\xi/4} \right) = O \left( e^{-(\log N)^{\eta_1}} \right),
$$

where $\xi$ is as in (1.2). In addition,

$$
e_1' W e_2 = o_h p (N\varepsilon).
$$
Lemma 4.4. If $e_1, e_2$ are as in Lemma 4.2, then

\begin{equation}
\text{Var} \left( e'_1 W e_2 \right) = O(\varepsilon),
\end{equation}

and

\begin{equation}
\mathbb{E} \left( e'_1 W^3 e_2 \right) = O(N\varepsilon).
\end{equation}

5. Proof of the main results. This section is devoted to the proof of the main results. At this point, it should be clarified that in this section, $e_j$ will always be as defined in (2.4). We start with showing that Theorem 2.1 is a corollary of Lemma 4.1.

Proof of Theorem 2.1. For a fixed $i \in \{1, \ldots, k\}$, it follows that

\[ |\lambda_i(A) - \lambda_i(E(A))| \leq \|W\| = O_{hp} \left( (N\varepsilon)^{1/2} \right), \]

by Lemma 4.1. In order to complete the proof, it suffices to show that

\[ \lim_{N \to \infty} (N\varepsilon)^{-1} \lambda_i(E(A)) = \theta_i, \]

which however follows from the observation that (2.2) implies that

\begin{equation}
\mathbb{E}(A) = N\varepsilon \sum_{j=1}^{k} \theta_j e_j e'_j.
\end{equation}

This completes the proof. \qed

Proceeding towards the proof of Theorem 2.2, let us fix $i \in I$, once and for all, denote

\[ \mu = \lambda_i(A), \]

and let $V$ be a $k \times k$ matrix, depending on $N$ which is suppressed in the notation, defined by

\[ V(j, l) = \begin{cases} \frac{N\varepsilon}{\sqrt{\theta_j \theta_l}} e'_j (I - \frac{1}{\mu} W)^{-1} e_l, & \text{if } \|W\| < \mu, \\ 0, & \text{else,} \end{cases} \]

for all $1 \leq j, l \leq k$. It should be noted that if $\|W\| < \mu$, then $I - W/\mu$ is invertible. The first step towards Theorem 2.2 is to show that $V/N\varepsilon$ converges to $\text{Diag}(\theta_1, \ldots, \theta_k)$, that is, the $k \times k$ diagonal matrix with diagonal entries $\theta_1, \ldots, \theta_k$, w.h.p.

Lemma 5.1. As $N \to \infty$,

\[ V(j, l) = N\varepsilon \theta_j \left( 1(j = l) + o_{hp}(1) \right), \quad 1 \leq j, l \leq k. \]
Proof. For fixed $1 \leq j, l \leq k$, writing

$$\left( I - \frac{1}{\mu} W \right)^{-1} = I + O_{hp} \left( \mu^{-1} \|W\| \right),$$

we get that

$$V(j, l) = N\varepsilon \sqrt{\theta_j \theta_l} \left( e'_j e_l + \frac{1}{\mu} O_{hp}(\|W\|) \right).$$

Since

$$\lim_{N \to \infty} e'_j e_l = 1 (j = l),$$

and

$$\|W\| = o_{hp}(\mu)$$

by Lemma 4.1 and Theorem 2.1, the proof follows.

The next step, which is one of the main steps in the proof of Theorem 2.2, shows that the $i$-th eigenvalues of $A$ and $V$ are equal w.h.p.

Lemma 5.2. With high probability,

$$\mu = \lambda_i(V).$$

The proof of the above lemma is based on the following fact which is a direct consequence of the Gershgorin circle theorem; see Theorem 1.6, pg 8 of [20].

Fact 5.1. Suppose that $U$ is an $n \times n$ real symmetric matrix. Define

$$R_l = \sum_{1 \leq j \leq n, j \neq l} |U(j, l)|, \ 1 \leq l \leq n.$$ 

If for some $1 \leq m \leq n$ it holds that

$$U(m, m) + R_m < U(l, l) - R_l, \ \text{for all} \ 1 \leq l \leq m - 1,$$

and

$$U(m, m) - R_m > U(l, l) + R_l, \ \text{for all} \ m + 1 \leq l \leq n,$$

then

$$\{\lambda_1(U), \ldots, \lambda_n(U)\} \setminus \left( \bigcup_{1 \leq l \leq k, l \neq m} [U(l, l) - R_l, U(l, l) + R_l] \right) = \{\lambda_m(U)\}. $$

Remark 5.1. The assumptions (5.3) and (5.4) of Fact 5.1 mean that the Gershgorin disk containing the $m$-th largest eigenvalue is disjoint from any other Gershgorin disk.
Proof of Lemma 5.2. The first step is to show that

\[(5.5) \mu \in \{ \lambda_1(V), \ldots, \lambda_k(V) \} \text{ w.h.p.} \]

To that end, fix \( N \geq 1 \) and a sample point for which \( \|W\| < \mu \). The following calculations are done for that fixed sample point.

Let \( v \) be an eigenvector of \( A \), with norm 1, corresponding to \( \lambda_i(A) \). That is,

\[(5.6) \mu v = Av = Wv + N\varepsilon \sum_{l=1}^{k} \theta_l(e'_lv)e_l, \]

by (5.1). Since \( \mu > \|W\| \), \( \mu I - W \) is invertible, and hence

\[(5.7) v = N\varepsilon \sum_{l=1}^{k} \theta_l(e'_lv)(\mu I - W)^{-1} e_l. \]

Fixing \( j \in \{1, \ldots, k\} \) and premultiplying the above by \( \sqrt{\theta_j} \mu e'_j \) yields

\[
\mu \sqrt{\theta_j}(e'_jv) = N\varepsilon \sum_{l=1}^{k} \sqrt{\theta_j} \theta_l(e'_lv)e'_j \left( I - \frac{1}{\mu} W \right)^{-1} e_l = \sum_{l=1}^{k} V(j, l) \sqrt{\theta_l}(e'_lv). 
\]

As the above holds for all \( 1 \leq j \leq k \), this means that if

\[(5.8) u = \left[ \sqrt{\theta_1}(e'_1v) \ldots \sqrt{\theta_k}(e'_kv) \right]', \]

then

\[(5.9) Vu = \mu u. \]

Recalling that in the above calculations a sample point is fixed such that \( \|W\| < \mu \), what we have shown, in other words, is that a vector \( u \) satisfying the above exists w.h.p.

In order to complete the proof of (5.5), it suffices to show that \( u \) is a non-null vector w.h.p. To that end, premultiply (5.6) by \( v' \) to obtain that

\[
\mu = v'Wv + N\varepsilon \|u\|^2. 
\]

Dividing both sides by \( N\varepsilon \) and using Lemma 4.1 implies that

\[
\|u\|^2 = \Theta_i + o_{hp}(1). 
\]

Thus, (5.5) follows.

Lemma 5.1 shows that for all \( l \in \{1, \ldots, i-1\} \),

\[
\left[ V(i, i) + \sum_{1 \leq j \leq k, j \neq i} |V(i, j)| \right] - \left[ V(l, l) - \sum_{1 \leq j \leq k, j \neq l} |V(l, j)| \right] = N\varepsilon \left( \Theta_i - \Theta_l \right) \left( 1 + o_{hp}(1) \right),
\]
as $N \to \infty$. Since $i \in I$, $\theta_i - \theta_l < 0$, and hence
\[
V(i, i) + \sum_{1 \leq j \leq k, j \neq i} |V(i, j)| < V(l, l) - \sum_{1 \leq j \leq k, j \neq l} |V(l, j)| \text{ w.h.p.}
\]
A similar calculation shows that for $l \in \{i + 1, \ldots, k\}$,
\[
V(i, i) - \sum_{1 \leq j \leq k, j \neq i} |V(i, j)| > V(l, l) + \sum_{1 \leq j \leq k, j \neq l} |V(l, j)| \text{ w.h.p.}
\]
In view of (5.5) and Fact 5.1, the proof would follow once it can be shown that for all $l \in \{1, \ldots, k\} \setminus \{i\}$,
\[
|\mu - V(l, l)| > \sum_{1 \leq j \leq k, j \neq l} |V(l, j)| \text{ w.h.p.}
\]
This follows, once again, by dividing both sides by $N\varepsilon$ and using Theorem 2.1 and Lemma 5.1. This completes the proof. □

The next step is to write
\[
(I - \frac{1}{\mu} W)^{-1} = \sum_{n=0}^{\infty} \mu^{-n} W^n,
\]
which is possible because $\|W\| < \mu$. Denote
\[
Z_{j,l,n} = e'_j W^n e_l, 1 \leq j, l \leq k, n \geq 0,
\]
and for $n \geq 0$, let $Y_n$ be a $k \times k$ matrix with
\[
Y_n(j, l) = \sqrt{\theta_j \theta_l N\varepsilon Z_{j,l,n}}, 1 \leq j, l \leq k.
\]
The following bounds will be used several times.

