Necessary and Sufficient Conditions for Almost Sure Convergence of the Largest Eigenvalue of a Wigner Matrix

Article *in* The Annals of Probability · October 1988

DOI: 10.1214/aop/1176991594 · Source: OAI

Z. D. Bai
Northeast Normal University

The Annals of Probability
1988, Vol. 16, No. 4, 1729–1741

NECESSARY AND SUFFICIENT CONDITIONS FOR ALMOST SURE CONVERGENCE OF THE LARGEST EIGENVALUE OF A WIGNER MATRIX

BY Z. D. BAI AND Y. Q. YIN[1]

University of Pittsburgh and University of Arizona

Let $W = (X_{ij};\ 1 \le i, j < \infty)$ be an infinite matrix. Suppose W is symmetric, entries on the diagonal are iid, entries off the diagonal are iid and they are independent. Then it is proved that the necessary and sufficient conditions for $\lambda_{\max}((1/\sqrt{n})W_n) \to a$ a.s. are (1) $E(X_{11}^+)^2 < \infty$; (2) $EX_{12}^4 < \infty$; (3) $EX_{12} \le 0$; (4) $a = 2\sigma$, $\sigma^2 = EX_{12}^2$. Here $W_n = (X_{ij};\ 1 \le i, j \le n)$.

1. Introduction. We will call matrix $W = (X_{ij})$ a Wigner matrix if it satisfies the following conditions:

(1.1) symmetric;

(1.2) entries above the main diagonal are iid random variables;

(1.3) entries on the diagonal are iid random variables;

(1.4) diagonal entries are independent of nondiagonal entries.

Wigner (1958) studied this kind of random matrix. He established the semicircle law. Juhász (1981) and Füredi and Komlós (1981) studied the asymptotic properties of the largest eigenvalues for symmetric random matrices. They assume the existence of moments of all orders. Sometimes they assume the uniform boundedness of entries.

In this paper, we confine ourselves to Wigner matrices and get the necessary and sufficient conditions for the convergence of the largest eigenvalue of a Wigner matrix.

Throughout this paper, we assume the following conditions are true. $W = \{X_{ij};\ 1 \le i < \infty,\ 1 \le j < \infty\}$ is an infinite matrix satisfying (1.1)–(1.4). And $W_n = (X_{ij}\colon 1 \le i \le n, 1 \le j \le n)$ is the $n \times n$ submatrix of W. The n eigenvalues of a symmetric $n \times n$ matrix A will be denoted

$$\lambda_1(A) \ge \lambda_2(A) \ge \cdots \ge \lambda_n(A).$$

Sometimes we write $\lambda_{\max}(A)$ instead of $\lambda_1(A)$.

The main purpose of this paper is to prove Theorem A.

Received March 1987; revised February 1988.

[1]Work of this author partially supported by NSF Grant DMS-87-03090.

AMS 1980 *subject classifications*. Primary 60F99; secondary 62E20.

Key words and phrases. Random matrix, Wigner matrix, largest eigenvalue, semicircle law.

Theorem A. *$\lambda_1((1/\sqrt{n})W_n)$ tends to a finite constant a as $n \to \infty$ with probability* 1, *if and only if*

$$E(X_{11}^{+})^2 < \infty; \tag{1.5}$$

$$EX_{12}^4 < \infty; \tag{1.6}$$

$$a = 2\sigma, \qquad \sigma^2 = EX_{12}^2; \tag{1.7}$$

$$EX_{12} \le 0. \tag{1.8}$$

2. Some graph theory. In order to prove Theorem C of Section 3, we need some graph theory. Let $e_1, \dots, e_{2k}$ be edges, $v_1, v_2, \dots, v_{2k}$ be vertices. v_i is the initial of e_i and v_{i+1} is the terminal of e_i, $i = 1, \dots, 2k$, but we define $v_{2k+1} = v_1$. We also assume $v_i \neq v_{i+1}$ for $i = 1, \dots, 2k$. Sometimes we write $v_i v_{i+1}$ instead of e_i.

If two sets $\{v_i, v_{i+1}\}$ and $\{v_j, v_{j+1}\}$ are equal, we say that e_i and e_j coincide. But the edges are always regarded as distinct, even if they are coincident.

