Two CLTs for Sparse Random Matrices

Simona Diaconu

Abstract

Let $G = G(n, p_n)$ be a homogeneous Erdos-Rényi graph, $A$ its adjacency matrix with eigenvalues $\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A)$. In such cases, it has been shown using local laws that $\lambda_k(A)$ can exhibit different behaviors: real-valued Tracy-Widom for $p_n \gg n^{-1/3}$, normal for $n^{-7/9} < p_n \ll n^{-2/3}$, or a mix of the two for $p_n = n^{-2/3}$. Additionally, this technique renders the largest eigenvalue $\lambda_1(A)$, separated from the rest of the spectrum by $\lambda_1(A) \gg \sqrt{n \log n}$, analyzed. This paper extends the range of the last convergence to $n^{-1} < p_n \leq 6\epsilon$, $\epsilon \in (0, 1)$ : the tool behind this is a CLT for the eigenvalue statistics of $A$, which is justified by the method of moments.

1 Introduction

Random graphs are crucial in disciplines such as computer science ([13]), algorithmic graph theory ([1]), and for several models, their adjacency matrices as well as their Laplacian transforms have been closely studied. The homogeneous Erdos-Rényi graph $G(n, p_n)$ is the most classical example in this field: a simple graph (i.e., no loops or multiple edges) on $n$ vertices, undirected, in which each edge is present independently with probability $p_n$. (see [8] for inhomogeneous graphs, i.e., $P(\ell \in E(G)) = p_{\ell,j}$). The focus of this paper is on the adjacency matrix $A$ of $G(n, p_n)$, which is sparse when $p_n \ll 1$, and as in expectation it has $o(n)$ nonzero entries per row. Denote its eigenvalues by $\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A)$: most of the existing results in the literature are concerned with $\lambda_1(A)$ or $\lambda_2(A)$. Furedi and Komlos [15] showed that if $p_n = p \in (0, 1)$,

$$\lambda_1(A) - [(n - 1)p + 1 - p] \Rightarrow N(0, 2p(1 - p)). \quad (1)$$

Subsequently, Krivelevich and Sudakov [17] proved

$$\lambda_1(A) = (1 + o_p(1)) \max(\sqrt{\Delta}, n p_n), \quad (2)$$

where $\Delta = \Delta(G(n, p_n))$ is the maximal vertex degree of $G(n, p_n)$: this together with what was known about $\Delta$ suggested a change in behavior for $\lambda_1(A)$ when $np_n \asymp \log n$ (in the sparse regime $np_n \ll \log n$, there is no concentration: some vertices have degrees much larger than the expectation $n p_n$ and other, much lower, whereas for the dense regime $np_n \gg \log n$, the graph is roughly regular: see [26], [19], in which the tails of $|A - E[A]|$ are upper bounded). The second largest eigenvalue has also been analyzed, especially the centered version $\lambda_2(A - E[A])$, for which the threshold $\log n$ is anew crucial and whose asymptotic law is fairly well understood. Benaych-Georges et al. [6], [7] showed for $np_n \gg \log n$,

$$\lambda_2(A - E[A]) \overset{p}{\rightarrow} 2, \quad (3)$$

while for $np_n \ll \log n$, the largest eigenvalues are of order greater than $\sqrt{\log n}$:

$$\lambda_k(A - E[A]) \approx \sqrt{\frac{\log (n/k)}{\log \log n/|np_n|}}, \quad k \geq n^{1-\epsilon}, \quad \epsilon > 0. \quad (4)$$

Alt et al. [2] complemented these two regimes by analyzing $np_n \approx c \log n$, and showing the transition from (3) to (4) occurs at $c = \frac{\log |E|}{\log n(k)}$ : this threshold has also been discovered through a different approach by Tikhomirov and Youssef [24].

Fluctuations of the largest eigenvalues in more general settings than $p_n = p$ covered by Furedi and Komlos have been obtained as well. Erdos et al. [11] extended their CLT when $p_n \gg n^{-1/3}$, under the natural normalization suggested by (1):

$$\frac{1}{\sqrt{p_n(1 - p_n)}} (\lambda_1(A) - E[\lambda_1(A)]) \Rightarrow N(0, 2). \quad (5)$$

It must be mentioned that despite the similarity between (1) and (5), they are justified through fundamentally different means. On the one hand, two stages render (1): finding the centering by the method of moments (i.e., computing $E[tr(A^m)]$ for $m \in \mathbb{N}$, and eventually letting $m \to \infty$), and obtaining the fluctuations, based on two observations: $\lambda_1(A)$ is much larger than all the other eigenvalues of $A$, and $\nu_n = [1, \ldots, 1]^T \in \mathbb{R}^n$ is

\[\text{Department of Mathematics, Stanford University, sdiaconu@stanford.edu}\]
roughly an eigenvector of $A$ with $Av_n \approx \lambda_n(A)v_n$. On the other hand, local laws (i.e., approximations of the Stieltjes transform at small scales) produce (5): after centering the entries of $A$ and normalizing them, the authors of [11] prove both the edge eigenvalues of the new matrix $H$ and its bulk spectrum behave like their counterparts in a Gaussian orthogonal ensemble. This together with Cauchy interlacing inequalities and the eigenvalues of $H$ being delocalized provide the behavior of $\lambda_1(A)$ as well as of $\lambda_2(A)$. Other uses of local laws led to the discovery of a transition for $\lambda_2(A)$ at $p_n \asymp n^{-2/3}$: Huang et al. [16] showed for $n^{-7/9} \ll p_n \ll n^{-2/3}$ its behavior is asymptotically normal:

$$n^{1/2}(\lambda_2(A) - [2(np_n)^{1/2} + (np_n)^{-1/2} - \frac{5(np_n)^{-3/2}}{4}]) \Rightarrow N(0, 1)$$

while at $p_n = cn^{-2/3}$, it is a combination of a normal and a real-valued Tracy-Widom law. This threshold is optimal as for $p_n \gg n^{-2/3}$ the fluctuations are given by the latter:

$$\frac{n^{3/2}}{\sqrt{np_n}}(\lambda_2(A) - 2\sqrt{np_n}(1 - p_n)) \Rightarrow TW_1$$

(Erdos et al. [12] covered $p_n \gg n^{-1/3}$ and Lee and Schnelli [18] extended the range to $p_n \gg n^{-2/3}$).

In this paper, the convergence (5) is justified under weakening $p_n \gg n^{-1/3}$ to $n^{-1} \leq p_n \leq \frac{1}{\epsilon}$ for $\epsilon \in (0, 1)$ (Theorem 2), coming close to the conjecture Erdos et al. made in [11] that (5) holds as long as $np_n \gg \log n$ (this seems the natural threshold in light of (2), and the results mentioned earlier on the concentration of the vertex degrees of $G(n, p_n)$). Theorem 2 is based on a CLT for the statistics of $A$, whose proof relies on the counting technique pioneered by Sinai and Soshnikov in [20]. This device was originally used to derive trace CLTs for real-valued Wigner matrices $A = (a_{ij})_{1 \leq i,j \leq n}$: suppose $a_{ij} \sim i.i.d.$, $a_{11} = -a_{11}$, $\mathbb{E}[a_{11}] = 0$, $\mathbb{E}[a_{11}^{2k}] \leq \left(\frac{2k}{n}\right)^k$ for all $k \in \mathbb{N}$; then

$$\text{tr}(A^m) - \mathbb{E}[\text{tr}(A^m)] \Rightarrow N(0, \frac{1}{m}).$$

In [20], this was shown for $m = o(n^{1/2})$, and the aforementioned authors extended the range to $m = o(n^{2/3})$ in [21]; moreover, Soshnikov in [22] employed it for $m \asymp n^{2/3}$ to derive universality of the joint distribution of the edge eigenvalues of Wigner matrices (both real- and complex-valued). This approach has been exploited in other contexts as well: upper bounding the trace expectation of large powers of $A$ (e.g., matrices with heavy-tailed entries in Auffinger et al. [3] and with $n^\mu, \mu \in (0, 1]$, nonzero entries per row in Benaych-Georges and Péché [5]) and obtaining universality results for one-rank perturbations of $A$ (Féral and Péché [14]).

A considerable difference between the current framework and the previous contexts in which this technique has been used is the lack of symmetry. As remarked in [11], one difficulty the method of moments faces with adjacency matrices is their entries not being even centered: particularly, universality oftentimes has been proven in two stages. First, justifying the claims in the Gaussian case (for which formulae can be utilized), and second, showing certain quantities, which determine the asymptotic behavior of the object of interest, depend solely on few moments of the underlying distributions: the second moment theorem by Tao and Vu is such an example (theorem 15 in [23]). Nevertheless, in the present case, relaxing symmetry to a vanishing first moment suffices to detect the primary contributors in the expectations of interest, leading to a trace CLT for sparse matrices that, to the author’s knowledge, has no close equivalent in the existing literature.

**Theorem 1.** Suppose $A = (a_{ij})_{1 \leq i,j \leq n}$ is symmetric with $a_{ij} \sim i.i.d.,$ Bernoulli distributed with parameter $p_0$. Assume there is $c_0 \in (0, 1)$ with

$$n^{1/2} \leq p_n \leq \frac{1}{2}.$$

Then for some $c_0 > 0$, as $n \to \infty$ and $m \in \mathbb{N}$, $\frac{1}{c_0 - c_0} \leq m \leq \frac{c_0 + \log\log n}{\log\log n}$,

$$\frac{\text{tr}(A^m) - \mathbb{E}[\text{tr}(A^m)]}{2m \cdot (np_n)^{2m-1} \sqrt{p_n(1 - p_n)}} \Rightarrow N(0, 2).$$

(6)

It can be easily shown using Lindeberg’s CLT

$$\frac{\text{tr}(A^2) - \mathbb{E}[\text{tr}(A^2)]}{n\sqrt{p_n(1 - p_n)}} \Rightarrow N(0, 2).$$

(7)
When \( m > 1 \), the terms dominating \( \mathbb{E}[\text{tr}(A^{2m})] \) differ from those for \( m = 1 \) : this distinction becomes apparent in the proof below (see (28) and (29) in subsection 2.4) and explains discrepancy between (6) and (7). Some comments on the inequalities on \( p_n, m \) are in order. The condition \( p_n \leq \frac{1}{2} \) ensures the centered entries of \( A \) have nonnegative moments, while the upper and lower bounds on \( m \) are needed for dealing with second-order terms in the trace expectations of powers of \( A \), and the moments of the ratio of interest, respectively: both restraints are likely suboptimal. Regarding odd powers of \( A \), although the current bounds discern the dominant components from the negligible ones solely when \( p_n \geq n^{-1/3+\epsilon} \), their overall contribution does not have a closed form as its even counterpart does (see end of section 2 for their description). Furthermore, despite \( A \) not being the adjacency matrix of a simple graph (it allows loops), the justification of Theorem 1 entails its diagonal entries play no role in the first-order term contributors, the result thus holding for adjacency matrices of simple graphs as well.