**Lemma 5.3.** It holds that
\[
E(\|Y_1\|) = O\left(\frac{N\varepsilon^{3/2}}{2}\right),
\]
and
\[
\|Y_1\| = o_{\text{w.p.}}\left((N\varepsilon)^2\right).
\]

**Proof.** Lemma 4.4 implies that
\[
\text{Var}(Z_{j,l,1}) = O(\varepsilon), 1 \leq j, l \leq k.
\]
Hence,
\[
E\|Y_1\| = O\left(\frac{N\varepsilon}{2} \sum_{j, l=1}^{k} E|Z_{j,l,1}|\right)
\]
\[
= O\left(\frac{N\varepsilon}{2} \sum_{j, l=1}^{k} \sqrt{\text{Var}(Z_{j,l,1})}\right)
\]
\[
= O\left(\frac{N\varepsilon^{3/2}}{2}\right),
\]
the equality in the second line using the fact that \( Z_{j,l,1} \) has mean 0. This proves the first claim. The second claim follows from (4.3) of Lemma 4.3.

The next step is to truncate the infinite sum in (5.10) to level \( L \), where \( L = \lceil \log N \rceil \) as defined before.

**Lemma 5.4.** It holds that

\[
\mu = \lambda_i \left( \sum_{n=0}^{L} \mu^{-n}Y_n \right) + o_{hp}(\sqrt{\varepsilon}).
\]

**Proof.** From the definition of \( V \), it is immediate that for \( 1 \leq j, l \leq k \),

\[
V(j, l) = N\varepsilon \sqrt{\theta_j \theta_l} \sum_{n=0}^{\infty} \mu^{-n}Y_n W^n e_l 1(\|W\| < \mu),
\]

and hence

\[
V = 1(\|W\| < \mu) \sum_{n=0}^{\infty} \mu^{-n}Y_n.
\]

For the sake of notational simplicity, let us suppress \( 1(\|W\| < \mu) \). Therefore, with the implicit understanding that the sum is set as zero if \( \|W\| \geq \mu \), for the proof it suffices to check that

\[
\left\| \sum_{n=L+1}^{\infty} \mu^{-n}Y_n \right\| = o_{hp}(\sqrt{\varepsilon}).
\]  

(5.11)

To that end, Theorem 2.1 and Lemma 4.1 imply that

\[
\left\| \sum_{n=L+1}^{\infty} \mu^{-n}Y_n \right\| \leq \sum_{n=L+1}^{\infty} |\mu|^{-n} \|Y_n\| = O_{hp} \left( (N\varepsilon)^{-\frac{(L-1)}{2}} \right).
\]

In order to prove (5.11), it suffices to show that as \( N \to \infty \),

\[
-\log \varepsilon = o \left( (L-1) \log (N\varepsilon) \right).
\]

(5.12)

To that end, recall (1.2) to argue that

\[
N^{-1} = o(\varepsilon)
\]

(5.13)

and

\[
\log \log N = O(\log (N\varepsilon)).
\]

(5.14)
By (5.13), it follows that

\[- \log \varepsilon = O(\log N) \]
\[= o(\log N \log \log N) \]
\[= o((L - 1) \log (N\varepsilon)) , \]

the last line using (5.14). Therefore, (5.12) follows, which ensures (5.11), which in turn completes the proof.

In the next step, \(Y_n\) is replaced by its expectation for \(n \geq 2\).

**Lemma 5.5.** It holds that

\[\mu = \lambda_i \left( Y_0 + \mu^{-1}Y_1 + \sum_{n=2}^{L} \mu^{-n}E(Y_n) \right) + o_{hp}(\sqrt{\varepsilon}) . \]

**Proof.** In view of Theorem 2.1 and Lemma 5.4, all that has to be checked is

\[(5.15) \quad \sum_{n=2}^{L} (N\varepsilon)^{-n} \| Y_n - E(Y_n) \| = o_{hp}(\sqrt{\varepsilon}) . \]

For that, invoke Lemma 4.3 to claim that

\[(5.16) \quad \max_{2 \leq n \leq L, 1 \leq j, l \leq k} P \left( |Z_{j,l,n} - E(Z_{j,l,n})| > N^{(n-1)/2} \varepsilon^{n/2} (\log N)^{n\xi/4} \right) = O \left( e^{-(\log N)^q_1} \right) , \]

where \(\xi\) is as in (1.2).

Our next claim is that there exists \(C_2 > 0\) such that for \(N\) large,

\[(5.17) \quad \bigcap_{2 \leq n \leq L, 1 \leq j, l \leq k} \left[ |Z_{j,l,n} - E(Z_{j,l,n})| \leq N^{(n-1)/2} \varepsilon^{n/2} (\log N)^{n\xi/4} \right] \]
\[\subseteq \left[ \sum_{n=2}^{L} (N\varepsilon)^{-n} \| Y_n - E(Y_n) \| \leq C_2 \sqrt{\varepsilon} \left( (N\varepsilon)^{-1} (\log N)^{\xi} \right)^{1/2} \right] . \]

To see this, suppose that the event on the left hand side holds. Then, for fixed \(1 \leq j, l \leq k\), and large \(N\),

\[\sum_{n=2}^{L} (N\varepsilon)^{-n} \| Y_n(j,l) - E(Y_n(j,l)) \| \]
\[\leq \theta_1 N\varepsilon \sum_{n=2}^{L} (N\varepsilon)^{-n} |Z_{j,l,n} - E(Z_{j,l,n})| \]
\[\leq \theta_1 \sum_{n=2}^{\infty} (N\varepsilon)^{-(n-1)} N^{(n-1)/2} \varepsilon^{n/2} (\log N)^{n\xi/4} \]
\[= \left[ 1 - (N\varepsilon)^{-1/2} (\log N)^{\xi/4} \right]^{-1} \theta_1 \sqrt{\varepsilon} (N\varepsilon)^{-1/2} (\log N)^{\xi/2} . \]
Thus, (5.17) holds for some $C_2 > 0$.

Combining (5.16) and (5.17), it follows that

\[
P \left( \sum_{n=2}^{L} (N\varepsilon)^{-n} \|Y_n - E(Y_n)\| > C_2 \sqrt{\varepsilon} \left( (N\varepsilon)^{-1} (\log N)^{\xi} \right)^{1/2} \right)
= O \left( \log N e^{-\left(\log N\right)^n} \right)
= o \left( e^{-\left(\log N\right)^{(1+n)/2}} \right).
\]

This, with the help of (1.2), establishes (5.15) from which the proof follows. \hfill \Box

The goal of the next two lemmas is replacing $\mu$ by a deterministic quantity in

\[
\sum_{n=2}^{L} \mu^{-n} E(Y_n).
\]

**Lemma 5.6.** For $N$ large, the deterministic equation

(5.18) \hspace{1cm} x = \lambda_i \left( \sum_{n=0}^{L} x^{-n} E(Y_n) \right), x > 0,

has a solution $\tilde{\mu}$ such that

(5.19) \hspace{1cm} 0 < \liminf_{N \to \infty} (N\varepsilon)^{-1} \tilde{\mu} \leq \limsup_{N \to \infty} (N\varepsilon)^{-1} \tilde{\mu} < \infty.

**Proof.** Define a function

\[ h : (0, \infty) \to \mathbb{R}, \]

by

\[ h(x) = \lambda_i \left( \sum_{n=0}^{L} x^{-n} E(Y_n) \right). \]

Our first claim is that for any fixed $x > 0$,

(5.20) \hspace{1cm} \lim_{N \to \infty} \frac{1}{(N\varepsilon)^{-1}} h (xN\varepsilon) = \theta_i.