A sequence $\{e_a, e_{a+1}, \dots, e_b\}$ of consecutive edges is called a chain. Sometimes we use $v_a v_{a+1} \cdots v_{b+1}$ to denote a chain; it is just the chain $\{e_a, e_{a+1}, \dots, e_b\}$.

If in the chain $v_1 v_2 \cdots v_b$, an edge e_a $(a < b)$ does not coincide with any other edge in this chain, then we say that e_a is single up to v_b. If v_{a+1} is distinct from $v_1, \dots, v_a$ we say that $e_a \in T_1$ or e_a is an innovation. If e_a is single up to v_{a+1} but $e_a \notin T_1$, we say that $e_a \in T_2$. If $e_a \in T_1$ and e_b $(b > a)$ is the first one which coincides with e_a, then e_b is said to belong to T_3. T_4 is the complement of $T_1 \cup T_3$. So, $T_2 \subset T_4$.

Lemma 2.1. *Let $1 \le a \le b < c$, $v_b v_{b+1}$ be single up to v_c and $v_c = v_a$. Then in the chain $v_b v_{b+1} \cdots v_c$ there is a T_2 edge. If in addition we also have*

$$v_a = v_b = v_c, \tag{2.1}$$

$$v_{a-1} v_a \in T_1, \text{ single up to } v_c, \tag{2.2}$$

$$\text{in } v_a \cdots v_b \text{ there is no } T_1 \text{ edge, single up to } v_b, \tag{2.3}$$

then the first edge $v_{c'-1} v_{c'}$ with $c' > b + 1$ and $v_{c'} = v_c$ is a T_4 edge.

Proof. We assume $c > b + 1$. Otherwise, $e_b \in T_2$. Let $v_{c'}$ be the first vertex with $c' > b + 1$ and $v_{c'} = v_c$. We assert that $v_{c'-1} v_{c'} \in T_2$.

At first we see that $v_{c'-1} v_{c'}$ cannot coincide with any edge in the chain $v_b v_{b+1} \cdots v_{c'-1}$. If it is single up to $v_{c'}$, it is already in T_2. If $v_{c'-1} v_{c'}$ coincides with an edge in $v_1 \cdots v_b$, then $v_b v_{b+1} \cdots v_{c'-1}$ is in the same situation as $v_b \cdots v_c$ but with shorter length and induction hypothesis applies.

Suppose $v_{a-1} v_a \in T_1$ is single up to v_c, $v_a = v_b = v_c$ and in the chain $v_a \cdots v_b$ there is no T_1 edge single up to v_b, $v_{c'-1} v_{c'} \notin T_1$. If $v_{c'-1} v_{c'} \in T_3$, it must coincide with a T_1 edge before v_{a-1}. Thus $v_{c'}$ must equal some v_d with $d < a$. But $v_d = v_{c'} = v_a$, contradicting $v_{a-1} v_a \in T_1$. Therefore, $v_{c'-1} v_{c'} \in T_4$. □

LEMMA 2.2. *If in the chain $v_1 \cdots v_a$, there are s edges with the properties,*

$$\text{they belong to } T_1, \tag{2.4}$$

$$\text{they are single up to } v_a, \tag{2.5}$$

$$\text{they all have one end equal to } v_a, \tag{2.6}$$

then $s \le t + 1$, where t is the number of T_2 edges contained in $v_1 \cdots v_a$.

PROOF. Let the s edges be $v_{a_1}v_{a_1+1}$ and $v_{a_2}v_{a_2+1}, \dots, v_{a_s}v_{a_s+1}$. Suppose $v_{a_2} = \cdots = v_{a_s} = v_{a_1}$ or v_{a_1+1} and $a_1 < a_2 < \cdots < a_s < a$. By Lemma 2.1 in each of the chains $v_{a_2}v_{a_2+1} \cdots v_{a_3}, v_{a_3}v_{a_3+1} \cdots v_{a_4}, \dots, v_{a_s}v_{a_s+1} \cdots v_a$, there is a T_2 edge. But if e_b, e_c ($b < c$) are T_2 edges, e_b, e_c cannot coincide. Thus if t is the number of T_2 edges in $v_1 \cdots v_a$, $s - 1 \le t$ or $s \le t + 1$. □

A T_3 edge e_a is called regular, if the number of T_1 edges satisfying (2.4)–(2.6) in Lemma 2.2 is at least 2.