The CLT in Theorem 1 together with the separation between \( \lambda_1(A) \) and the rest of the spectrum of \( A \) \((\lambda_i(A) \approx np_{n},\text{ whereas } \lambda_2(A) \approx 2\sqrt{np_{n}})\) allow to recoup the asymptotic behavior of the former.

**Theorem 2.** *Under the assumptions of Theorem 1 on \( A \) and \( p_n \),*

\[
\frac{1}{\sqrt{p_n(1-p_n)}}(\lambda_1(A) - \mathbb{E}[\lambda_1(A)]) \xrightarrow{\text{d}} N(0, 2).
\]

The remainder of the paper is organized as follows: section 2 contains the proof of Theorem 1, and section 3 employs (6) to deduce Theorem 2.

### 2 Trace CLT

This section justifies (6). For the sake of simplicity, take \( p = p_n \), and let \( \nu_n = [11...1]^T \in \mathbb{R}^n, \bar{a}_{ij} = a_{ij} - p \), for which

\[
A = (p + \bar{a}_{ij})_{1 \leq i, j \leq n} := p_n^{1/2} \bar{A} + \bar{A},
\]

\[
\mathbb{E}[\bar{a}^2_{11}] = p(1 - p) \cdot [(1 - p)^{n-1} - (-p)^{n-1}], \quad q \in \mathbb{N};
\]

in particular,

\[
\mathbb{E}[\bar{a}_{11}] = 0, \quad \mathbb{E}[\bar{a}^2_{11}] = q(1 - p),
\]

and by convention, \( n, p, m \) satisfy the conditions of Theorem 1 in all the limits below unless otherwise stated. As mentioned in the introduction, (6) is proved computing moments of the ratio of interest: for \( l \in \mathbb{N} \), it is shown

\[
\lim_{n \to \infty} \frac{\mathbb{E}[\text{tr}(A^{2m})]}{\sqrt{2m \cdot (np)^{2m-1}} \sqrt{p(1 - p)}} = 0, \quad l = 2l_0 - 1, l_0 \in \mathbb{N}, l = 2l_0, l_0 \in \mathbb{N}.
\]

The technique developed by Sinai and Soshnikov in [20] is key for these convergences since it constructs a change of summation in

\[
\mathbb{E}[\text{tr}(M^q)] = \sum_{(i_0, i_1, ..., i_{q-1})} \mathbb{E}[m_{i_0 i_1} m_{i_1 i_2} ... m_{i_{q-1} i_0}] := \sum_{k=(i_0, i_1, ..., i_{q-1}, i_0)} \mathbb{E}[m_k],
\]

from cycles \( 1 := (i_0, i_1, ..., i_{q-1}, i_0) \) to tuples of nonnegative integers \((n_1, n_2, ..., n_q)\), which can be employed to infer the main contributors in (10). This transition relies on the distribution of \( a_{11} \) being symmetric, a property entailing \( l, \mathbb{E}[m_k] \neq 0 \) must have all undirected edges of even multiplicity: i.e., for \( u, v \in \{1, 2, ..., n\} \),

\[
\{0 \leq t < q - 1, i_t i_{t+1} = uv\} \in \{2m, m \in \mathbb{Z}\},
\]

where \( uv = vu \) denotes an undirected edge between \( v, u \). In the present situation, this property is not valid either for \( a_{11} \) or \( \bar{a}_{11} \). However, the advantage of the latter over the former is being centered. Up to a large extent, the primary upper hand of symmetric distributions is turning most of the summands on the right-hand side of (10) negligible. A weaker assumption, a vanishing first moment, accomplishes a considerable reduction as well: although some cycles \( 1 \) with \( \mathbb{E}[m_k] \neq 0 \) contain edges of odd multiplicity, each such graph can have at most \( q/2 \) pairwise distinct undirected edges, which coincides with the maximal number of edges of the first-order contributors in the symmetric case. This observation and the growth of the moments of \( a_{11} \) imply the dominant terms in (10) for \( M = \bar{A} \) are the same as if \( a_{11} \) were symmetric when \( q \) is even.

Before proceeding with the proof of (6), let us recall some terminology from Sinai and Soshnikov [20], necessary in what is to come. Interpret \( i := (i_0, i_1, ..., i_{q-1}, i_0) \) as a directed cycle with vertices among \( \{1, 2, ..., n\} \), and call \((i_{k-1}, i_k)\) its \( k^{th} \) edge for \( 1 \leq k \leq q \), where \( i_q := i_0 \); for \( u, v \in \{1, 2, ..., n\} \), \((u, v)\)
denotes a directed edge from $u$ to $v$, whereas $uv$ is undirected (the former are the building blocks of the cycles underlying the trace in (10), while the latter determine their expectations): in particular, $uv = vu$.

Call I an even cycle if each undirected edge appears an even number of times in it, and odd otherwise. If the entries of $M$ are symmetric, then solely the first category contributes to (10): when they are just centered, any cycle I consisting only of edges of multiplicity at least 2 can have $E[N(t)] \neq 0$.

The change of summation mentioned earlier is achieved by mapping $I$ with $E[N(t)] \neq 0$ to a tuple of nonnegative integers $(n_1, n_2, ..., n_q)$ together with bounding the sizes of the preimages of this transformation and the expectations of their elements. For I, call an edge $(i_k, i_{k+1})$ and its right endpoint $i_{k+1}$ marked if an even number of copies of $i_k,i_{k+1}$ precedes it: i.e., $\{t \in \mathbb{Z} : 0 \leq t \leq k-1, i_k,i_{k+1} = i_k,i_{k+1}\} \in 2\mathbb{Z}$. Each even cycle I has $q/2$ marked edges, whereas for its odd counterparts, there are strictly more marked edges than unmarked. In both cases, any vertex $i \in \{1, 2, ..., n\}$ of I, apart perhaps from $i_0$, is marked at least once (the first edge of containing $j$ is of the form $(i,j)$ since $i_0 \neq j$, and no earlier edge is adjacent to $j$). For $0 \leq k \leq q$, denote by $N_I(k)$ the set of $j \in \{1, 2, ..., n\}$ marked exactly $k$ times in I with $n_k := |N_I(k)|$. Then

$$\sum_{0 \leq k \leq q} n_k = n, \quad \sum_{1 \leq k \leq q} k n_k = \lfloor q/2 \rfloor + o, \quad (11)$$

where $o$ is the number of pairwise distinct undirected edges in I of odd multiplicity. Having constructed $(n_1, n_2, ..., n_q)$, the final task is bounding the number of cycles mapped to a given tuple and their expectations, which steps 1–5 below undertake.

Subsection(s)

- 2.1 and 2.2 deal with the primary contributors in (10) for $M = \check{A}$ when $q$ is even, odd, respectively;
- 2.3 introduces lacunary cycles, used to transition from $\check{A}$ to $A$;
- 2.4 completes the analysis of $E[\operatorname{tr}(A^{2m})]$;
- 2.5 concludes the proof of Theorem 1.

Before proceeding, some remarks on what occurs when $q$ is odd are in order. In such situations, the upper bounds obtained by the approach below are tight as long as $p \geq n^{-1/3+\epsilon}$ for some $\epsilon > 0$; in (21), the second summand in the bound is much larger than the first when $n \leq n^\epsilon$ because the exponent $3/4$ can be replaced by $3(1 - \epsilon)/2$ for any $\epsilon > 0$ (see its proof), and $(p n^{-1-\epsilon})^{3(1-\epsilon)/2} \ll n^{-1}$ is tantamount to $p \geq n^{1/3+\epsilon}$; this entails even cycles of length $q - 3$ with three identical loops attached generate the second term in (21); however, their overall contribution does not have a closed form as the even cycles do. Beyond the range $p \geq n^{1/3+\epsilon}$, the bound in (21) is not tight, creating difficulties in pinpointing the cycles in (10) which provide the order of the expectation: judging by its even counterpart, the same collection of cycles should be the leading constituents (the fewer edges of odd multiplicity, the larger the contribution: when the number of edges is even, this can be 0, whereas when it is odd, it must be at least 1: this latter case is equivalent to an even cycle of length $q - 2$ with three identical loops appended since any vertex is adjacent to an even number of edges of odd multiplicity, unless the loop at it is among them).

### 2.1 Even Powers of $\check{A}$

There exist $C, \epsilon > 0$ such that if $w \in (0,1), y \in (0, c), n^{n-1} \leq p \leq \frac{4}{3} m \leq y \min(n^{n/2}, \sqrt{m^{n-1}})$, then

$$C_m(p(1-p))^m n^{m+1}(1 - \frac{4m^2}{n}) \leq E[\operatorname{tr}(\check{A}^{2m})] \leq C_m(p(1-p))^{m+1}(1 + C \cdot y^{2/3}), \quad (12)$$

where $C_m = \frac{1}{2} \binom{2m}{m}$ is the $m^{th}$ Catalan number.

To justify (12), apart from the terminology from Sinai and Soshnikov [20] presented in the beginning of the section, call $(e, v)$ marked jointly if there exists a marked copy of $e$ in I with $v$ as its right endpoint. By convention, $e$ is always an undirected edge of multiplicity $m(e)$ in I; i.e., $m(e) := \{(0 \leq t < q-1, i_t,i_{t+1} = e)\}$, $u, v$ are vertices of I; i.e., $\exists 0 \leq t_1, t_2 \leq q, u = i_{t_1}, v = i_{t_2}$, and I is a fixed cycle satisfying certain properties.

To perform the change of summation in

$$E[\operatorname{tr}(\check{A}^{2m})] = \sum_{(i_0, i_1, ..., i_{2m-1})} E[\check{a}_{i_0i_1} \check{a}_{i_1i_2} ... \check{a}_{i_{2m-1}i_0}] := \sum_{k = \{(0, 1, ..., 2m-1, 0)\}} E[\check{a}_k],$$

let

$$\mathcal{O} := \mathcal{O}(I) = \{e, 2|m(e)+1\} \quad 2e \in \mathcal{O},$$

4
and suppose no undirected edge in $I$ has multiplicity one (else $E[\hat{o}] = 0$). Recall

$$
\sum_{1 \leq k \leq m+o} kn_k = m + o
$$
as the left-hand side is the number of marked edges (an edge $e$ contributes $|\langle m(e) + 1 \rangle|$). Fix $(o, n_1, n_2, \ldots, n_{m+o})$; analyze next how many cycles are mapped to this tuple, and the size of their corresponding expectations.