To that end, observe that since $E(Y_1) = 0$,

\[ h (xN\varepsilon) = \lambda_i \left( E(Y_0) + \sum_{n=2}^{L} \frac{1}{x^n} E(Y_n) \right). \]

Recalling that

\[ Y_0(j, l) = N\varepsilon \sqrt{\theta_j \theta_l} e_j^i e_l^i, 1 \leq j, l \leq k, \]

it follows by (5.2) that

(5.21) \hspace{1cm} \lim_{N \to \infty} (N\varepsilon)^{-1} E(Y_0) = \text{Diag}(\theta_1, \ldots, \theta_k).
Lemma 4.2 implies that
\[ \text{E}(Z_{j,t,n}) \leq (\text{O}(N\varepsilon))^{n/2}, \]
uniformly for \(2 \leq n \leq L\), and hence there exists \(0 < C_3 < \infty\) with
\[ (5.22) \quad \|\text{E}(Y_n)\| \leq (C_3 N\varepsilon)^{n/2+1}, \quad 2 \leq n \leq L. \]
Therefore,
\[ \left\| \sum_{n=2}^{L} (x N\varepsilon)^{-n} \text{E}(Y_n) \right\| \leq \sum_{n=2}^{\infty} (x N\varepsilon)^{-n} (C_3 N\varepsilon)^{n/2+1} \rightarrow C_3^2 x^{-2}, \]
as \(N \rightarrow \infty\). With the help of \((5.21)\), this implies that
\[ \lim_{N \rightarrow \infty} (N\varepsilon)^{-1} \left( \sum_{n=0}^{L} (x N\varepsilon)^{-n} \text{E}(Y_n) \right) = \text{Diag}(\theta_1, \ldots, \theta_k), \]
and hence \((5.20)\) follows. It follows that for a fixed \(0 < \delta < \theta_i\).
\[ \lim_{N \rightarrow \infty} (N\varepsilon)^{-1} [N\varepsilon(\theta_i + \delta) - h ((\theta_i + \delta) N\varepsilon)] = \delta, \]
and thus, for large \(N\),
\[ N\varepsilon(\theta_i + \delta) > h ((\theta_i + \delta) N\varepsilon). \]
Similarly, again for large \(N\),
\[ N\varepsilon(\theta_i - \delta) < h ((\theta_i - \delta) N\varepsilon). \]
Hence, for \(N\) large, \((5.18)\) has a solution \(\tilde{\mu}\) in \([N\varepsilon(\theta_i - \delta), N\varepsilon(\theta_i + \delta)]\), which trivially satisfies \((5.19)\). Hence the proof.

**Lemma 5.7.** If \(\tilde{\mu}\) is as in Lemma 5.6, then
\[ \mu - \tilde{\mu} = O_{h_p} ((N\varepsilon)^{-1} \|Y_1\| + \sqrt{\varepsilon}). \]

**Proof.** Lemmas 5.5 and 5.6 imply that
\[
|\mu - \tilde{\mu}| \\
= |\lambda_i \left( Y_0 + \mu^{-1} Y_1 + \sum_{n=2}^{L} \mu^{-n} \text{E}(Y_n) \right) - \lambda_i \left( \sum_{n=0}^{L} \tilde{\mu}^{-n} \text{E}(Y_n) \right) | + o_{h_p}(\sqrt{\varepsilon}) \\
\leq \|\mu^{-1} Y_1\| + |\mu - \tilde{\mu}| \sum_{n=2}^{L} \mu^{-n} \tilde{\mu}^{-n} \|\text{E}(Y_n)\| \|\sum_{j=0}^{n-1} \mu^j \tilde{\mu}^{n-1-j} + o_{h_p}(\sqrt{\varepsilon}) \\
= |\mu - \tilde{\mu}| \sum_{n=2}^{L} \mu^{-n} \tilde{\mu}^{-n} \|\text{E}(Y_n)\| \|\sum_{j=0}^{n-1} \mu^j \tilde{\mu}^{n-1-j} + O_{h_p} ((N\varepsilon)^{-1} \|Y_1\| + \sqrt{\varepsilon}). \]
Thus,

\[ |\mu - \bar{\mu}| \left[ 1 - \sum_{n=2}^{L} \mu^{-n} \|E(Y_n)\| \sum_{j=0}^{n-1} \mu^j \bar{\mu}^{n-1-j} \right] \leq O_{hp} \left( (N\varepsilon)^{-1} \|Y_1\| + \sqrt{\varepsilon} \right). \]

Equations (5.19) and (5.22) imply that

\[ \left| \sum_{n=2}^{L} \mu^{-n} \|E(Y_n)\| \sum_{j=0}^{n-1} \mu^j \bar{\mu}^{n-1-j} \right| = O_{hp} \left( \sum_{n=2}^{\infty} n(N\varepsilon)^{(n+1)}(C_3 N\varepsilon)^{n/2+1} \right) = O_{hp} \left( (N\varepsilon)^{-1} \right) = o_{hp}(1), \quad N \to \infty. \]

This completes the proof with the help of (5.23).

The next lemma is arguably the most important step in the proof of Theorem 2.2, the other major step being Lemma 5.2.

**Lemma 5.8.** There exists a deterministic \( \bar{\mu} \), which depends on \( N \), such that

\[ \mu = \bar{\mu} + \mu^{-1} Y_1(i,i) + o_{hp} \left( (N\varepsilon)^{-1} \|Y_1\| + \sqrt{\varepsilon} \right). \]

**Proof.** Define a \( k \times k \) deterministic matrix

\[ X = \sum_{n=0}^{L} \bar{\mu}^{-n} E(Y_n), \]

which, as usual, depends on \( N \). Lemma 5.7 and (5.24) imply that

\[ \left\| X - \sum_{n=0}^{L} \mu^{-n} E(Y_n) \right\| \leq |\mu - \bar{\mu}| \sum_{n=2}^{L} \mu^{-n} \|E(Y_n)\| \sum_{j=0}^{n-1} \mu^j \bar{\mu}^{n-1-j} = o_{hp} \left( (N\varepsilon)^{-1} \right) = o_{hp} \left( (N\varepsilon)^{-1} \|Y_1\| + \sqrt{\varepsilon} \right). \]

By Lemma 5.5 it follows that

\[ \mu = \lambda_i \left( \mu^{-1} Y_1 + X \right) + o_{hp} \left( (N\varepsilon)^{-1} \|Y_1\| + \sqrt{\varepsilon} \right). \]

Let

\[ H = X + \mu^{-1} Y_1 - \left( X(i,i) + \mu^{-1} Y_1(i,i) \right) I, \]

\[ M = X - X(i,i) I, \]

and

\[ \bar{\mu} = \lambda_i(X). \]
Clearly,
\[
\lambda_i \left( \mu^{-1} Y_1 + X \right) = X(i, i) + \mu^{-1} Y_1(i, i) + \lambda_i(H) \\
= \bar{\mu} - \lambda_i(M) + \mu^{-1} Y_1(i, i) + \lambda_i(H).
\]

Thus, the proof would follow with the aid of (5.25) if it can be shown that
\[
\lambda_i(H) - \lambda_i(M) = o_{hp} \left( (N\varepsilon)^{-1} \|Y_1\| \right).
\]

If \( k = 1 \), then \( i = 1 \) and hence \( H = M = 0 \). Thus, the above is a tautology in that case. Therefore, assume without loss of generality that \( k \geq 2 \).

Proceeding towards proving (5.26) when \( k \geq 2 \), set
\[
U_1 = (N\varepsilon)^{-1} M,
\]
and
\[
U_2 = (N\varepsilon)^{-1} H.
\]

The main idea in the proof of (5.26) is to observe that the eigenvector of \( U_1 \) corresponding to \( \lambda_i(U_1) \) is same as that of \( M \) corresponding to \( \lambda_i(M) \), and likewise for \( U_2 \) and \( X \). Hence, the first step is to use this to get a bound on the differences between the eigenvectors in terms of \( \|U_1 - U_2\| \).

An important observation that will be used later is that
\[
\|U_1 - U_2\| = O_{hp} \left( (N\varepsilon)^{-2} \|Y_1\| \right).
\]

The second claim of Lemma 5.3 implies that the right hand side above is \( o_{hp}(1) \). The same implies that for \( m = 1, 2 \) and \( 1 \leq j, l \leq k \),
\[
U_m(j, l) = (\theta_j - \theta_i)1(j = l) + o_{hp}(1), \ N \to \infty.
\]

In other words, as \( N \to \infty \), \( U_1 \) and \( U_2 \) converge to \( \text{Diag}(\theta_1 - \theta_i, \ldots, \theta_k - \theta_i) \) w.h.p. Therefore,
\[
\lambda_i(U_m) = o_{hp}(1), m = 1, 2.
\]

Let \( \tilde{U}_m \), for \( m = 1, 2 \), be the \((k - 1) \times (k - 1)\) matrix (recall that \( k \geq 2 \)) obtained by deleting the \( i \)-th row and the \( i \)-th column of \( U_m \), and let \( \tilde{u}_m \) be the \((k - 1) \times 1\) vector obtained from the \( i \)-th column of \( U_m \) by deleting its \( i \)-th entry. It is worth recording, for possible future use, that
\[
\|\tilde{u}_m\| = o_{hp}(1), m = 1, 2,
\]
which follows from (5.30), and that
\[
\|\tilde{u}_1 - \tilde{u}_2\| = O_{hp} \left( (N\varepsilon)^{-2} \|Y_1\| \right),
\]
follows from (5.29).
Equations (5.30) and (5.31) imply that \( \tilde{U}_m - \lambda_i(U_m)I_{k-1} \) converges w.h.p. to

\[
\text{Diag}(\theta_1 - \theta_i, \ldots, \theta_{i-1} - \theta_i, \theta_{i+1} - \theta_i, \theta_k - \theta_i).
\]