A chain $v_b v_{b+1} \cdots v_c$ will be called a ★-cycle (with head v_b) if

$$v_b v_{b+1} \in T_1 \text{ and single up to } v_c, \tag{2.7}$$

$$c \text{ is the smallest integer such that } c > b + 1 \text{ and } v_c = v_b, \tag{2.8}$$

$$\text{there is a } T_1 \text{ edge } e_a \text{ such that } a < b, \text{ single up to } v_c \text{ and } v_a = v_b \text{ or } v_{a+1} = v_b. \tag{2.9}$$

If in (2.9) we have $v_{a+1} = v_b$ and in the chain $v_{a+1} \cdots v_b$ there is no T_1 edge single up to v_b, then the ★-cycle $v_b \cdots v_c$ is called a ★-cycle of the first kind. Other ★-cycles are called ★-cycles of the second kind.

If C is a ★-cycle of the first kind, let $\Phi(C)$ be the last T_4 edge in C. If C is a ★-cycle of the second kind, let $\Phi(C)$ be the first T_4 edge in C.

LEMMA 2.3. *The number of regular T_3 edges is bounded by twice the number of T_4 edges.*

PROOF. It is evident that the number of regular T_3 edges is not larger than the number of ★-cycles. So, it is sufficient to prove that given any three ★-cycles, their Φ values cannot be identical.

Suppose they have different heads. Then one of the following two cases must occur:

CASE 1. At least two of them are of the first kind.

CASE 2. At least two of them are of the second kind.

Let the two ★-cycles in either cases be $C = v_b v_{b+1} \cdots v_c$ and $C' = v_{b'} v_{b'+1} \cdots v_{c'}$, $b < b'$ and $v_b \ne v_{b'}$. Now suppose we are with Case 1. C, C' are two ★-cycles of the first kind. By definition of Φ and the second part of Lemma

2.1, we know that $v_c = v_b \neq v_{b'} = v_{c'}$, so

$$\Phi(C) = v_{c-1}v_c \neq v_{c'-1}v_{c'} = \Phi(C').$$

Now suppose C, C' are both of the second kind. There are three possibilities to consider: (i) $b < c < b' < c'$; (ii) $b < b' < c' < c$; (iii) $b < b' < c < c'$. For (i), $\Phi(C) \neq \Phi(C')$ is obvious. For (ii), we will show that $\Phi(C)$ is in $v_b \cdots v_{b'}$. But $\Phi(C')$ is obviously in $v_{b'} \cdots v_{c'}$, so we will get $\Phi(C) \neq \Phi(C')$.

By definition, since $C' = v_{b'} \cdots v_{c'}$ is of the second kind we have a T_1 edge $v_a v_{a+1}$, single up to $v_{c'}$, $a < b'$, either $v_a = v_{b'}$ or $v_{a+1} = v_{b'}$ but in $v_{a+1}v_{a+2} \cdots v_{b'}$ there is a T_1 edge single up to $v_{b'}$. At first, suppose $v_a = v_{b'}$. If $a \leq b$, the chain $v_b \cdots v_{b'}$ contains a T_2 edge by Lemma 2.1. Thus $\Phi(C)$ is in $v_b \cdots v_{b'}$. If $b < a < b'$, since $v_a \cdots v_{b'}$ contains a T_2 edge by Lemma 2.1, $\Phi(C)$ is also in $v_b \cdots v_{b'}$. Now suppose $v_{a+1} = v_{b'}$ and in $v_{a+1}v_{a+2} \cdots v_{b'}$ there is a T_1 edge single up to $v_{b'}$. The chain from this T_1 edge to $v_{b'}$ satisfies the requirement for Lemma 2.1; it contains a T_2 edge, in this case, $\Phi(C)$ also belongs to $v_b \cdots v_{b'}$.