Step 1: Map the edges of $I$ to a random walk with $m + o$, $m - o$ increments $1$ and $-1$, respectively, according to $s_k = +1$ if $(i_{k-1}, i_k)$ is marked, $s_k = -1$ if $(i_{k-1}, i_k)$ is unmarked. There are

$$
\sigma(2m, m - o) = 2m - 1 \quad m - o - 2
$$
such paths:

**Lemma 1.** For $m, n \geq 0, m + n > 0$, let $\sigma(m, n)$ be the number of sequences with $m + n$ of them $1$, $n$ of them $-1$, and all partial sums nonnegative: i.e.,

$$
\sigma(m, n) = |(s_1, s_2, \ldots, s_{m+n}), s_i \in \{-1, 1\}, \sum_{j \leq i} s_j \geq 0, 0 \leq m + 2n, \sum_{j = m+2n} s_j = m|.
$$

Then

$$
\sigma(m, n) = \binom{m + 2n - 1}{n} - \binom{m + n - 1}{n - 2},
$$

where $\binom{m}{n} = \binom{m - 1}{2} = 0$.

**Proof.** Use induction on $n$ : note $\sigma(0, n) = C_n$, and $\sigma(m, 0) = 1, \sigma(m, 1) = m + 1$. The claim ensues for $n \leq 1$. Consider now $n \geq 2$ : looking at whether the last term in the sequence is $-1$ or $1$ entails

$$
\sigma(m, n) = \sigma(m + 1, n - 1) + \sigma(m + 1, n).
$$

(13)

Suppose the result holds for $m - 1 \geq 0$; since the claim for $\sigma(0, n)$ is also true, (13) implies showing the right-hand side of the desired formula satisfies this recurrence suffices:

$$
(m + n - 1) \binom{m + 2n - 1}{n} - \binom{m + n - 1}{n - 2} - \binom{m + 2n - 2}{n - 2} = \binom{m + 2n - 2}{n - 1} - \binom{m + 2n - 2}{n - 3},
$$

employing $\binom{m}{n} = \binom{m - 1}{n - 1} + \binom{m - 1}{n - 2}$. This completes the induction step and the proof of the lemma. □

Step 2: Order the marked vertices: the number of ways in which this can be done is

$$
\frac{(m + o)!}{\prod_{k \geq 1} (k)^{n_k}} \frac{1}{\prod_{k \geq 1} n_k!}.
$$

Step 3: Choose the vertices of $I$,

$$
V(i) = \bigcup_{0 \leq j \leq m - 1} \{i_j\}
$$

$|V(i)| \leq 1 + \sum_{k \geq 1} n_k$ because any vertex in $I$, apart perhaps from the first, is marked at least once.

Step 4: Establish the remaining vertices: as for even cycles, only the right endpoints of unmarked edges are yet to be chosen. Lemma 1 in [9] still holds as it does not use the parity of $I$, giving there are at most

$$
\prod_{k \geq 2} (2k)^{n_k}
$$
such configurations.

Step 5: For $q \in \mathbb{N}, q \geq 2$,

$$
0 \leq E[|\hat{d}|] = p(1 - p) - (1 - p)^{q - 1} - (1 - p)^{q - 1} \leq p(1 - p)^{q} (1 + (\frac{p}{1 - p})^{q - 1}) \leq p(1 - p)^{q} (1 + \frac{p}{1 - p}).
$$

(14)

Let $e(I)$ be the number of undirected edges appearing in $I$ : note $e(I) \geq \sum_{k \geq 1} n_k - 1$ with the inequality strict if $i_0$ is unmarked (take $v \in \cup_{j \geq 1} N_{i_0}(v)$, $v \neq i_0$ to $e = i_{t+1}$, where $t$ is minimal with $i_t = v$ or $i_{t+1} = v$; then $i_{t+1} = v$ because $v \neq i_0$, and $i_j \neq v$ for $0 \leq j \leq t$; this yields the mapping is injective since $v$ is the right endpoint of the first copy of $e$ in $I$). This last inequality in conjunction with (14) and $e(I) \leq m$ entails

$$
E[|\hat{a}|] \leq p|e(I)| (1 - p)^{2m} \cdot \frac{1}{1 - p} \leq p^{2m} \sum_{k \geq 1} n_k (1 - p)^{m}.
$$

(15)

It becomes apparent, when putting together steps 1–5, that a connection between $(n_k)_{k \geq 1}$ and $o$ is necessary.
Lemma 2.

\[ \sum_{k \geq 2} kn_k \geq 3o. \]

Proof. Let \( T \) be the number of \( e = uv \in \mathcal{O} \) with \( u \in N_1(1) \) and \( (e, u) \) marked jointly. It suffices to show:

\[ \sum_{k \geq 2} kn_k \geq 2 \cdot 2o - T = 4o - T, \]  

\[ \sum_{k \geq 2} kn_k \geq 3 \cdot T. \]  

(16) is a consequence of all marked endpoints of \( e \in \mathcal{O} \) being elements of \( \cup_{k \geq 2} N_1(k) \), apart from \( T \) of them, which have just one endpoint in \( \cup_{k \geq 2} N_1(k) \). To see this, suppose \( e = uv \in \mathcal{O}, u, v \in N_1(1) \) (this is the only situation in which a marked copy of \( e = uv \) can have an endpoint outside of \( \cup_{k \geq 2} N_1(k) \)) and not be among the \( T \) edges described above since \( (\frac{m(1, 1)}{k}) \geq 2 \) duplicates of \( e \) are marked); then \( u \neq v \), and let \( t \) be minimal with \( i_t \in \{u, v\} \); if \( t > 0 \), then \( (i_{t-1}, i_t) \) is marked (it is the first edge in \( t \) adjacent to \( i_t \)) and \( i_{t-1}i_t \neq e \), implying \( i_t \in \{u, v\} \cap (\cup_{k \geq 2} N_1(k)) \), contradiction. It is shown next \( e = i_0i_1 \in \mathcal{O}, i_0, i_1 \in N_1(1) \), both marked jointly with \( i_0i_1 \) cannot happen. Suppose this were possible: because \( m(i_0i_1) \geq 3 \) and the first marked copy of \( e \) has \( i_1 \) as its right endpoint, there are \( 0 < t_1 < t_2 \) with \( i_t, i_{t+1} = i_0i_1 \) unmarked, \( i_{t_2}, i_{t_2+1} = (i_1, i_0) \) marked; since the second time \( i_0 \) appears is marked unless its predecessor is \( i_1 \), this must occur at \( t_1 + 1 \) and \( (i_{t_1}, i_{t_1+1}) = (i_1, i_0) \); if \( t_1 > 1 \), consider the first appearance of \( i_1 \) in \( [2, t_1] \); this is the right endpoint of a marked edge \( (i_0 \) does not show up in this interval), contradicting \( i_1 \in N_1(1) \) hence \( t_1 = 1 \) and the first three vertices of \( i \) are \( (i_0, i_1, i_0) \); take now the second time \( i_1 \) appears in \( t \) (there is a third copy of \( i_1i_0 \), and \( i_0 \neq i_1 \)) this is the right endpoint of a marked edge, again absurd.

Having completed the proof of (16), continue with (17). Let \( e = uv \in \mathcal{O} \) if \( u, v \notin N_1(1) \), then its marked copies contribute at least \( 2 \) to the sum on the left-hand side. Else, assume without loss of generality \( u \in N_1(1), v \notin N_1(k), k \geq 3 \) (the previous paragraph entails \( u, v \in N_1(1) \) cannot occur). Since \( i \) is a cycle and \( u \neq v \), there are at least two elements of \( \mathcal{O} \) to which \( u \) is adjacent, only one of them having a marked jointly with \( u \); this implies the marked copies of \( e' \in \mathcal{O} \), adjacent to \( u \) contribute at least \( 3 \) to the sum above because there is no \( e \in \mathcal{O} \) with both endpoints in \( N_1(1) \), one contributes at least \( 2 \), and another \( 1 \), concluding (17).

The final bound becomes, using steps \( 1 - 5 \) above and \( \sum_{k \geq 2} n_k = m + o - \sum_{k \geq 2} (k - 1)n_k, \)

\[ \left[ \frac{2m - 1}{m - o} - \frac{2m - 1}{m - o - 2} \right] \cdot n^{1+m}(p(1-p))^m. \]

\[ \sum_{(n_k)_{k\geq 1}} p^{\sum_{k \geq 2} (k-1)n_k - n} \cdot \sum_{k \geq 2} (k-1)n_k \cdot \frac{(m + o)!}{(k-1)!^m} \cdot \frac{1}{\prod_{k \geq 2} n_k!} \cdot \prod_{k \geq 2} (2k)^{kn_k}; \]

if \( i_0 \) is marked (unmarked), then a factor of \( n \) \( p^{-1} \) can be left out in step 3 (5), and because \( np \geq 1, \) the latter could be dispensed with. Lemma 2 and \( k - 1 \geq \frac{1}{4} \) for \( k \geq 2 \) give

\[ \sum_{k \geq 2} (k - 1)n_k \geq w \frac{3o}{2} + (1 - w) \sum_{k \geq 2} (k - 1)n_k, \]

whereby the summand is at most

\[ (m + o)! \cdot \sum_{k \geq 1} n_k! \cdot \frac{(m + o)!}{(k-1)!^n} \cdot \frac{1}{\prod_{k \geq 2} n_k!} \cdot \prod_{k \geq 2} (2k)^{kn_k}; \]

\[ \sum_{(n_k)_{k\geq 1}} p^{\sum_{k \geq 2} (k-1)n_k - n} \cdot \sum_{k \geq 2} (k-1)n_k \cdot \frac{(m + o)!}{(k-1)!^n} \cdot \frac{1}{\prod_{k \geq 2} n_k!} \cdot \prod_{k \geq 2} (2k)^{kn_k} \leq \]