Since \( i \in I \), the above matrix is invertible. Fix \( \delta > 0 \) such that every matrix in the closed \( \delta \)-neighborhood \( B_\delta \), in the sense of operator norm, of the above matrix is invertible. Let

\[
C_4 = \sup_{E \in B_\delta} \| E^{-1} \|.
\]

Then, \( C_4 < \infty \). Besides, there exists \( C_5 < \infty \) satisfying

\[
\| E_1^{-1} - E_2^{-1} \| \leq C_5 \| E_1 - E_2 \|, \quad E_1, E_2 \in B_\delta.
\]

Fix \( N \geq 1 \) and a sample point such that \( \tilde{U}_m - \lambda_i(U_m)I_{k-1} \) belongs to \( B_\delta \). Then, it is invertible. Define a \((k - 1) \times 1\) vector

\[
\tilde{v}_m = - \left[ \tilde{U}_m - \lambda_i(U_m)I_{k-1} \right]^{-1} \tilde{u}_m, \quad m = 1, 2,
\]

and a \( k \times 1 \) vector

\[
v_m = [\tilde{v}_m(1), \ldots, \tilde{v}_m(i - 1), 1, \tilde{v}_m(i), \ldots, \tilde{v}_m(k - 1)]', \quad m = 1, 2.
\]

It is immediate that

\[
\| \tilde{v}_m \| \leq C_4 \| \tilde{u}_m \|, \quad m = 1, 2.
\]

Our next claim is that

\[
U_m v_m = \lambda_i(U_m) v_m, \quad m = 1, 2.
\]

This claim is equivalent to

\[
[U_m - \lambda_i(U_m)I_k] v_m = 0.
\]

Let \( \tilde{U}_m \) be the \((k - 1) \times k\) matrix obtained by deleting the \( i \)-th row of \( U_m - \lambda_i(U_m)I_k \). Since the latter matrix is singular, and \( \tilde{U}_m - \lambda_i(U_m)I_{k-1} \) is invertible, it follows that the \( i \)-th row of \( U_m - \lambda_i(U_m)I_k \) lies in the row space of \( \tilde{U}_m \). In other words, the row spaces of \( U_m - \lambda_i(U_m)I_k \) and \( \tilde{U}_m \) are the same, and so do their null spaces. Thus, (5.38) is equivalent to

\[
\tilde{U}_m v_m = 0.
\]

To see the above, observe that the \( i \)-th column of \( \tilde{U}_m \) is \( \tilde{u}_m \), and hence we can partition

\[
\tilde{U}_m = \begin{bmatrix} \tilde{U}_{m1} \tilde{u}_m \tilde{U}_{m2} \end{bmatrix},
\]

where \( \tilde{U}_{m1} \) and \( \tilde{U}_{m2} \) are of order \((k - 1) \times (i - 1)\) and \((k - 1) \times (k - i)\), respectively. Furthermore,

\[
[\tilde{U}_{m1} \tilde{U}_{m2}] = \tilde{U}_m - \lambda_i(U_m)I_{k-1}.
\]
Therefore,
\[ \tilde{U}_m v_m = \tilde{u}_m + [\tilde{U}_m \tilde{U}_m^*] \tilde{v}_m = \tilde{u}_m + \left( \tilde{U}_m - \lambda_i(U_m)I_{k-1} \right) \tilde{v}_m = 0. \]

Hence, (5.38) follows, which proves (5.37).

Next, we note
\[
\|v_1 - v_2\| = \|\tilde{v}_1 - \tilde{v}_2\| \\
\leq \left\| \left( \tilde{U}_1 - \lambda_i(U_1)I_{k-1} \right)^{-1} \right\| \|\tilde{u}_1 - \tilde{u}_2\| \\
+ \left\| \left( \tilde{U}_1 - \lambda_i(U_1)I_{k-1} \right)^{-1} - \left( \tilde{U}_2 - \lambda_i(U_2)I_{k-1} \right)^{-1} \right\| \|\tilde{u}_2\| \\
\leq C_4 \|\tilde{u}_1 - \tilde{u}_2\| + C_5 \left\| \left( \tilde{U}_1 - \lambda_i(U_1)I_{k-1} \right) - \left( \tilde{U}_2 - \lambda_i(U_2)I_{k-1} \right) \right\| \|\tilde{u}_2\|,
\]

where $C_4$ and $C_5$ being as in (5.34) and (5.35), respectively. Recalling that the above calculation was done on an event of high probability, what we have proven, with the help of (5.29) and (5.33), is that
\[
\|v_1 - v_2\| = O_{hp} \left( (N\varepsilon)^{-1} \|Y_1\| \right).
\]

Furthermore, (5.32) and (5.36) imply that
\[
\|\tilde{v}_m\| = o_{hp}(1).
\]

Finally, noting that
\[ U_m(i, i) = 0, m = 1, 2, \]
and that
\[ v_m(i) = 1, m = 1, 2, \]
it follows that
\[
|\lambda_i(U_1) - \lambda_i(U_2)| \\
= \left| \sum_{1 \leq j \leq k, j \neq i} U_1(i, j)v_1(j) - \sum_{1 \leq j \leq k, j \neq i} U_2(i, j)v_2(j) \right| \\
\leq \sum_{1 \leq j \leq k, j \neq i} |U_1(i, j)||v_1(j) - v_2(j)| + \sum_{1 \leq j \leq k, j \neq i} |U_1(i, j) - U_2(i, j)||v_2(j)| \\
= O_{hp} \left( \|\tilde{u}_1\| \|v_1 - v_2\| + \|U_1 - U_2\|\|\tilde{v}_2\| \right) \\
= o_{hp} \left( (N\varepsilon)^{-1} \|Y_1\| \right).
\]

Recalling (5.27) and (5.28), (5.26) follows, which completes the proof in conjunction with (5.25).

Now, we are in a position to prove Theorem 2.2.
Proof of Theorem 2.2. Recalling that
\[ Y_1(i, i) = \theta_i N \varepsilon e_i' W e_i, \]
it suffices to show that
\[ (5.39) \quad \mu - \mathbb{E}(\mu) = \mu^{-1} Y_1(i, i) + o_p(\sqrt{\varepsilon}). \]

Lemma 5.8 implies that
\[ (5.40) \quad \mu - \bar{\mu} = \mu^{-1} Y_1(i, i) + o_{hp} \left( (N \varepsilon)^{-1} \| Y_1 \| + \sqrt{\varepsilon} \right), \]
a consequence of which, combined with Lemma 5.3, is that
\[ (5.41) \quad \lim_{N \to \infty} (N \varepsilon)^{-1} \bar{\mu} = \theta_i. \]

Thus,
\[
\left| \frac{1}{\mu} Y_1(i, i) - \frac{1}{\bar{\mu}} Y_1(i, i) \right| = O_{hp} \left( (N \varepsilon)^{-2} \| \mu - \bar{\mu} \| \| Y_1 \| \right)
\]
\[
= o_{hp} (|\mu - \bar{\mu}|)
\]
\[
= o_{hp} \left( (N \varepsilon)^{-1} \| Y_1 \| + \sqrt{\varepsilon} \right)
\]
\[
= o_p(\sqrt{\varepsilon}),
\]
Lemma 5.3 implying the second line, the third line following from (5.40) and the fact that
\[ (5.43) \quad \| Y_1 \| = O_p \left( N \varepsilon^{3/2} \right), \]
which is also a consequence of the former lemma, being used for the last line. Using Lemma 5.8 once again, we get that
\[ (5.44) \quad \mu = \bar{\mu} + \frac{1}{\bar{\mu}} Y_1(i, i) + o_{hp} \left( (N \varepsilon)^{-1} \| Y_1 \| + \sqrt{\varepsilon} \right). \]

Let
\[ R = \mu - \bar{\mu} - \frac{1}{\bar{\mu}} Y_1(i, i). \]

Clearly,
\[ \mathbb{E}(R) = \mathbb{E}(\mu) - \bar{\mu}, \]
and (5.44) implies that for \( \delta > 0 \) there exists \( \eta > 1 \) with
\[
\mathbb{E}|R| \leq \delta \left( \sqrt{\varepsilon} + (N \varepsilon)^{-1} \mathbb{E}\| Y_1 \| \right) + E^{1/2} \left( \mu - \bar{\mu} - \frac{1}{\bar{\mu}} Y_1(i, i) \right) O \left( e^{-(\log N)^\eta} \right).
\]