Now we consider case (iii). We need only to prove $\Phi(C)$ is in $v_b \cdots v_{b'}$. Let $v_a v_{a+1} \in T_1$ be single up to $v_{c'}$, $a < b'$, and either $v_a = v_{b'}$ or $v_{a+1} = v_{b'}$, but in $v_{a+1} \cdots v_{b'}$ there is a T_1 edge single up to $v_{b'}$. Just as in the proof for case (ii), $\Phi(C)$ is in $v_b \cdots v_{b'}$ in either situation.

Thus we get the conclusion that the number of ★-cycles is less than or equal to $2|T_4|$. On the other hand, the number of regular T_3 edges is less than or equal to the number of ★-cycles. The proof of Lemma 2.3 is finished. □

3. Two auxiliary theorems.

Theorem B. *If $EX_{11}^2 < \infty$, $EX_{12} = 0$, $EX_{12}^2 = \sigma^2$ and $EX_{12}^4 < \infty$, then the spectral distribution of the matrix $(1/\sqrt{n})W_n$ approaches the semicircle law (with density $(1/2\sigma^2\pi)\sqrt{4\sigma^2 - x^2}\,1_{(|x| \leq 2\sigma)}$) as $n \to +\infty$, a.s.*

This theorem is a consequence of Theorem 3.2.3 of Girko (1975), or cf. Arnold (1971).

Theorem C. *If $EX_{12} = 0$, $EX_{12}^2 = \sigma^2$, $EX_{12}^4 < \infty$ and $X_{ii} \equiv 0$, then $\lambda_{\max}((1/\sqrt{n})W_n) \to 2\sigma$ and $\lambda_{\min}((1/\sqrt{n})W_n) \to -2\sigma$ a.s.*

Lemma 3.1. (Truncation Lemma). *If $EX_{12}^4 < \infty$ and $X_{ii} = 0$, then there is a sequence $\delta_n \downarrow 0$ such that $P(W_n \neq \hat{W}_n \text{ i.o.}) = 0$ where $\hat{W}_n = (\hat{X}_{ij}^{(n)})$ and $\hat{X}_{ij}^{(n)} = X_{ij}1(|X_{ij}| \leq \sqrt{n}\,\delta_n)$. Evidently the speed of $\delta_n \downarrow 0$ can be made arbitrarily slow. Here $1(A)$ denotes the indicator function of A.*

Proof. Since $E|X_{12}|^4 < \infty$, we have for any $\varepsilon > 0$,

$$\sum_{m=1}^{\infty} 2^{2m} P\big(|X_{12}| \geq \varepsilon 2^{m/2}\big) < \infty.$$

We even have

$$\sum_{m=1}^{\infty} 2^{2m} P\big(|X_{12}| \ge \varepsilon_m 2^{m/2}\big) < \infty,$$

for some sequence $\varepsilon_m > 0$, converging to 0. Of course we can assume ε_m tends to 0 arbitrarily slowly.

Define $\delta = \delta_n = 2\varepsilon_m$ for $2^{m-1} \le n < 2^m$. Define $\tilde{W}_n = (\tilde{X}_{ij})_{1 \le i, j \le n}$, where $\tilde{X}_{ij} = X_{ij} 1(|X_{ij}| \le \sqrt{n}\,\delta)$. Thus for any $k \ge 1$,

$$\begin{aligned} P\big(W_n \ne \hat{W}_n \text{ i.o.}\big) &\le \sum_{m=k}^{\infty} P\Big(\bigcup_{2^{m-1} \le n < 2^m} \bigcup_{1 \le i < j \le n} \big(|X_{ij}| \ge \sqrt{n}\,\delta\big)\Big) \\ &\le \sum_{m=k}^{\infty} P\Big(\bigcup_{2^{m-1} \le n < 2^m} \bigcup_{1 \le i < j \le n} \big(|X_{ij}| \ge 2^{m/2}\varepsilon_m\big)\Big) \\ &\le \sum_{m=k}^{\infty} P\Big(\bigcup_{1 \le i < j \le 2^m} \big(|X_{ij}| \le 2^{m/2}\varepsilon_m\big)\Big) \\ &\le \sum_{m=k}^{\infty} 2^{2m} P\big(|X_{ij}| \ge 2^{m/2}\varepsilon_m\big) \to 0, \quad \text{as } k \to \infty. \qquad \square \end{aligned}$$

Now we prove Theorem C. Because of symmetry we prove only that $\lambda_{\max}((1/\sqrt{n})W_n) \to 2\sigma$, a.s. We can also assume that $\sigma = 1$. Because of Theorem B, $\liminf_{n\to\infty} \lambda_{\max}((1/\sqrt{n})W_n) \ge 2$, a.s. We need only to prove that $\limsup_{n\to\infty} \lambda_{\max}((1/\sqrt{n})W_n) \le 2$, a.s.