\[ \sum_{(n_k)_{k\geq 1}} p^{\sum_{k \geq 2} (k-1)n_k - n} \cdot \sum_{k \geq 2} (k-1)n_k \cdot \frac{(m + o)!}{(k-1)!^n} \cdot \frac{1}{\prod_{k \geq 2} n_k!} \cdot \prod_{k \geq 2} (2k)^{kn_k}. \]
For any \( \epsilon \in [0, 1] \),
\[
\prod_{k \geq 2} \frac{1}{nk} \left( \frac{(2e)^{k}(m + o)^{k}}{p^{k-1}(1-w)(k-1)} \right)^{\epsilon} \leq \exp\left( \sum_{k \geq 2} \frac{(2e)^{k}(m + o)^{k}}{p^{k-1}n^{k-1}(1-w)(k-1)} \right),
\]  
(19)
and Lemma 2 implies for \( p \geq n^{-w-1}, 16m \leq p^{1/2}(1-w)/2 \)
\[
\prod_{k \geq 2} \frac{(2e)^{k}(m + o)^{k}}{p^{k-1}(1-w)(k-1)} \leq \prod_{k \geq 2} \frac{2e \cdot (m + o)^{k}}{p^{k-1}n^{k-2}(1-w)}/2 \leq \left( \frac{4em}{p^{1/2}(1-w)/2} \right)^{3}\cdot\left( \frac{n}{p^{1-3w}/2} \right)^{2}.
\]
As
\[
\left( \frac{4em}{p^{1/2}(1-w)/2} \right)^{3}\cdot\left( \frac{n}{p^{1-3w}/2} \right)^{2} = (4em)^{3}\cdot p^{(3-1)/2}n^{(3-1)/2-3w+2}/2,
\]
\( \epsilon = \frac{1}{2} \) and (18) entail the contribution for \( \alpha \geq 0, m \leq c_1 n^{w}/4, m^2 \leq c_2 p^{1-w} \) is at most
\[
\left[ \frac{2m-1}{m-o} - \frac{2m-1}{m-o-2} \right] \cdot n^{1+m}(p(1-p))^{m} \cdot ((4em)^{3}\cdot p^{(3-1)/2}n^{(3-1)/2-3w+2}/2).
\]
For \( c \leq \frac{1}{2} \), \( f(o) = \left[ \frac{2m-1}{m-o} - \frac{2m-1}{m-o-2} \right] \cdot c^o \) is decreasing in \( o \in \mathbb{Z}, o \geq 0 \) because \( f(o) = c^o - \frac{(2m+1)(2m)!}{(m+1)!} \), whereby
\[
\frac{f(o+1)}{f(o)} = c \cdot \frac{2o+3}{2o+1} \cdot \frac{m-o}{m+o+2} \leq c \cdot 3 \leq 1.
\]
Note this bound is tight when \( \alpha = 0 \) inasmuch as the contribution of the simple even cycles \( C(m) \) (i.e., \( n_1 = m, i_0 \) unmarked: see section 2.3 in [9] for a recursive description of \( C(m) \)) is
\[
C_{m}(p(1-p))^{m} \prod_{0 \leq i < m} \left( \frac{(2m-1)}{m-i} - \frac{(2m-1)}{m-2} \right) \cdot n^{1+m}(p(1-p))^{m} \cdot (1-O(m^2/n)),
\]
whereby the lower bound in (12) ensures because all the moments of \( \bar{a}_{11} \) are nonnegative (\( p \leq \frac{1}{2} \)), \( e^{-x} \geq 1 - 2x, x \in [0, \log 2], \) and for \( m \leq 16 \),
\[
1 \geq \prod_{0 \leq i < m} \frac{n-i}{n} \geq \exp(-2) \sum_{0 \leq i < m} \frac{i}{n} \geq \exp(-2m^2/n) \geq 1 - \frac{4m^2}{n}.
\]  
(20)
If \( o = 0 \), and the cycles are not simple, then \( \sum_{k \geq 2} kn_k > 0 \), from which the product in (19) is at most
\[
\exp\left( \sum_{k \geq 2} \frac{(2e)^{k}(m + o)^{k}}{p^{k-1}n^{k-1}(1-w)(k-1)} \right) - 1 \leq \frac{c_1 m^{2/3}}{(p^{1-w})^{1/3}}.
\]
This concludes the justification of (12).

### 2.2 Odd Powers of \( \bar{A} \)

Under the conditions stated in the beginning of subsection 2.1,
\[
\mathbb{E}(|\{\bar{A}^{2m+1}\}|) \leq C_{m+1}n^{2m}(p(1-p))^{m+1} \cdot \frac{2Cm^2}{np} \cdot \frac{Cm^2}{p^{1-w}n^2} \cdot \frac{(16m)^4}{n^2(1-p)}.
\]  
(21)
For cycles of length \( 2m+1 \), let \( 2o+1 := |C|, o \geq 0 \); there are \( m + o + 1 \) marked edges, and reasoning as done previously, the analogues of steps 1 – 5 are:

Step 1: Map the edges of \( i \) to a random walk with \( m + o + 1, m - o \) increments 1 and \( -1 \), respectively, according to \( s_e = +1 \) if \( (i_{k-1}, i_k) \) is marked, \( s_e = -1 \) if \( (i_{k-1}, i_k) \) is unmarked. The number of such paths is
\[
\sigma(2o+1, m-o) = \left( \frac{2m}{m-o} \right) - \left( \frac{2m}{m-o-2} \right).
\]
Step 2: Order the marked vertices: the number of ways in which this can be done is

\[
\frac{(m + o + 1)!}{\prod_{k \geq 1} (k!)^{n_k} \cdot \prod_{k \geq 1} n_k!}
\]

Steps 3.4: Idem.

Step 5: There is an extra factor of \(1 - p\) on the right-hand side of (15), and the analogue of Lemma 2 gives

\[
\sum_{k \geq 2} k n_k \geq 3o + 2
\]

(22)

since (16) becomes

\[
\sum_{k \geq 2} k n_k \geq 4o + 2 - T,
\]

from which (22) ensues:

\[
\sum_{k \geq 2} k n_k \geq \frac{3}{4} (4o + 2 - T) + \frac{1}{4} 3T = 3o + \frac{3}{2}.
\]

The final bound becomes

\[
\sum_{(n_k)_{k \geq 1}} p^{n - \sum_{k \geq 2} (k-1)n_k} \cdot \left( \frac{Cm}{pn^{1-m}} \right)^{1+3o/2} \leq \left( \frac{Cm^2}{pn^{1-m}} \right)^{1+3o/2}.
\]

The previous reasoning yields, since \(\sum_{k \geq 2} k n_k \geq 3o + 2 > 0\), the sum is at most

\[
2Cm^3 \left( \frac{pn^{1-m}}{2} \right)^{1+3o/2} \leq \left( \frac{Cm^2}{pn^{1-m}} \right)^{1+3o/2}.
\]

Since the ratio is smaller than 1, for \(g(o) = \left( \frac{2m}{m-o} \right) - \left( \frac{2m}{m-o-2} \right) \cdot c^o = c^o \cdot \frac{2(2m+1)!}{(m-o)(m-o+2)!},
\]

\[
g(o+1)/g(o) = c \cdot \frac{o+2}{o+3} \cdot \frac{m-o+2}{m+1-o} \leq c \cdot \frac{2}{3} \cdot 1 < 2c.
\]

Hence, when \(o > 0\), there is some decay by choosing \(\epsilon = 1/3: this yields the first term in (21) since

\[
2\left( \frac{2m}{m} \right) - \left( \frac{2m}{m-2} \right) \leq n^{2m-o} \cdot (p(1-p))^{m+1} \cdot \left( \frac{Cm^2}{pn^{1-m}} \right)^{7/4} = C_m \cdot n^{2m-o} \cdot (p(1-p))^{m+1} \cdot \left( \frac{Cm^2}{pn^{1-m}} \right)^{7/4}.
\]

When \(o = 0\), \(O = \{u\}\) because any vertex is adjacent either to an even number of elements of \(O\) and has the loop at it in the set. By erasing three copies of \(u\), what remains is an even cycle of length \(2m - 2: (12)\) then yields the second term in (21).

2.3 Lacunary Cycles

In the cycles \(i\) underlying \(tr(A^{2m}) = tr(\tilde{A} + (p)_{1 \leq i < j \leq 2m})\), the contribution of \(ij \in I\) is either \(\tilde{a}_{ij}\) or \(p\) (call the latter erased, and I lacunary).

Suppose \(s \geq 1\) segments with endpoints \((l_i, r_i)\) \(1 \leq s\) are erased (s minimal), treat the remaining \(s\) with endpoints \((r_i, l_{i+1})\) \(1 \leq s\), as a graph \((l_{i+1} := l_i)\), denoted by \(l_n\), with edge and vertex labels set and inherited from the bigger cycle in which it is embedded (i.e., fix \((l_i, r_i)\) \(1 \leq s\)). There is \(c > 0\) such that if \(w \in (0,1), n^{1-w} \leq p \leq \frac{1}{2}, m \leq \min \{ \left( p n^{1-w} \right) \}^{1/2}, \left( \frac{1}{ \log \left( \frac{m}{m-n} \right) } \right)\), then the contribution of any such configuration with \(2m, 2m + 1\) uneased edges is

\[
r^{m+s} \left( 1 - O\left( \frac{m^2}{n} \right) \right) (p(1-p))^m C_m \cdot \left( \sum_{m_i + \ldots + m_n = m, m_i \geq 1} C_{m_1 \ldots m_n} + O\left( \frac{m^4}{(pn^{1-w})^{1/2}} \right) \right),
\]

(23)

\[
O(n^{m+s} + m^4 C_{m+1}),
\]

(24)
respectively (the first term dominating the second in (23) when \(s \leq m\)), where by convention, the constants underlying \(O(\cdot)\) are universal.

Consider first the case in which there are 2m unerasable edges. Repeat the analysis from subsection 2.1 for \(I_v\) : begin marking edges and their right endpoints at \(r_1\), let \(2\alpha := |O|\), and recall
\[
\sum_{k \geq 1} kn_k = m + \alpha.
\]

Step 1: This remains true as at any point there are at least as many marked edges as unmarked even in such (potentially) disconnected graphs \(I_v\).

Step 2: Iden.

Step 3: Any vertex is marked at least once apart from the right endpoints: for \(v \notin \cup_{1 \leq i \leq s}\{I_v\} \cup \{r_j\}\) argue in the same vein as for cycles; suppose \(v \in \cup_{1 \leq i \leq s}\{I_v\}, v \notin \cup_{1 \leq i \leq s}\{r_j\} : \) then the first appearance of \(v\) in \(I_v\) is the right endpoint of an edge and must be marked. Hence \(n(0) \leq s\) vertices are unmarked, whereby
\[
|V(I)| := n(0) + \sum_{i \leq c \leq k} n_k.
\]

Step 4: Anew solely the right endpoints of the unmarked edges are yet to be chosen. Lemma 1 from [9] continues to hold for \(v \notin \cup_{1 \leq i \leq s}\{r_j\} : \) since the inequalities are driven by unmarked edges \((v, s)\) being numerous; when \(v \in \cup_{1 \leq i \leq s}\{r_j\}\), a trivial bound can be taken as \(m!\) (i contains \(m - o \leq m\) unmarked edges), which can be substituted by 1 if \(v \notin \cup_{1 \leq i \leq s}\{r_j\}\) is never marked.