Lemma 5.3 implies that
\[
\mathbb{E}|R| \leq o(\sqrt{\varepsilon}) + E^{1/2} \left( \mu - \bar{\mu} - \frac{1}{\bar{\mu}} Y_1(i, i) \right) O \left( e^{-(\log N)^\eta} \right).
\]
Next, (5.41) and that $|\mu| \leq N^2$ a.s. imply that
\[
E^{1/2} \left( \mu - \bar{\mu} - \frac{1}{\mu} Y_1(i,i) \right)^2 = O \left( N^2 \right) = o \left( \varepsilon^{1/2} N^3 \right) = o \left( \varepsilon^{1/2} e^{(\log N)^\eta} \right).
\]
Thus,
\[
E|R| = o(\sqrt{\varepsilon}),
\]
and hence
\[
E(\mu) = \bar{\mu} + o(\sqrt{\varepsilon}).
\]
This, in view of (5.44), implies that
\[
\mu = E(\mu) + \frac{1}{\bar{\mu}} Y_1(i,i) + o_p \left( \left( N\varepsilon \right)^{-1} \|Y_1\| + \sqrt{\varepsilon} \right)
\]
\[
= E(\mu) + \frac{1}{\bar{\mu}} Y_1(i,i) + o_p \left( \sqrt{\varepsilon} \right),
\]
the second line following from (5.43). This establishes (5.39) with the help of (5.42), and hence the proof.

Theorems 2.1 and 2.2 establish Theorem 2.3 with the help of (2.6). Now we shall proceed toward proving Theorem 2.4. For the rest of this section, (2.7) will be assumed, that is, $f$ is Lipschitz continuous. As a consequence, the functions $r_1, \ldots, r_k$, which are eigenfunctions of the integral operator $I_f$, are also Lipschitz.

The following lemma essentially proves Theorem 2.4.

**Lemma 5.9.** If $f$ is a Lipschitz function, then
\[
\mu = \lambda_i \left( Y_0 + (N\varepsilon \theta_i)^{-2} E(Y_2) \right) + O_p \left( \sqrt{\varepsilon} + (N\varepsilon)^{-1} \right).
\]

**Proof.** Lemma 5.5 implies that
\[
\mu = \lambda_i \left( \sum_{n=0}^{3} \mu^{-n} E(Y_n) \right) + O_p \left( \mu^{-1} \|Y_1\| + \sum_{n=4}^{L} \mu^{-n} \|E(Y_n)\| \right) + o_p(\sqrt{\varepsilon}).
\]
Equation (5.43) implies that
\[
\mu = \lambda_i \left( \sum_{n=0}^{3} \mu^{-n} E(Y_n) \right) + O_p \left( \sqrt{\varepsilon} + \sum_{n=4}^{L} \mu^{-n} \|E(Y_n)\| \right).
\]
From (5.22), it follows that
\[
\sum_{n=4}^{L} \mu^{-n} \|E(Y_n)\| = O_p \left( (N\varepsilon)^{-1} \right).
\]
and hence

\[ \mu = \lambda_i \left( \sum_{n=0}^{3} \mu^{-n} \mathbb{E}(Y_n) \right) + O_p \left( \sqrt{\varepsilon} + (N\varepsilon)^{-1} \right). \]

Lemma 4.4, in particular (4.5) therein, implies that

\[ \|\mathbb{E}(Y_3)\| = O \left( (N\varepsilon)^2 \right), \]

and hence

\[ \mu^{-3}\|\mathbb{E}(Y_3)\| = O_p \left( (N\varepsilon)^{-1} \right). \]

This, in conjunction with (5.45), implies that

\[ \mu = \lambda_i (Y_0 + \mu^{-2}\mathbb{E}(Y_2)) + O_p \left( \sqrt{\varepsilon} + (N\varepsilon)^{-1} \right). \]

An immediate consequence of the above and (5.22) is that

\[ \mu = \lambda_i (Y_0) + O_p(1). \]

Applying Fact 5.1 as in the proof of Lemma 5.2, it can be shown that

\[ |\lambda_i(Y_0) - Y_0(i, i)| \leq \sum_{1 \leq j \leq k, j \neq i} |Y_0(i, j)|. \]

Since \( r_i \) and \( r_j \) are Lipschitz functions, it holds that

\[ e_i' e_j = 1(i = j) + O \left( N^{-1} \right). \]

Hence, it follows that

\[ Y_0(i, i) = N\varepsilon \left( \theta_i + O(N^{-1}) \right) = N\varepsilon\theta_i + O(\varepsilon), \]

and similarly,

\[ Y_0(i, j) = O(\varepsilon), j \neq i. \]

Combining these findings with (5.48) yields that

\[ \lambda_i(Y_0) = N\varepsilon\theta_i + O(\varepsilon). \]

Equations (5.47) and (5.49) together imply that

\[ \mu = N\varepsilon\theta_i + O_p(1). \]

Therefore,

\[ \|\mu^{-2}\mathbb{E}(Y_2) - (N\varepsilon\theta_i)^{-2}\mathbb{E}(Y_2)\| \]

\[ = O_p \left( (N\varepsilon)^{-3}\|\mathbb{E}(Y_2)\| \right), \]

\[ = O_p \left( (N\varepsilon)^{-1} \right). \]

This in conjunction with (5.46) completes the proof. \( \square \)
Theorem 2.4 is a simple corollary of the above lemma, as shown below.

**Proof of Theorem 2.4.** A consequence of Theorem 2.2 is that

\[ \mu - E(\mu) = O_p(\sqrt{\varepsilon}) . \]

The claim of Lemma 5.9 is equivalent to

\[ \lambda_i(B) - \mu = O_p \left( \sqrt{\varepsilon} + (N\varepsilon)^{-1} \right) . \]

The proof follows by adding the two equations, and noting that \( B \) is a deterministic matrix.

Next we proceed towards the proof of Theorem 2.5, for which the following lemma will be useful.

**Lemma 5.10.** If \( f \) is Lipschitz continuous, then as \( N \to \infty \),

\[ e_j' (I - \mu^{-1}W)^{-n} e_l = 1(j = l) + O_p \left( (N\varepsilon)^{-1} \right) , 1 \leq j, l \leq k, n = 1, 2 . \]

**Proof.** For a fixed \( n = 1, 2 \), expand

\[ (I - \mu^{-1}W)^{-n} = I + n\mu^{-1}W + O_p \left( \mu^{-2}\|W\|^2 \right) . \]

The proof can be completed by proceeding along similar lines as in the proof of Lemma 5.9.

Now we are in a position to prove Theorem 2.5.

**Proof of Theorem 2.5.** Theorem 2.1 implies that (2.8) holds for any \( i \in I \). Fix such an \( i \), denote

\[ \mu = \lambda_i(A) , \]

and let \( v \) be the eigenvector of \( A \), having norm 1, corresponding to \( \mu \), which is uniquely defined with probability close to 1.

Fix \( k \geq 2 \), and \( j \in \{1, \ldots, k\} \setminus \{i\} \). Premultiplying (5.7) by \( e_j' \) yields that

\[ e_j' v = N\varepsilon \sum_{l=1}^{k} \theta_l (e_l'v)e_j' (\mu I - W)^{-1} e_l , \text{ w.h.p.} \]

Therefore,

\[ e_j' v \left( 1 - \theta_j \frac{N\varepsilon}{\mu} e_j' (I - \mu^{-1}W)^{-1} e_j \right) \]

\[ = \frac{N\varepsilon}{\mu} \sum_{1 \leq l \leq k, l \neq j} \theta_l (e_l'v)e_j' (I - \mu^{-1}W)^{-1} e_l , \text{ w.h.p.} \]
Lemma 5.10 implies that as \( N \to \infty \),

\[
1 - \theta_j \frac{N \varepsilon}{\mu} e_j' (I - \mu^{-1}W)^{-1} e_j \overset{p}{\to} 1 - \frac{\theta_j}{\theta_i} \neq 0.
\]

Therefore,

\[
e_j' v = O_p \left( \frac{N \varepsilon}{\mu} \sum_{1 \leq l \leq k, l \neq j} \theta_l e_l' (I - \mu^{-1}W)^{-1} e_l \right)
\]

\[
= O_p \left( \sum_{1 \leq l \leq k, l \neq j} |e_j' (I - \mu^{-1}W)^{-1} e_l| \right)
\]

\[
= O_p \left( (N \varepsilon)^{-1} \right),
\]

the last line being another consequence of Lemma 5.10. Thus, (2.10) holds.