We can assume that the entries $X_{ij} = X_{ij}^{(n)}$ of W_n satisfy the conditions: $|X_{ij}| \le \sqrt{n}\,\delta$, where $\delta = \delta_n \downarrow 0$ very slowly; $EX_{ij} = 0$, for $i \ne j$; and $X_{ii} = 0$, $1 \le i \le j \le n$. In fact, we can replace the original entries of W_n by $X_{ij}1(|X_{ij}| \le \sqrt{n}\,\delta)$, denoted by $\tilde{X}_{ij}$, then define $\hat{X}_{ij} = \tilde{X}_{ij} - E\tilde{X}_{ij}$, for $i \ne j$, $1 \le i, j \le n$ and $\hat{X}_{ii} = 0$ for $1 \le i \le n$. By Lemma 3.1, $\tilde{W}_n = W_n$ for large n, a.s. For any vector x, $\|x\| = 1$, we have

$$x'\hat{W}_n x = x'\tilde{W}_n x - x'(J - I)xE\tilde{X}_{12} = x'\tilde{W}_n x - u.$$

Here J is $n \times n$ matrix with all entries be equal to 1. But

$$\begin{aligned} |u| &\le \Big(\Big(\sum x_i\Big)^2 + 1\Big)\big|EX_{12}1\big(|X_{12}| \le \sqrt{n}\,\delta\big)\big| \\ &\le (n+1)E|X_{12}|1\big(|X_{12}| > \sqrt{n}\,\delta\big), \end{aligned}$$

by $EX_{12} = 0$. Since $EX_{12}^4 < \infty$, $u \to 0$. So we can replace W_n by $\hat{W}_n$. Of course, now we have $|X_{ij}| \le 2\sqrt{n}\,\delta$, but we still write $|X_{ij}| \le \sqrt{n}\,\delta$. We can also assume $EX_{ij}^2 \le 1$ and tends to 1 and $EX_{ij}^4 \le d < \infty$ for some $d > 0$.

In order to show that $\limsup \lambda_{\max} \le 2$, a.s., it is sufficient to prove that for any constant $z > 2$ there is a sequence of positive integers $k = k(n)$ such that

$$\sum \frac{E\,\mathrm{tr}\big((1/\sqrt{n})W_n\big)^{2k}}{z^{2k}} < \infty.$$

Note that $[\lambda_{\max}((1/\sqrt{n})W_n)]^{2k} \le \mathrm{tr}[((1/\sqrt{n})W_n)^{2k}]$.

Now fix $z > 2$ arbitrarily. Let $k = k_n$ be any sequence of positive integers satisfying

$$k_n/\log n \to \infty, \qquad \delta_n^{1/3} k_n/\log n \to 0.$$

We are going to show that

$$\sum_n \frac{E\,\mathrm{tr}\big((1/\sqrt{n})W_n\big)^{2k_n}}{z^{2k_n}} < \infty.$$

We first estimate

$$E\,\mathrm{tr}\left(\frac{1}{\sqrt{n}}W_n\right)^{2k} = n^{-k}\sum{}^{*} E X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_{2k} i_1}.$$

Here Σ^* means the summation is taken for $i_1, \dots, i_{2k}$ running from 1 to n and subject to the condition that $i_1 \neq i_2$, $i_2 \neq i_3, \dots, i_{2k} \neq i_1$ and $X_{ij} = X_{ij}^{(n)}$ depends on n, but for simplicity we do not write the n explicitly.