Step 5: Let \(e(I_v)\) be the number of undirected edges appearing in \(I_v\) : note \(e(I_v) \geq \sum_{k \geq 1} k - s + n(0)\) (same rationale as for cycles: now the correction is given by the number of marked right endpoints, \(s - n(0)\)). Together with \(e(I_v) \leq m\),
\[
\mathbb{E}[n] \leq p^{e(I_v)}(1 - p)^{2m - p^{(e(I_v)) \leq n(m(1 - p))^{m - \sum_{k \geq 1} n_k}} \cdot \sum_{k \geq 1} n_k \sum_{k \geq 1} n_k \geq 3T - n(l),
\]
\[
\text{for the analogue of Lemma 2, the exact reasoning gives}
\]
\[
\sum_{k \geq 2} n_k \geq \sum_{k \geq 2} \frac{3}{4} (4o - T - n(l)) + \frac{1}{4} (3T - n(l)) = 3o - n(l).
\]

The bound for the contribution of \(I_v\) becomes, using steps 1 - 5, \(\sum_{k \geq 2} n_k = m + o - \sum_{k \geq 2} (k - 1) n_k,\)
\[
\left(\begin{array}{cc}
2m - 1 \\
m - o
\end{array}\right) \cdot n(m(1 - p))^{m - \sum_{k \geq 1} n_k} \cdot \sum_{k \geq 1} n_k \sum_{k \geq 1} n_k \geq \sum_{k \geq 2} (k - 1) n_k,
\]
\[
(n^2 - \sum_{k \geq 2} (k - 1) n_k) \cdot \prod_{k \geq 2} (2k + 2s) \cdot \frac{1}{k \geq 2} \cdot \frac{1}{k \geq 2} \cdot \frac{1}{k \geq 2} \cdot \frac{1}{k \geq 2} \cdot \frac{1}{k \geq 2}
\]

because for \(v = r_i \in N(v)\), marked, a trivial bound in step 4 is \((2s + 2)!\), while for the rest \(2k + 2s\) suffices (see Lemma 2 in [9]). (27) gives
\[
\sum_{k \geq 2} (k - 1) n_k \geq w \cdot \frac{3o - n(l)}{2} + (1 - w) \sum_{k \geq 2} (k - 1) n_k,
\]
whereby the sum is at most
\[
\left(\begin{array}{cc}
2m - 1 \\
m - o
\end{array}\right) \cdot n(m(1 - p))^{m - \sum_{k \geq 1} n_k} \cdot \sum_{k \geq 1} n_k \sum_{k \geq 1} n_k \geq \sum_{k \geq 2} (k - 1) n_k,
\]
\[
\left(\begin{array}{cc}
2m - 1 \\
m - o
\end{array}\right) \cdot n(m(1 - p))^{m - \sum_{k \geq 1} n_k} \cdot \sum_{k \geq 1} n_k \sum_{k \geq 1} n_k \geq \sum_{k \geq 2} (k - 1) n_k,
\]

9
2.4 Even Powers of $A$

Having pinned down the dominant components of lacunary cycles, proceed with even powers of $A$ for $w \in (0, 1)$, $n^{w-1} \leq p \leq \frac{1}{2}$, $m \leq c \min((pn^{1-w})^{1/4} \log(\frac{pn^{1-w}}{n^{w-1}}), (np)^{1/5}, n^{w/4})$.

\[ \mathbb{E}[\tau(A^m)] \geq (np)^{2m} + 2m \cdot (np)^{2m-1} (1-p) + C_m n^{m+1} (p(1-p))^m (1-o(1)), \]

\[ \mathbb{E}[\tau(A^m)] \leq (np)^{2m} + 2m \cdot (np)^{2m-1} (1-p) + C_m n^{m+1} (p(1-p))^m (1 + C \cdot e), \]

(28) (29)

The third term is necessary when $m = 2$; note the first strictly dominates the rest solely when $m > 1$.
Suppose a fixed configuration has \( q \), \( q \), \( s \) contribution for is, by (24),

\[
\sum_{i= \{a_{j_1}, \ldots, a_{j_{2m-1}}, h_0\}} E[a_{j_1}a_{j_2} \ldots a_{j_{2m-1}}h_0],
\]

where \( a_{j_i} \in \{p, a_k\} \), consider the extremal lacunary cycles first. If no edge is erased, then the contribution is \( E[tr(A^{2m})] \), given by (12). If all edges are erased, then \( (np)^{2m} \) arises. It remains to see what occurs when there is a mix of both categories: it is shown that the more erased edges, the larger the overall contribution; in particular, the second term in (28) comes from cycles with all but two consecutive edges erased. Besides, as the moments of the centered trace are of interest, obtaining the primary contributors from random cycles (i.e., not all edges erased) is crucial.

**Case 1:** Suppose a fixed configuration has \( s \geq 1 \) erased segments (s minimal), which amount to a total of \( 2m - 2q - 1 \) edges (i.e., the endpoints of these segments are given). If \( q = 0 \), then all the expectations are 0; else \( q \geq 1 \). The erased edges with their interior vertices (i.e., those that are not endpoints of the segments) contribute \( n^{2m-2q-1}s^{2m-2q-1} \) (each edge gives a factor of \( p \), and there are \( 2m - 2q - 1 - s \) hidden vertices, arbitrary elements of \( \{1, 2, \ldots, n\} \)). What remains is a lacunary cycle with \( 2q + 1 \) edges, whose contribution is, by (24),

\[
O(n^{p+1}C_{q+1}^{(s)}) \cdot n^{2m-2q-1}s^{2m-2q-1} = O((n^{2m-1}p^{2m-1}C_{q+1}^{(s)}),
\]

The number of configurations with \( E \) edges and \( s \) segments is at most

\[
(2m)! \cdot \left( \frac{E}{s-1} \right)^s - 1
\]

because there are \((E-s+r-1)_{s-1}^n = \{(x_1, x_2, \ldots, x_s) : x_1 + \ldots + x_s = E, x_i \in \mathbb{N}\} \) possibilities for the lengths of the segments and at most \((2m)!^s\) ways to choose the left endpoints from \( \{1, \ldots, 2m\} \), from which the overall contribution for \( q, s \) fixed is

\[
O(n^{2m-1}p^{2m-1}C_{q+1}) \cdot (2m)!\left( \frac{2q}{s-1} \right)
\]

using

\[
\sum_{s \leq 2q} (2m)!\left( \frac{2q}{s-1} \right) \leq 2m(2m+1)^{2q} \leq (2m+1)^{2q+1},
\]

for given \( q \), this is at most

\[
O(n^{2m-1}p^{2m-1}C_{q+1}) \cdot (2m+1)^{2q+1}.
\]

**Case 2:** Suppose a fixed configuration has \( s = 1 \) erased segments (s minimal), which amount to a total of \( 2m - 2q \) edges (i.e., the endpoints of these segments are given) with \( 1 \leq q \leq m - 1 \). The erased edges with their interior vertices (i.e., those that are not endpoints of the segments) contribute \( n^{2m-2q-s}p^{2m-2q} \). The remainder is a lacunary cycle with \( 2q \) edges, giving at most

\[
n^{p+1}(p(1-p))^sC_{q+s} \cdot n^{2m-2q-s}p^{2m-2q} = C_{q+s}n^{2m-2q-s}(1-p)^q
\]

from (23) and

\[
\sum_{m_1, \ldots, m_{q+s}=m, \sum m_i \geq 0} C_{m_1} \ldots C_{m_{q+s}} = C_{m+s-1}.
\]

As above, the overall contribution is at most

\[
C_{q+s}n^{2m-2q-s}(1-p)^q \cdot (2m)!\left( \frac{2q}{s-1} \right)
\]

and summing over \( s \) yields at most

\[
C_{q+s}n^{2m-2q-s}(1-p)^q \cdot (2m+1)^{2q}.
\]

When \( q = 1 \), the dominant graphs are even cycles of length 2 with \( s = 1 \); this gives the second term in (28) (2m comes from labeling the first apparition of the unerased edge). All the other terms are dominated by this subfamily since \( m^5 \leq c_{10} \cdot np \), and

\[
\frac{n^{2m-1}p^{2m-1}C_{q+1}(2m+1)^{2q+1}}{n^{2m-1}p^{2m-1}(1-p)} = \frac{C_{q+1}(2m+1)^{2q+1}}{(np)^q(1-p)} \leq 8(2m+1) \cdot \left( \frac{(4m+1)^2}{np} \right)^q,
\]

\[
\frac{n^{2m-(1-p)^q}p^{2m-1}C_{q+1}(2m+1)^{2q}}{n^{2m-1}p^{2m-1}(1-p)} = \frac{(1-p)^{q-1}C_{q+1}(2m+1)^{2q}}{(np)^{q-1}} \leq 64(2m+1)^2 \cdot \left( \frac{64m+1}{np} \right)^{q-1}.
\]

11
2.5 Moments of $\text{tr}(A^{2m}) - \mathbb{E}[\text{tr}(A^{2m})]$

Consider now the moments of the centered trace: Carleman condition (lemma 2.3 in [4]) entails showing the $l\text{th}$, $l \geq 2$ moment converges to 0 if $2l > 1$, and $2l/(l-1)!$ if $2l$ suffices to conclude the asymptotic behavior of the ratio of interest is normal ($l$ is fixed). To do so, use the results from subsections 2.1-2.4 concerning the dominant components of the traces, and a variant of the gluing technique in [20].

Begin with $l = 2$:

$$\mathbb{E}[(\text{tr}(A^{2m}) - \mathbb{E}[\text{tr}(A^{2m})])^2] = \sum_{(i,j)} (\mathbb{E}[a_i^*a_j^*] - \mathbb{E}[a_i^*] \cdot \mathbb{E}[a_j^*]),$$  \hspace{1cm} (32)

where $*$ indicates the cycles of length $2m$ are lacunary. By independence, the contribution of $(i,j)$ is nonzero only if $1, j$ share an unerased edge. As in [20], merge this pair into a cycle $\mathcal{C}$ of length $4m$ by employing the first shared edge $e$ as a bridge from one to the other: let $i_{1-1} = j_{1-1}$, with $s$ minimal in this order (i.e., $s = \min \{1 \leq k \leq 2m, 31 \leq q \leq 2m, i_{k-1} = j_{k-1}\}$, $s = \min \{1 \leq q \leq 2m, j_{s-1} = i_{s-1}\}$, where only unerased edges are considered). Walk along $\mathcal{C}$ up to $i_{s-1}$, use it as a bridge to switch to $j_{s-1}$, traverse all of it, and get back to the rest of $i$ upon returning to $j_{s-1}$.

In [20], the two copies of this common edge are deleted, giving rise to a cycle of length $4m - 2$; however, doing so in the current situation is an issue if $m(e) = 3$ (the expectation of the remaining cycle is 0). The previous subsection yields the contribution of cycles of length $4m$ is dominated by those having exactly 2 erased edges (all have at least 2 erased by this procedure), hence, what remains is finding the pairs $(i,j)$ which are their preimages. Each must have one unerased edge $e = uv$, which is shared: once this is fixed in $2m$ possibilities, its position must be chosen in $1 (2m - 2)$ possibilities as it can be $(u,v)$ or $(v,u)$ when $v \neq u$; cycles with $u = v$ are negligible). This yields the first order term in (32) is $2(2m)^2(np)^{m-2}p(1-p)$, entailing the desired result for $l = 2$.