Using (5.7) once again, we get that

\[
1 = (N \varepsilon)^2 \sum_{l', m=1}^k \theta_{l,m} (e_l' v)(e_m' v)e_l' (\mu I - W)^{-2} e_m,
\]

that is,

\[
\theta_i^2 (e_i' v)^2 e_i' (I - \mu^{-1}W)^{-2} e_i
\]

\[
= (N \varepsilon)^{-2} \mu^2 - \sum_{(l,m) \in \{1, \ldots, k\}^2 \setminus \{(i,i)\}} \theta_{l,m} (e_l' v)(e_m' v)e_l' (I - \mu^{-1}W)^{-2} e_m.
\]

Using Lemma 5.10 once again, it follows that

\[
e_i' (I - \mu^{-1}W)^{-2} e_i = 1 + O_p \left( (N \varepsilon)^{-1} \right).
\]

Thus, (2.9) would follow once it’s shown that

\[
(N \varepsilon)^{-2} \mu^2 = \theta_i^2 + O_p \left( (N \varepsilon)^{-1} \right),
\]

and that for all \((l,m) \in \{1, \ldots, k\}^2 \setminus \{(i,i)\},

\[
(e_l' v)(e_m' v)e_l' (I - \mu^{-1}W)^{-2} e_m = O_p \left( (N \varepsilon)^{-1} \right).
\]

Equation (5.53) is a trivial consequence of (5.50). For (5.54), assuming without loss of generality that \( l \neq i \), (2.10) implies that

\[
\left| (e_l' v)(e_m' v)e_l' (I - \mu^{-1}W)^{-2} e_m \right| = \left| (e_m' v)e_l' (I - \mu^{-1}W)^{-2} e_m \right| O_p \left( (N \varepsilon)^{-1} \right)
\]

\[
\leq \left| e_l' (I - \mu^{-1}W)^{-2} e_m \right| O_p \left( (N \varepsilon)^{-1} \right)
\]

\[
= O_p \left( (N \varepsilon)^{-1} \right),
\]

the last line following from Lemma 5.10. Thus, (5.54) follows, which in conjunction with (5.53) establishes (2.9). This completes the proof. \(\square\)
Finally, Theorem 2.6 is proved.

Proof of Theorem 2.6. Fix \( i \in I \). Recall (5.8) and (5.9), and let \( u \) be as defined in the former. Let \( \tilde{u} \) be the column vector obtained by deleting the \( i \)-th entry of \( u \), \( \tilde{V}_i \) be the column vector obtained by deleting the \( i \)-th entry of the \( i \)-th column of \( V \), and \( \tilde{V} \) be the \((k - 1) \times (k - 1)\) matrix obtained by deleting the \( i \)-th row and \( i \)-th column of \( V \). Then, (5.9) implies that

\[
(5.55) \quad \mu \tilde{u} = \tilde{V}\tilde{u} + u(i)\tilde{V}_i \text{, w.h.p.}
\]

Lemma 5.1 implies that

\[
\| I_{k} - \mu^{-1}V - \text{Diag} \left( 1 - \frac{\theta_1}{\theta_i}, \ldots, 1 - \frac{\theta_k}{\theta_i} \right) \| = o_p(1),
\]

and hence \( I_{k-1} - \mu^{-1}\tilde{V} \) is non-singular w.h.p. Thus, (5.55) implies that

\[
(5.56) \quad \tilde{u} = u(i)\mu^{-1} \left( I_{k-1} - \mu^{-1}\tilde{V} \right)^{-1}\tilde{V}_i \text{, w.h.p.}
\]

The next step is to show that

\[
(5.57) \quad \| \mu^{-1}V - \text{Diag} \left( \frac{\theta_1}{\theta_i}, \ldots, \frac{\theta_k}{\theta_i} \right) \| = o_p \left( \sqrt{\epsilon} \right).
\]

To see this, use the fact that \( f \) is Lipschitz to write for a fixed \( 1 \leq j, l \leq k \),

\[
V(j, l) = N\varepsilon \sqrt{\varepsilon} \theta_j \theta_l \left( e'_j e_l + \mu^{-1}e'_j We_l + O_p \left( \mu^{-2} \| W \|^2 \right) \right)
\]

\[
= N\varepsilon \sqrt{\varepsilon} \theta_j \theta_l \left( e'_j e_l + O_p \left( (N\varepsilon)^{-1} \right) \right)
\]

\[
= N\varepsilon \theta_l \left( 1(j = l) + O_p \left( (N\varepsilon)^{-1} \right) \right)
\]

\[
(5.58) \quad = N\varepsilon \theta_l \left( 1(j = l) + o_p \left( \sqrt{\varepsilon} \right) \right),
\]

the last line following from the fact that

\[
(5.59) \quad (N\varepsilon)^{-1} = o \left( \sqrt{\varepsilon} \right),
\]

which is a consequence of (2.11). This along with (5.50) implies that

\[
(5.60) \quad (N\varepsilon \theta_l)^{-1} \mu = 1 + o_p \left( \sqrt{\varepsilon} \right).
\]

Combining this with (5.58) yields that

\[
\mu^{-1}V(j, l) = \theta_i^{-1} \theta_j 1(j = l) + o_p \left( \sqrt{\varepsilon} \right).
\]

Thus, (5.57) follows, an immediate consequence of which is that

\[
(5.61) \quad \left\| \left(I_{k-1} - \mu^{-1}\tilde{V}\right)^{-1} - \tilde{D} \right\| = o_p \left( \sqrt{\varepsilon} \right),
\]
where
\[ \tilde{D} = \left[ \text{Diag} \left( 1 - \frac{\theta_1}{\theta_i}, \ldots, 1 - \frac{\theta_{i-1}}{\theta_i}, 1 - \frac{\theta_{i+1}}{\theta_i}, \ldots, 1 - \frac{\theta_k}{\theta_i} \right) \right]^{-1}. \]

Next, fix \( j \in \{1, \ldots, k\} \setminus \{i\} \). By similar arguments as above, it follows that
\[
V(i, j) = N \varepsilon \sqrt{\theta_i \theta_j} \left( \sum_{n=0}^{3} \mu^{-n} e_i^\prime W^n e_j + O_p (\mu^{-4} \|W\|^4) \right)
\]
\[
= N \varepsilon \sqrt{\theta_i \theta_j} \sum_{n=0}^{2} \mu^{-n} e_i^\prime W^n e_j + O_p ((N \varepsilon)^{-1})
\]
\[
= N \varepsilon \sqrt{\theta_i \theta_j} \sum_{n=1}^{2} \mu^{-n} e_i^\prime W^n e_j + o_p (\sqrt{\varepsilon}),
\]
using (5.59) once again, because
\[
N \varepsilon e_i^\prime e_j = O(\varepsilon) = o (\sqrt{\varepsilon}),
\]
and
\[
N \varepsilon \mu^{-3} e_i^\prime W^3 e_j = O_p ((N \varepsilon)^{-2} E(e_i^\prime W^3 e_j)) = o_p (\sqrt{\varepsilon}),
\]
by (4.5). Thus,
\[
V(i, j) - N \varepsilon \sqrt{\theta_i \theta_j} \mu^{-1} e_i^\prime W e_j = N \varepsilon \sqrt{\theta_i \theta_j} \mu^{-2} e_i^\prime W^2 e_j + o_p (\sqrt{\varepsilon})
\]
\[
= N \varepsilon \sqrt{\theta_i \theta_j} \mu^{-2} E(e_i^\prime W^2 e_j) + o_p (\sqrt{\varepsilon})
\]
\[
= (N \varepsilon)^{-1} \theta_j^{-1/2} \theta_i^{-3/2} E(e_i^\prime W^2 e_j) + o_p (\sqrt{\varepsilon}),
\]
the second line following from Lemma 4.3, and the last line from (5.59), (5.60) and Lemma 4.2. In particular,
\[
V(i, j) = O_p (1).
\]
The above in conjunction with (5.61) implies that
\[
\left[ (I_{k-1} - \mu^{-1} \tilde{V})^{-1} \tilde{V} \right] (j)
\]
\[
= \left( 1 - \frac{\theta_j}{\theta_i} \right)^{-1} \sqrt{\theta_i \theta_j} \left[ (N \varepsilon)^{-1} \theta_i^{-2} E(e_i^\prime W^2 e_j) + N \varepsilon \mu^{-1} e_i^\prime W e_j \right] + o_p (\sqrt{\varepsilon}).
\]
In light of (5.56), the above means that
\[
e_i^\prime v
\]
\[
= (e_i^\prime v) \mu^{-1} \left( 1 - \frac{\theta_j}{\theta_i} \right)^{-1} \left[ (N \varepsilon)^{-1} \theta_i^{-1} E(e_i^\prime W^2 e_j) + N \varepsilon \theta_i \mu^{-1} e_i^\prime W e_j + o_p (\sqrt{\varepsilon}) \right]
\]
\[
= \mu^{-1} \left( 1 - \frac{\theta_j}{\theta_i} \right)^{-1} \left[ (N \varepsilon)^{-1} \theta_i^{-1} E(e_i^\prime W^2 e_j) + N \varepsilon \theta_i \mu^{-1} e_i^\prime W e_j + o_p (\sqrt{\varepsilon}) \right],
\]
the last line following from (2.9) and (5.59). Using (5.60) once again yields that
\[
N \varepsilon (e_i^\prime v) = \frac{1}{\theta_j - \theta_i} \left[ (N \varepsilon)^{-1} \theta_i^{-1} E(e_i^\prime W^2 e_j) + N \varepsilon \theta_i \mu^{-1} e_i^\prime W e_j \right] + o_p (\sqrt{\varepsilon}).
\]
This completes the proof. \( \square \)
6. Appendix.

**Lemma 6.1.** The eigenfunctions \( \{ r_i : 1 \leq i \leq k \} \) of the operator \( I_f \) are Riemann integrable.

**Proof.** Let \( D_f \subset [0, 1] \times [0, 1] \) be the set of discontinuity points of \( f \). Since \( f \) is Riemann integrable, the Lebesgue measure of \( D_f \) is 0. Let

\[
D_f^x = \{ y \in [0, 1] : (x, y) \in D_f \}, \quad x \in [0, 1].
\]

If \( \lambda \) is the one dimensional Lebesgue measure, then Fubini’s theorem implies that

\[
E = \{ x \in [0, 1] : \lambda(D_f^x) = 0 \}
\]

has full measure. Fix an \( x \in E \) and consider \( x_n \to x \) and observe that

\[
f(x_n, y) \to f(x, y) \quad \text{for all } y \notin D_f^x.
\]

Fix \( 1 \leq i \leq k \) and let \( \theta_i \) be the eigenvalue with corresponding eigenfunction \( r_i \), that is,

\[
r_i(x) = \frac{1}{\theta_i} \int_0^1 f(x, y)r_i(y) \, dy.
\]

(6.1)

Using \( f \) is bounded and \( r \in L^2[0, 1] \), dominated convergence theorem implies

\[
r_i(x_n) = \frac{1}{\theta_i} \int_{(D_f^x)^c} f(x_n, y)r_i(y) \, dy \to \frac{1}{\theta_i} \int_0^1 f(x, y)r_i(y) \, dy = r_i(x)
\]

and hence \( r \) is continuous at \( x \). So the discontinuity points of \( r_i \) form a subset of \( E^c \) which has Lebesgue measure 0. Further, (6.1) shows that \( r_i \) is bounded and hence Riemann integrability follows.

The following result is a version of the Perron-Frobenius theorem in the infinite dimensional setting (also known as the Krein-Rutman theorem). Since our integral operator is positive, self-adjoint and finite dimensional so the proof in this setting is much simpler and can be derived following the work of [16]. In what follows, we use for \( f, g \in L^2[0, 1] \), the inner product

\[
\langle f, g \rangle = \int_0^1 f(x)g(x) \, dx.
\]

**Lemma 6.2.** Suppose \( f > 0 \) a.e. on \( [0, 1] \times [0, 1] \). Then largest eigenvalue \( \theta_1 \) of \( T_f \) is positive and the corresponding eigenfunction \( r_1 \) can be chosen such that \( r_1(x) > 0 \) for almost every \( x \in [0, 1] \). Further, \( \theta_1 > \theta_2 \).

**Proof.** First observe that

\[
0 < \theta_1 = \langle r_1, r_1 \rangle = \langle r_1, I_f(r_1) \rangle = |\langle r_1, I_f(r_1) \rangle| \leq \langle u_1, I_f(u_1) \rangle \leq \theta_1
\]
where \( u_1(x) = |r_1|(x) \) and the last inequality follows from the Rayleigh-Ritz formulation of the largest eigenvalue. Hence note that the string of inequalities is actually an equality, that is,
\[
\langle r_1, I_f(r_1) \rangle = \langle u_1, I_f(u_1) \rangle.
\]
Breaking \( r_1 = r_1^+ - r_1^- \) implies either \( r_1^+ = 0 \) or \( r_1^- = 0 \) almost everywhere. Without loss of generality assume that \( r_1 \geq 0 \) almost everywhere. Using
\[
\theta_1 r_1(x) = \int_0^1 f(x,y) r_1(y) \, dy
\]
Note that if \( r_1(x) \) is zero for some \( x \) then due to the positivity assumption on \( f, r_1(y) = 0 \) for almost every \( y \in [0,1] \) which is a contradiction. Hence we have that \( r_1(x) > 0 \) almost every \( x \in [0,1] \).

For the final claim, without loss of generality assume that \( \int_0^1 r_1(x) \, dx \geq 0 \). If \( \theta_1 = \theta_2 = 0 \), then the previous argument would give us \( r_2(x) > 0 \) and this will contradict the orthogonality of \( r_1 \) and \( r_2 \).

Lemmas 4.1 – 4.4 are proved in the rest of this section. Therefore, the notations used here should refer to those in Section 4 and should not be confused with those in Section 5. For example, \( e_1 \) and \( e_2 \) are as in Lemma 4.2.

**Proof of Lemma 4.1.** Note that for any even integer \( k \)
\[
E(\|W\|^k) \leq E(\text{Tr}(W^k)).
\]
Using \( E(W(i,j)^2) \leq \varepsilon M \) and condition (1.2) it is immediate that conditions of Theorem 1.4 of [21] are satisfied. We shall use the following estimate from the proof of that result. It follows from [21, Section 4]
\[
E(\text{Tr}(W^k)) \leq K_1 N (2\sqrt{\varepsilon MN})^k
\]
where \( K_1 \) is some positive constant and there exists a constant \( a > 0 \) such that \( k \) can be chosen as
\[
k = \sqrt{2} a (\varepsilon M)^{1/4} N^{1/4}.
\]
Using (6.2), (6.3) and \((1 - x)^k \leq e^{-kx} \) for \( k, x > 0 \),
\[
P \left( \|W\| \geq 2\sqrt{MN\varepsilon} + C_1 (N\varepsilon)^{1/4}(\log N)^{\xi/4} \right)
\]
\[
= K_1 N \left( 1 - \frac{C_1 (N\varepsilon)^{1/4}(\log N)^{\xi/4}}{2\sqrt{MN\varepsilon} + C_1 (N\varepsilon)^{1/4}(\log N)^{\xi/4}} \right)^k
\]
\[
\leq K_1 N \exp \left( -\frac{k C_1 (N\varepsilon)^{1/4}(\log N)^{\xi/4}}{2\sqrt{MN\varepsilon} + C_1 (N\varepsilon)^{1/4}(\log N)^{\xi/4}} \right).
\]
Now plugging in the value of \( k \) in the bound (6.4) and using
\[
2\sqrt{M} + C_1 (N\varepsilon)^{-1/4}(\log N)^{\xi/4} \leq 2\sqrt{M} + C_1
\]
we have
\[(6.4) \leq K_1 N \exp \left( -\frac{C_1 aM^{1/4} \sqrt{2} (\log N)^{\xi/4}}{2\sqrt{M} + C_1} \right) \leq e^{-C_2 (\log N)^{\xi/4}}\]
for some constant \(C_2 > 0\) and \(N\) large enough. This proves (4.1) and hence the lemma.

**Proof of Lemma 4.2.** Let \(A\) be the event where Lemma 4.1 holds, that is, \(\|W\| \leq C \sqrt{N\varepsilon}\) for some constant \(C\). Since the entries of \(e_1\) and \(e_2\) are in \([-1/\sqrt{N}, 1/\sqrt{N}]\) so \(\|e_i\| \leq 1\) for \(i = 1, 2\). Hence on the high probability event it holds that
\[|E(e'_1 W^n e_2 1_A)| \leq (CN\varepsilon)^{n/2}.\]
We show that the above expectation on the low probability event \(A^c\) is negligible. For that first observe
\[|E[(e'_1 W^n e_2)^2]| \leq N^{nC'}\]
for some constant \(0 < C' < \infty\). Thus using Lemma 4.1 one has
\[|E(e'_1 W^n e_2 1_{A^c})| \leq \left| E \left[ (e'_1 W^n e_2)^2 \right]^{1/2} \right| P(A^c) \leq \exp \left( nC' \log N - 2^{-1} C_2 (\log N)^{\xi/4} \right)\]
Since \(n \leq \log N\) and \(\xi > 8\) the result follows.