Consider now $l > 2$:

$$\mathbb{E}[(\text{tr}(A^{2m}) - \mathbb{E}[\text{tr}(A^{2m})])^2] = \sum_{(i,j,k_1, \ldots, k_l)} (\mathbb{E}[a_i^*a_{k_1}^* \ldots a_{k_l}^*] - \mathbb{E}[a_i^*] \cdot \mathbb{E}[a_{k_1}^*] \cdot \ldots \cdot \mathbb{E}[a_{k_l}^*]).$$ \hspace{1cm} (33)

Recall first the gluing strategy in [20] for high moments when the underlying distributions are symmetric and subgaussian: symmetry takes care of $2l + 1$; assume now $2l$. Construct a simple undirected graph $\mathcal{G}$ with vertices $1, 2, \ldots, k$ (cycles of length $m$), in which $i, j \in \mathcal{G}$ if only if $i, j$ are a shared edge. The contributors to (33) are tuples for which all connected components of $\mathcal{G}$ have size at least two (else the expectation is 0 by independence), and the contributions of the components are independent as no two share an edge. Glue next $1, k_2, \ldots, k_l$ into one cycle in a series of steps. The key observation is that the first-order term is given by tuples whose associated graphs contain $l/2$ components, each of size 2: this generates $(l-1)!$, which is the number of possibilities of splitting $\{1, 2, \ldots, l\}$ in $l/2$ unordered pairs. To merge the tuples into one, employ two types of unions. A regular step is gluing two cycles along their first shared edge $e$ and erasing it: this works well unless $e$ is the only edge shared with other cycles in the connected component, in which case erasing $e$ disconnects it. A modified step is designed to account for this last possibility; consider all cycles in which $e$ appears, and let $(e)$ be their number; if $2l(e)$, then merge them into an even cycle with no copy of $e$, by using it $l(e)/2$ times as a bridge (as for $l = 2$); else $2l(e) + 1$, and since $e$ has even multiplicity in the union of the cycles (otherwise the contribution is 0), there is one among them with at least two copies of $e$; use a regular step and erase a copy of $e$ from it and another from its pair; this reduces the number of cycles (two turn into one) and repeat this procedure until each connected component of $\mathcal{G}$ is transformed into one cycle $1$. If $\mathcal{G}$ is connected, then $\mathcal{G}$ has length $l - m + q, l \leq q \leq 2l, q = 2l + 2q + \ldots + 2l + 2g$, where $2g$ is the number of regular steps and $(2l)/l \leq g \leq 2l$, the numbers of paths combined in the modified steps when their number is even. The authors in [20] finish by proving each connected component with $l > 2$ vertices is negligible, whereas for $l = 2$, there are about $\frac{1}{4}$ of them (up to normalization), implying the claim.

In the current situation, proceed with a simplified version of this gluing that does not throw away the shared edges, which are unerased, in the process. For a connected component in $\mathcal{G}$ of size $L$, merge its vertices in $L - 1$ regular steps (to make it canonical, always add to the previous outcome the vertex of smallest index). The final cycle has length $2Lm$, and at least $L - 1$ pairs of copies of unerased edges at distance $2m - 1$ (i.e.,
there are 2m − 1 edges between them traversing the cycle one way, and at least 2m − 1 going the other way). Denote by α the number of pairs with at least one edge among the 2m − 1 between them erased, and let β = L − 1 − α. Inequalities (30) and (31) in subsection 2.4 entail such graphs with exactly t unerased edges contribute at most

\[(np)^{2mL−\lfloor(t+1)/2\rfloor}C_m(2mL + 1)^t\).

Consider now the preimages of this mapping: for any cycle of length 2mL, there are at most \(P(L)m^L\) possibilities for splitting it into L cycles by reversing the gluing described above and labeling them (by induction on \(L\)). Hence, the total over all configurations of connected components is at most

\[
\left(\max_{2 \leq t \leq l} P(L)^t \sum_{(L_i,t_i) \sum L_i = t_i \geq L} \prod 64^t (np)^{2mL_{i−\lfloor(t+1)/2\rfloor}}(2mL_{i} + 1)^t m_{L_{i}−\lfloor(t+1)/2\rfloor} \leq \right.
\]

\[
\left(\max_{2 \leq L \leq l} P(L)^{\lfloor L/2 \rfloor} \cdot (256n^2)^{\lfloor L/2 \rfloor} \cdot (np)^{2mL} \sum_{(L_i,t_i) \sum L_i = t_i \geq L} \prod (np)^{−\lfloor(L+1)/2\rfloor} \cdot \right)
\]

Note \(t_i \geq 2m\beta + \alpha_i\); each pair of edges underlying \(\beta_i\) (i.e., with all the \(2m\) edges between them unerased) generates a cycle (among the original ones) whose edges are unerased (by the procedure), and each of the remaining \(\alpha_i\) give rise at least one unerased edge (the one belonging to the cycle that is merged): the positions of any two such edges are pairwise distinct in the final cycle, whereby the claim ensues. Moreover, either the final cycle contains at least \(2mL_{i} = 2m\) unerased edges or each pair behind \(\alpha_i\) generates a factor of at most \(m^{2mC(L)^m}L_{i}−\alpha_i−1\) because there exists an erased edge between their components no matter how the cycle is traversed, implying the two copies belong to different segments (if it is not possible to glue the remaining segments in a collection of pairwise edge disjoint even cycles, then there is overall some polynomial decay in \(n\); else, the cycle containing these edges is formed gluing at least two segments; then by erasing these two copies, what remains is still an even cycle, for which the claim follows using (12)). In the latter case, this gives a factor of at most \(m^{mC(L)^m}L_{i}−n−\alpha_i\), while in the former the contribution is at most

\[(np)^{2mL_{i}−(mL_{i}−m)m(L_{i}−1)^{mL_{i}−(mL_{i}−m)}C_{mL_{i}−1}(nL_{i})^{mC(L_{i})}}\]

In the former, the final bound for given \(\alpha_i, \beta_i\) is at most \((np)^{2mL_{i}}\) times

\[n^{−n}(np)^{−(2mL_{i}−1)^{2}−(2mL_{i})}C_{mC(L_{i})}^{(nL_{i})^{mC(L_{i})}} \leq n^{−L_{i}−(mL_{i}−1)^{2}}C_{mC(L_{i})}^{(nL_{i})^{mC(L_{i})}} \]

recall the main order term is of size

\[
(np)^{2mL_{i}−1/p^{L_{i}} = (np)^{2mL_{i}−1/p^{L_{i}}}}\]

Notice \((np)^{−\lfloor(L+1)/2\rfloor} \leq (np)^{−\sum L_i/2} = (np)^{−L_i/2}\), since \(n^{\beta_i}(np)^{−(mL_{i}−1)^{2}/2} = \sqrt{np} (\sqrt{np})^{−(mL_{i}−1)^{2}/2}\), some polynomial decay in \(n\) is generated if \(\sum \beta_i > 0, m \geq m(w)\). This is also the case if any \(L_i > 2\) as in these situations \(L_i − 1 > L_i/2\); what remains is the case \(L_i = 2\): the cycles are paired, and as it was previously noticed, the dominant configurations have each pair sharing one unerased edge and all the rest erased: the case \(l = 0\) concludes the claim: note the factors in \(m\) can always be absorbed by a power of \(n\) as \(m^{mC(L)} \leq n^{c(m)}\) for \(m \leq \frac{\log \log n}{\log \log \log n}\) and \(c(m) > 0\) sufficiently small (the summation over \((\alpha_i, \beta_i)\) generates at most a factor of \(L^L\) since \(\alpha_i + \beta_i = L − 1 \leq L − 1\)). This completes the analysis of higher moments of the centered trace and the proof of Theorem 1.

3 CLT for \(\lambda_1(A)\)

This section contains the proof of Theorem 2, in which (6) plays a fundamental role. Take \(\mathbb{A} = \frac{1}{\sqrt{p}}A\), (6) can be restated as

\[\frac{\text{tr}(\mathbb{A}^{2m})}{2m \cdot (n\sqrt{p})^{2m−1} \cdot \sqrt{1−p}} = \mathbb{N}(0, 2)\]

and the claim of Theorem 2 becomes

\[\frac{1}{\sqrt{A−p}}(\lambda_1(\mathbb{A}) − \mathbb{E}[\lambda_1(\mathbb{A})]) = \mathbb{N}(0, 2)\].
In what follows, assume $\frac{1}{n \log n} \leq m \leq \frac{c(n, \epsilon_0)}{n \log n}$ unless otherwise stated. (39) below and Slutsky’s lemma entail $n \sqrt{\tilde{p}}$ can be replaced by $E[\lambda_1(A)]$ in (34):

$$\frac{tr(A^2m) - E[tr(A^2m)]}{2m \sqrt{1 - p \cdot (E[\lambda_1(A)])^{2m-1}}} \Rightarrow N(0, 2).$$

Let $\xi = \lambda_1(A) - E[\lambda_1(A)]$ :

$$\frac{tr(A^2m) - E[tr(A^2m)]}{2m \sqrt{1 - p \cdot (E[\lambda_1(A)])^{2m-1}}} = \frac{\lambda_2^2m(A) - E[\lambda_2^2m(A)] + \sum_{k \geq 2} \lambda_k^2m(A) - E[\sum_{k \geq 2} \lambda_k^2m(A)]}{2m \sqrt{1 - p \cdot (E[\lambda_1(A)])^{2m-1}}} =$$

$$= \frac{\frac{\xi^2}{2m} + (E[\lambda_1(A)])^{2m} - E[\lambda_1(A)]^{2m}}{2m \sqrt{1 - p \cdot (E[\lambda_1(A)])^{2m-1}}} + \frac{\sum_{k \geq 2} \lambda_k^2m(A) - E[\sum_{k \geq 2} \lambda_k^2m(A)]}{2m \sqrt{1 - p \cdot (E[\lambda_1(A)])^{2m-1}}}.$$

For $m \geq m(\epsilon_0),$

$$\frac{\sum_{k \geq 2} \lambda_k^2m(A) - E[\sum_{k \geq 2} \lambda_k^2m(A)]}{2m \cdot (E[\lambda_1(A)])^{2m-1}} = o_p(1);$$

Recall $A = \sqrt{p}w_+^t + \frac{1}{\sqrt{p}} A, w_n = [11 \ldots 1]^t \in \mathbb{R}^n$: Weyl’s inequalities yield for $2 \leq j \leq n,$

$$\lambda_j(\sqrt{p}A) \leq \lambda_1(\sqrt{p}A) \leq \lambda_j(\sqrt{p}A),$$

whereby

$$\sum_{j \geq 2} \lambda_j^2m(A) \leq 2tr(\sqrt{p}A^{2m}),$$

which together with Markov inequality and (12) concludes (36) (recall (39)): for $t > 0,$

$$\mathbb{P}(tr(\sqrt{p}A^{2m}) \geq t(n \sqrt{p})^{2m-1}) \leq \frac{2C_m(1 - p)^m n^{m+1}}{t(n \sqrt{p})^{2m-1}} \leq \frac{2 \cdot 4^n n^2}{t(n \sqrt{p})^m} = o(1).$$

Return to (35): (36) entails for $m(\epsilon_0) \leq m \leq np,$

$$\frac{tr(A^2m) - E[tr(A^2m)]}{2m \sqrt{1 - p \cdot (E[\lambda_1(A)])^{2m-1}}} + \frac{(E[\lambda_1(A)])^{2m} - E[\lambda_1(A)]^{2m}}{2m \sqrt{1 - p \cdot (E[\lambda_1(A)])^{2m-1}}} + o_p(1).$$

Lemma 3 yields $\xi = o_p(E[\lambda_1(A)]n^{-\alpha/4})$, whereby

$$\frac{(E[\lambda_1(A)])^{2m} - E[\lambda_1(A)]^{2m}}{2m \sqrt{1 - p \cdot (E[\lambda_1(A)])^{2m-1}}} = \frac{\xi}{\sqrt{1 - p}} \leq 2 \sum_{k \geq m-1} \binom{2m}{k} (E[\lambda_1(A)])^{-k+1} |\xi|^k \leq 2^{2m}E[\lambda_1(A)]^{-2} |\xi|^2 \leq 2^{2m}E[\lambda_1(A)]^{-2} \xi^2 = o_p(\xi),$$

from which Slutsky’s lemma gives the desired result as long as

$$E[\lambda_1^{2m}(A)] - E[\lambda_1(A)]^{2m} \geq \frac{1}{2m \cdot (E[\lambda_1(A)])^{2m-1}} \Rightarrow 0.$$

Subsection 3.1 derives two concentration results for $\lambda_1(A)$; however, these solely imply the ratio above is $o(1)$ plus some terms depending on few moments of $\lambda_1(A)$. Subsection 3.2 uses this simplification and (37) for $m(\epsilon_0) \leq m \leq c(n, \epsilon_0)$ with $\lim_{n \to \infty} c(n, \epsilon_0) = \infty$ to derive the eigenvalue CLT.

### 3.1 Concentration of $\lambda_1(A)$

Begin with a concentration result for $\lambda_1(A)$.

**Lemma 3.** There exists $c > 0$ such that for $t \geq 4(\epsilon_1 + \epsilon_2), E[\lambda_1(A)] \in [np(1 - \epsilon_1), np(1 + \epsilon_2)]$, $\epsilon_1, \epsilon_2 \in (0, \frac{1}{4})$, and $w \in (0, 1), n^{\alpha - 3} \leq p \leq \frac{1}{2}, m \in \mathbb{N}, m \leq c \min(n^{\alpha/4}, \sqrt{pm^{1-\alpha}}),$

$$\mathbb{P}\left(\frac{\lambda_1(A)}{E[\lambda_1(A)]} - 1 \geq t\right) \leq \frac{4 \cdot 16^m n}{(np)^m t^{2m}}.$$  

(38)
Proof. It suffices to show each tail has probability at most \( \frac{24n^m}{m^{w-1}} \). Begin on the right:

\[
\Pr \left( \frac{\lambda_1(A)}{E[\lambda_1(A)]} - 1 \geq t \right) \leq \Pr(||A|| \geq (1 + t)E[\lambda_1(A)]) \leq \Pr(||A|| \geq np(1 + t)(1 - \epsilon_1)).
\]

Since

\[
||A|| = ||pv_nv_n^T + \tilde{A}|| \leq ||pv_nv_n^T|| + ||\tilde{A}|| = np + ||\tilde{A}||,
\]
an upper bound for

\[
\Pr(||\tilde{A}|| \geq np(t(1 - \epsilon_1) - \epsilon_1)) \leq \Pr(||\tilde{A}|| \geq np(t/2))
\]
is enough. Use (12): for \( c > 0 \) sufficiently small and \( m \leq c \min \{n^{w/4}, \sqrt{m^{1-w}}\} \),

\[
\Pr(||\tilde{A}|| \geq np(t/2)) \leq \frac{2C_m(p(1 - p))^m}{(np/2)^{2m}} \leq \frac{2 \cdot 16mn}{(np)^{m/2}}.
\]

Similarly, for the left tail,

\[
\lambda_1(A) = \lambda_1(pv_nv_n^T + \tilde{A}) \geq \lambda_1(pv_nv_n^T) - ||\tilde{A}|| = np - ||\tilde{A}||,
\]

whereby

\[
\Pr \left( \frac{\lambda_1(A)}{E[\lambda_1(A)]} - 1 \leq -t \right) \leq \Pr(\lambda_1(A) \leq np(1 + \epsilon_2)(1 - t)) \leq \Pr(||\tilde{A}|| \geq np(t + \epsilon_2) - \epsilon_2)) \leq \Pr(||\tilde{A}|| \geq np(t/2)).
\]

Under the binomial expansion given by \( \lambda_1(\tilde{A}) = E[\lambda_1(\tilde{A})] + (\lambda_1(\tilde{A}) - E[\lambda_1(\tilde{A})]) \), Lemma 3 and (39) yield most of the terms in

\[
\frac{E[\lambda_1^2(\tilde{A})] - (E[\lambda_1(\tilde{A})])^{2m}}{2m \cdot E[\lambda_1^2(\tilde{A})]^{2m-1}}
\]

are negligible as \( n \rightarrow \infty \).

Lemma 4. There exists \( c > 0 \) such that for \( w \in (0, 1) \), \( n^{w-1} \leq p \leq \frac{1}{2} \), \( m(w) \leq m \leq c \min \{n^{w/4}, \sqrt{m^{1-w}}\} \),

\[
\lim_{n \rightarrow \infty} n^2p \cdot \frac{E[\lambda_1(A)]}{E[\lambda_1(A)] - 1} 2m = 0.
\]

Proof. Take \( \delta > 0 \) (chosen subsequently): Lemma 3 for \( m_0 \leq c \min \{n^{w/4}, \sqrt{m^{1-w}}\} \), \( m_0 \geq 2m \), and \( \delta^{1/2m} > 4(\epsilon_1 + \epsilon_2) \) yields

\[
E[\frac{\lambda_1(A)}{E[\lambda_1(A)]} - 1]^{2m} \leq \delta + E[\frac{\lambda_1(A)}{E[\lambda_1(A)]} - 1]^{2m} (\lambda_1(A) - 1|_{\lambda_1(A) = \delta^{1/2m}}) \leq \delta + 4 \cdot \frac{16mn}{m_0} \int_\delta^{\delta^{1/2m}} \frac{1}{m_0/d} \delta^{1-m_0/m} - \frac{4 \cdot 16mn}{m_0} \delta^{1-m_0/m} m_0/ - m_0/m - 1.
\]

For \( \delta = \frac{1}{np \log w} \),

\[
n^2p \cdot \frac{E[\lambda_1(A)]}{E[\lambda_1(A)] - 1} 2m \leq \frac{1}{\log n} + 4 \frac{16mn}{(np)^m} \frac{n^2p \cdot \delta^{1-m_0/m}}{m_0/m - 1} \leq \frac{1}{\log n} + 4n \cdot \frac{(16n^2p \log n)^{1/m}}{np} = o(1)
\]

for \( m \geq m(w) \).

The result follows by letting \( \epsilon_2 = \frac{\delta}{np} \), \( \epsilon_1 = O(\frac{1}{np}) \), increasing \( m(w) \) to ensure \( 4(\epsilon_1 + \epsilon_2) < \delta^{1/2m} \), and showing

\[
E[\lambda_1(A)] \in [np(1 - \epsilon_1), np(1 + \epsilon_2)].
\]

(39)

For the upper bound, use (28), in which \( (np)^{2m-1} > 2C_m n^{m+1}(p(1 - p))^m \) for \( m \geq m(w) \), and so

\[
E[\lambda_1(A)] \leq (E[tr(A^{2m})])^{1/2m} \leq [(np)^{2m} + 4m \cdot (np)^{2m-1}]^{1/2m} \leq np + 2,
\]

whereas for the lower,

\[
\lambda_1(A) \geq \frac{1}{n} v_n^T Av_n = np + \frac{1}{n} \sum_{1 \leq i, j \leq n} \tilde{a}_{ij}.
\]

15
and Hoeffding’s inequality (theorem 2.6.2 in Vershynin [25]) with
\[ |\tilde{a}_{11}| |v_1| := \inf\{ t > 0, E[\exp(\tilde{a}_{11}^2/t^2)] \leq 2 \} \leq \frac{1 - p}{\sqrt{c' \cdot \log p}}, \]
since for \( t = \frac{1 - p}{\sqrt{c' \cdot \log p}} \) \( c' \) sufficiently small and any \( p \leq \frac{1}{2} \),
\[ E[\exp(\tilde{a}_{11}^2/t^2)] = pe^{(1-p)t^2/2} + (1-p)e^{t^2/2} = p^{1-c'} + (1-p)e^{t^2/2} \leq 2, \]
yielding
\[ P\left( \frac{1}{n} \sum_{1 \leq i, j \leq n} \tilde{a}_{ij} \leq -t \right) \leq 2 \exp\left( -\frac{Cn^2t^2}{(1-p)t^2} \right) \leq 2 \exp\left( -C't^2 \log \frac{1}{p} \right), \]
from which
\[ E[(\lambda_1(A) - np)\_+ \leq 2 \int_0^{\infty} \exp(-C't^2 \log \frac{1}{p})dt \leq 2 \int_0^{\infty} \frac{1}{1 + C't^2 \log \frac{1}{p}}dt = O((\log \frac{1}{p})^{-1/2}) = O(1). \]

3.2 Asymptotic Law of \( \lambda_1(A) \)
Return to (35): given (37), consider the first ratio on its right-hand side.

**Lemma 5.** There exists \( c > 0 \) such that if \( w \in (0, 1), n^{w-1} \leq p \leq \frac{1}{2}, m \leq c \min(n^{w/4}, \sqrt{pn^{1-w}}) \), then
\[ \frac{(\xi + E[\lambda_1(\tilde{A}))])^{2m} - E[(\lambda_1(\tilde{A}))])^{2m}}{2m(E[\lambda_1(A)])^{2m-1}} - \xi = o_p(1). \]