**Proof of Lemma 4.3.** The proof is similar to the proof of Lemma 6.5 of [11]. The exponent in the exponential decay is crucial, so the proof is briefly sketched. Observe that
\[(6.5) e'_1 W^n e_2 - E(e'_1 W^n e_2) = \sum_{i \in \{1, \ldots, N\}^{n+1}} e_1(i_1) e_2(i_{n+1}) \left( n \prod_{l=1}^{n} W(i_l, i_{l+1}) - E \left[ n \prod_{l=1}^{n} W(i_l, i_{l+1}) \right] \right)\]
To use the independence, one can split the matrix \(W\) as \(W' + W''\) where the upper triangular matrix \(W'\) has entries \(W'(i, j) = W(i, j) 1(i \leq j)\) and the lower triangular matrix \(W''\) with entries \(W''(i, j) = W(i, j) 1(i > j)\). Therefore the above quantity under the sum breaks into \(2^n\) terms each having similar properties. Denote one such term as
\[L_n = \sum_{i \in \{1, \ldots, N\}^{n+1}} e_1(i_1) e_2(i_{n+1}) \left( n \prod_{l=1}^{n} W'(i_l, i_{l+1}) - E \left[ n \prod_{l=1}^{n} W'(i_l, i_{l+1}) \right] \right)\]
Using the fact that each entry of \(e_1\) and \(e_2\) are bounded by \(1/\sqrt{N}\), it follows by imitating the proof of Lemma 6.5 of [11] that
\[E[|L_n|^p] \leq \frac{(Cnp)^{np} (N\varepsilon)^{np/2}}{N^{np/2}},\]
where \( p \) is an even integer and \( C \) is a positive constant, independent of \( n \) and \( p \). Rest of the \( 2^n - 1 \) terms arising in (6.5) have the same bound and hence

\[
P \left( \left| e'_1 W^n e_2 - E(e'_1 W^n e_2) \right| > N^{(n-1)/2} \varepsilon^{n/2} (\log N)^{n\xi/4} \right) 
\leq \frac{(2Cnp)^{np} (N\varepsilon)^{np/2}}{Np/2 \varepsilon^p n^{n/2} (\log N)^{pm\xi/4}} = \frac{(2Cnp)^{np}}{(\log N)^{pm\xi/4}}.
\]

Choose \( \eta \in (1, \xi/4) \) and consider

\[
p = \frac{(\log N)^\eta}{2Cn},
\]

(with \( N \) large enough to make \( p \) an even integer) to get

\[
P \left( \left| e'_1 W^n e_2 - E(e'_1 W^n e_2) \right| > N^{(n-1)/2} \varepsilon^{n/2} (\log N)^{n\xi/4} \right) 
\leq \exp \left( -\frac{1}{2C} (\log N)^\eta (\frac{\xi}{4} - \eta) \log \log N \right).
\]

Note that \( n \leq L \), ensures that \( p > 1 \). Since the bound is uniform over all \( 2 \leq n \leq L \), the first bound (4.2) follows.

For (4.3) one can use Hoeffding’s inequality [14, Theorem 2] as follows. Define

\[
\bar{A}(k,l) = A(k,l)e_1(k)e_2(l), \quad 1 \leq k \leq l \leq N.
\]

Since \( A(k,l) \) are Bernoulli random variables, so one has \( \{\bar{A}(k,l) : 1 \leq k \leq l \leq N\} \) are independent random variables taking values in \([-1/N, 1/N]\) and hence by Hoeffding’s inequality we have, for any \( \delta > 0 \),

\[
P \left( \left| \sum_{1 \leq k \leq l \leq N} \bar{A}(k,l) - E\left( \sum_{1 \leq k \leq l \leq N} \bar{A}(k,l) \right) \right| > \delta N \varepsilon \right) 
\leq 2 \exp \left( -\delta^2 (N \varepsilon)^2 \right) \leq 2 \exp \left( -\delta^2 (\log N)^{2\xi} \right).
\]

Dealing with the case \( k > l \) similarly, the desired bound on \( e'_1 W e_2 \) follows. \( \Box \)

**Proof of Lemma 4.4.** Follows by a simple moment calculation. \( \Box \)

**Acknowledgment.** RSH thanks Kavita Ramanan for a fruitful discussion. The research of AC and RSH was supported by the MATRICS grant of SERB. SC thanks Matteo Sfragara for helpful discussions.

**References.**

[1] F. Benaych-Georges, A. Guionnet, and M. Maida. Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices. *Electronic Journal of Probability*, 16:1621–1662, 2011.

[2] F. Benaych-Georges, C. Bordenave, and A. Knowles. Spectral radii of sparse random matrices. *arXiv preprint arXiv:1704.02945*, 2017.
LARGEST EIGENVALUE OF INHOMOGENEOUS ERG

[3] F. Benaych-Georges, C. Bordenave, and A. Knowles. Largest eigenvalues of sparse inhomogeneous erdős–rényi graphs. The Annals of Probability, 47(3):1653–1676, 2019.

[4] S. Bhamidi, R. Van Der Hofstad, J. van Leeuwaarden, et al. Scaling limits for critical inhomogeneous random graphs with finite third moments. Electronic Journal of Probability, 15:1682–1702, 2010.

[5] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. Random Structures & Algorithms, 31(1):3–122, 2007.

[6] M. Capitaine, C. Donati-Martin, D. Féral, et al. The largest eigenvalues of finite rank deformation of large wigner matrices: convergence and nonuniversality of the fluctuations. The Annals of Probability, 37(1):1–47, 2009.

[7] M. Capitaine, C. Donati-Martin, and D. Féral. Central limit theorems for eigenvalues of deformations of wigner matrices. In Annales de l’IHP Probabilités et statistiques, volume 48, pages 107–133, 2012.

[8] A. Chakrabarty, R. S. Hazra, F. den Hollander, and M. Sfragara. Spectra of adjacency and laplacian matrices of inhomogeneous Erdős-Rényi random graphs. arXiv preprint arXiv:1807.10112. To appear in Random Matrices:Theory and Applications, 2018.

[9] X. Ding, T. Jiang, et al. Spectral distributions of adjacency and laplacian matrices of random graphs. Annals of Applied Probability, 20(6):2086–2117, 2010.

[10] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Spectral statistics of erdős–rényi graphs ii: Eigenvalue spacing and the extreme eigenvalues. Communications in Mathematical Physics, 314(3):587–640, Sep 2012. ISSN 1432-0916. . URL https://doi.org/10.1007/s00220-012-1527-7.

[11] 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. ISSN 0091-1798. . URL https://doi.org/10.1214/11-AOP734.

[12] D. Féral and S. Péché. The largest eigenvalue of rank one deformation of large wigner matrices. Communications in mathematical physics, 272(1):185–228, 2007.

[13] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. Combinatorica, 1(3): 233–241, 1981. ISSN 0209-9683. . URL https://doi.org/10.1007/BF02579329.

[14] W. Hoeffding. Probability inequalities for sums of bounded random variables. Journal of the American Statistical Association, 58(301):13–30, 1963.

[15] J. O. Lee and K. Schnelli. Local law and Tracy–Widom limit for sparse random matrices. Probability Theory and Related Fields, 171(1-2):543–616, 2018.

[16] F. Ninio. A simple proof of the Perron-Frobenius theorem for positive symmetric matrices. Journal of Physics A: Mathematical and General, 9(8):1281, 1976.

[17] A. Pizzo, D. Renfrew, and A. Soshnikov. On finite rank deformations of wigner matrices. In Annales de l’IHP Probabilités et statistiques, volume 49, pages 64–94, 2013.

[18] L. V. Tran, V. H. Vu, and K. Wang. Sparse random graphs: eigenvalues and eigenvectors. Random Structures and Algorithms, 42(1):110–134, 2013. ISSN 1042-9832. . URL https://doi.org/10.1002/rsa.20406.

[19] R. van der Hofstad. Critical behavior in inhomogeneous random graphs. Random Structures & Algorithms, 42(4):480–508, 2013.

[20] R. S. Varga. Geršgorin and his circles, volume 36 of Springer Series in Computational Mathematics. Springer-Verlag, Berlin, 2004. ISBN 3-540-21100-4. . URL https://doi.org/10.1007/978-3-642-17798-9.

[21] V. H. Vu. Spectral norm of random matrices. Combinatorica, 27(6):721–736, 2007. ISSN 0209-9683. URL https://doi.org/10.1007/s00493-007-2190-z.

[22] Y. Zhu. Graphon approach to limiting spectral distributions of Wigner-type matrices. arXiv.1806.11246, 2018.