**Proof.** Hoeffding’s inequality (theorem 2.6.2 in Vershynin [25]) gives
\[ P\left( \frac{1}{n} |v_1^T \tilde{A}v_n| \geq t \right) \leq 2 \exp(-C't^2 \log \frac{1}{p}) \]
(similarly to (40)). Take \( u = \alpha \cdot \frac{1}{\sqrt{m}}v_n + \sqrt{1 - \alpha^2}v_1 \) with \( ||u|| = 1, v_n \cdot v_1 = 0, Au = \lambda_1(A)u \). Then
\[ \lambda_1(A) = u^TAu = \alpha p + \frac{\alpha}{n} V_1^TAu_n + \sqrt{1 - \alpha^2} \cdot (v_1)^T \tilde{A}v_1. \]

With high probability
\[ |\lambda_1(A)| \leq \alpha p + \sqrt{1 - \alpha^2} \cdot ||\tilde{A}|| + \alpha \leq np + 1 \]
as (12) implies for \( m_0 \geq m(w) \),
\[ P(||\tilde{A}|| \geq (np)^{1/4}) \leq 2C_m(p(1-p))^{m_0} np^{m_0+1}/(np)3^{m_0/2} \leq \frac{2n^{-4m_0}}{(pn^{-m_0/2})} = o(1). \]

By the same token, \( \lambda_1(A) \geq np - 1 \) with high probability, whereby using (39) and letting \( C > 0 \) be sufficiently large,
\[ \sqrt{p}\xi \leq |\lambda_1(A) - np| + |np - E[\lambda_1(A)]| \leq C, \]
from which for \( k \geq 2 \),
\[ (E[\lambda_1(\tilde{A})])^{k-1} \cdot |\xi|^k \leq (np/2)^{-k+1}(p/C^2)^{-k/2} = n^{-k+1}C^{-k+1}2^{k-1} \sqrt{p} \leq n^{-k+1}(2C)^k. \]
The claim follows from \( (2m)^k \leq (2n)^k, m^2 = o(n) \),
\[ \frac{(\xi + E[\lambda_1(\tilde{A}))])^{2m} - E[(\lambda_1(\tilde{A}))])^{2m}}{2m(E[\lambda_1(A)])^{2m-1}} - \xi = \frac{1}{2m} \sum_{2 \leq k \leq m} \left( \frac{2m}{k} \right) (E[\lambda_1(\tilde{A})])^{k-1} |\xi|^k. \]
Set \( w = \epsilon_0 / 2 \). Lemma 5 transforms (37) into

\[
\frac{\text{tr}(A^{2m}) - E[\text{tr}(A^{2m})]}{2m \sqrt{1 - p} \cdot (E[|\lambda_1(A)|])^{2m-1}} = \frac{\xi}{\sqrt{1 - p}} + \frac{E[|\lambda_1(A)|^{2m}] - (E[|\lambda_1(A)|])^{2m}}{2m \sqrt{1 - p} \cdot (E[|\lambda_1(A)|])^{2m-1}} + o_p(1).
\]

Lemma 4, (39), and \( E[X] \leq \sqrt{E[X^2]} \) imply for \( m(w) \leq m \leq C_1(n, w) \), \( \lim_{n \to \infty} C_1(n, w) = \infty \),

\[
\frac{E[|\lambda_1(T)|]}{2m \sqrt{1 - p}} - \frac{E[|\lambda_1(T)|^{2m}] - (E[|\lambda_1(T)|])^{2m}}{2m \sqrt{1 - p} \cdot (E[|\lambda_1(T)|])^{2m-1}} = o(1) + \frac{E[|\lambda_1(T)|]}{2m \sqrt{1 - p}} \sum_{2 \leq k < m(w)} \left( \frac{2m}{k} \right) \cdot E[|\lambda_1(T)|^k].
\]

Denote the second term by \( P(m, n, p) : \) note this can be written as

\[
P(m, n, p) = \sum_{2 \leq k < m(w)} \frac{(2m - 1)(2m - 2) \cdots (2m - k + 1)}{k!} \cdot c_k(n, p),
\]

and thus for \( m(w) \leq m \leq C_1(n, w) \),

\[
\frac{\xi}{\sqrt{1 - p}} + P(n, m, p) \Rightarrow N(0, 2).
\]

This entails for \( m(w) \leq m \leq C_1(n, w) - 1 \), \( P(n, m + 1, p) - P(n, m, p) \to 0 \) : for any \( C > 0 \),

\[
P(P(n, m + 1, p) - P(n, m, p) > C) \leq P\left( \left( P(n, m + 1, p) + \frac{\xi}{\sqrt{1 - p}} \right) - \left( P(n, m, p) + \frac{\xi}{\sqrt{1 - p}} \right) > C \right)\]

\[
\leq P\left( P(n, m + 1, p) + \xi > C/2 \right) + P\left( P(n, m + 1, p) + \xi < -C/2 \right) \leq (1 + e(C))P(|N(0, 1)| > C/2) < 1.
\]

Let \( P_1(m, n, p) = P(m + 1, n, p) - P(m, n, p) \), and consider it as a polynomial in \( m \). After \( d_1(w) \) iterations of the type \( P_{k+1}(m, n, p) = P(m + 1, n, p) - P(m, n, p) \), what is obtained is \( c_{m(w)-1}(n, p) - d_2(w) \) for \( d_2(w) > 0 \); this entails \( c_{m(w)-1} \to 0 \), and hence the last term in (42) converges to 0 for any \( m = m(w) \). Proceed similarly with all the remaining coefficients \( c_k(n, p), 2 \leq k < m(w) - 1 \) : since all appear in \( P_1 \), it follows all converge to 0, whereby \( P(m(w), n, p) \to 0 \). The desired result then ensues.

References

1. N. Alon. *Spectral techniques in graph algorithms*, Lecture Notes in Computer Science, Springer, Berlin, 296 – 215, 1998.

2. J. Alt, R. Ducatez, and A. Knowles. *Extremal eigenvalues of critical Erdős-Rényi graphs*, Ann. Probab. 49(3) : 1347 – 1401, 2021.

3. A. Auffinger, G. Ben-Arous, and S. Péché. *Poisson convergence for the largest eigenvalues of heavy tailed random matrices*, Ann. Inst. H. Poincaré Probab. Statist., Vol. 45, No. 3, 589 – 610, 2009.

4. Z. Bai, and J. Silverstein. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, Springer Series in Mathematics, Second Edition, 2010.

5. F. Benaych-Georges, and S. Péché. *Localization and Delocalization for Band Matrices*, Ann. Inst. H. Poincaré Probab. Statist., Vol. 50, No. 4, 1385 – 1403, 2014.

6. F. Benaych-Georges, C. Bordenave, and A. Knowles. *Largest eigenvalues of sparse inhomogeneous Erdős-Rényi graphs*, The Annals of Probability, Vol. 47, No. 3, 1653 – 1676, 2019.

7. F. Benaych-Georges, C. Bordenave, and A. Knowles. *Spectral radius of sparse random matrices*, Ann. Inst. H. Poincaré Probab. Statist. 56(3) : 2141 – 2161, 2020.

8. B. Bollobás, S. Janson, and O. Riordan. *The phase transition in inhomogeneous random graphs*, Random Structures and Algorithms, 31 : 3 – 122, 2007.

9. S. Diaconu. *More Limiting Distributions for Eigenvalues of Wigner Matrices*, arXiv:2203.08712, 2022.

10. S. Diaconu. *Finite Rank Perturbations of Heavy-Tailed Random Matrices*, arXiv:2208.02756, 2022.
11. L. Erdos, A. Knowles, Horng-Tzer Yau, and J. Yin. *Spectral Statistics of Erdos-Renyi Graphs I: Eigenvalue Spacing and the Extreme Eigenvalues*, The Annals of Probability, Vol. 41, No. 3B, 2279 – 2375, 2013.

12. L. Erdos, A. Knowles, Horng-Tzer Yau, and J. Yin. *Spectral Statistics of Erdos-Renyi Graphs II: Eigenvalue Spacing and the Extreme Eigenvalues*, Communications in Mathematical Physics, Volume 3, 14, 587 – 640, 2012.

13. U. Feige, and E. Ofek. *Spectral techniques applied to sparse random graphs*, Random Structures and Algorithms, 27, 251 – 275, 2005.

14. D. Féral, and S. Péché. *The largest eigenvalue of rank one deformation of large Wigner matrices*, Communications in Mathematical Physics, Volume 272, 185 – 228, 2007.

15. Z. Furedi, and J. Komlos. *The Eigenvalues of Random Symmetric Matrices*, Combinatorica 1 (3), 233 – 241, 1981.

16. J. Huang, B. Landon, and H.-T. Yau. *Transition from Tracy–Widom to Gaussian fluctuations of extremal eigenvalues of sparse Erdos–Rényi graphs*, The Annals of Probability, Vol. 48, No. 2, 916–962, 2020.

17. M. Krivelevich, and B. Sudakov. *The Largest Eigenvalue of Sparse Random Graphs*, Combinatorics, Probability and Computing, 12, 61 – 72, 2003.

18. J. Lee, and K. Schnelli. *Local law and Tracy–Widom limit for sparse random matrices*, Probability Theory and Related Fields, Volume 171, 543 – 616, 2018.

19. L. Lu, and X. Peng. *Spectra of edge-independent random graphs*, Electronic Journal of Combinatorics, 20(4), 2013.

20. Ya. Sinai, and A. Soshnikov. *Central Limit Theorem for Traces of Large Random Symmetric Matrices With Independent Matrix Elements*, Bol. Soc. Brasil. Mat., Vol. 29, No. 1, 1 – 24, 1998.

21. Ya. Sinai, and A. Soshnikov, *A Refinement of Wigner’s Semicircle Law in a Neighborhood of the Spectrum Edge for Random Symmetric Matrices*, Functional Analysis and Its Applications, Vol. 32, No. 2, 1998.

22. A. Soshnikov, *Universality at the edge of the spectrum in Wigner random matrices*, Comm. Math. Phys., Vol. 207, No. 3, 697 – 733, 1999.

23. T. Tao and V. Vu. *Random matrices: Universality of local eigenvalue statistics*, Acta Math., 206, 127 – 204, 2011.

24. K. Tikhomirov, and P. Youssef. *Outliers in spectrum of sparse Wigner matrices*, Random Structures and Algorithms, Volume 58, Issue 3, 517 – 605, 2021.

25. R. Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*, https://www.math.uci.edu/~rvershyn/papers/HDP-book/HDP-book.pdf.

26. V. Vu. *Spectral norm of random matrices*, Combinatorica, 27(6) : 721 – 736, 2007.