We can regard $i_1, i_2, \dots, i_{2k}$ as vertices for the graph described in Section 2. The $2k$ edges are automatically determined so that if e_a is the ath edge, then the initial and the terminal of e_a are i_a and i_{a+1}, respectively. Thus

$$E\,\mathrm{tr}\left(\frac{1}{\sqrt{n}}W_n\right)^{2k} = n^{-k}\sum{}'\sum{}''\sum{}''' E X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_{2k} i_1}.$$

Here Σ' is the summation for different arrangements of three different types (T_1, T_3, T_4) of edges at $2k$ consecutive positions; Σ'' is the summation for all canonical graph Γ with given arrangement of the three types of edges (a graph is called canonical if $i_r \in T_1$ implies $i_r = \max(i_1, \dots, i_{r-1}) + 1$, but $i_1 = 1$); Σ''' is (for a given canonical graph) the summation over all graphs isomorphic to it [two graphs $(v_1, \dots, v_{2k})$ and $(v_1', \dots, v_{2k}')$ are called isomorphic if $v_i = v_j \Leftrightarrow v_i' = v_j'$].

Let l denote the number of edges of class T_1, so l is also the number of edges in T_3. So, Σ' can be replaced by $\sum_{l=1}^{k}\binom{2k}{l}\binom{2k-l}{l}$. If there is an edge single throughout the graph, the corresponding mean is 0. We do not consider such graphs. Σ''' is evidently bounded by n^{l+1}.

Suppose we know the $2k$ edges which belong to T_1, which belong to T_3 or T_4. Now we estimate the bound of the number of canonical graphs with the given arrangement of three types of edges. If an edge $i_r i_{r+1}$ should be in T_3 and it is not regular, it can only coincide with the only T_1 edge which has a common end with it. If it is a regular edge, then it has at most $t + 1$ possible choices by Lemma 2.2, where t is the maximal number of noncoincident T_4 edges. But the maximal number of regular T_3 edges is bounded by $2(2k - 2l)$ by Lemma 2.3. If the maximal number of noncoincident T_4 edges is t, the $2(k - l)T_4$ edges have at most $\binom{k^2}{t}t^{2(k-l)}$ choices. Thus Σ''' can be replaced by $\sum_{t=0}^{2k-2l}(t + 1)^{2(2k-2l)}\binom{k^2}{t}t^{2(k-l)}$.

Finally, we bound the mean $EX_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_{2k} i_1}$. If $t = 0$, it is $[E(X_{12}^2)]^k \le 1$. For $t \ge 1$, let μ be the number of T_1 edges which coincide with T_4 edges. Let n_i

denote the number of T_4 edges which coincide with the ith such T_1 edge, $i = 1, \dots, \mu$.

Let m_j be the number of T_4 edges which are coincident of each other but not with any T_1 edge, $j = 1, \dots, t - \mu$. Then we have

$$EX_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_{2k} i_1} = \left[EX_{12}^2\right]^{l-\mu} \prod_{i=1}^{\mu} E\left(X_{12}^{n_i+2}\right) \prod_{j=1}^{t-\mu} EX_{12}^{m_j},$$

where $2(l - \mu) + \sum_{i=1}^{\mu}(n_i + 2) + \sum_{i=1}^{t-\mu} m_i = 2k$ and $\mu \le t$.

It is easy to see that $E|X_{12}|^l \le (\sqrt{n}\,\delta)^{l-2}$ for $l \ge 2$ and $E|X_{12}|^l \le C(\sqrt{n}\,\delta)^{l-3}$ for $l \ge 3$, where $C > 0$ ($C = E|X_{12}|^3$) is a constant, so we have

$$|EX_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_{2k} i_1}| \le C^{\mu}\left(\sqrt{n}\,\delta\right)^{2k-2l-t} \le k^t\left(\sqrt{n}\,\delta\right)^{2k-2l-t}.$$

Combining the preceding considerations, we have

$$\begin{aligned} E\,\mathrm{tr}\left(\frac{1}{\sqrt{n}} W_n\right)^{2k} &\le n^{-k} \sum_{l=1}^{k} \binom{2k}{l}\binom{2k-l}{l} n^l \\ &\quad \times \sum_{t=0}^{2(k-l)} \binom{k^2}{t} t^{2(k-l)} (t+1)^{2(2k-2l)} k^t \left(\sqrt{n}\,\delta\right)^{2k-2l-t} \\ &\le \sum_{l=1}^{k} \frac{(2k)!}{l!\,l!\,(2k-2l)!} \sum_{t=0}^{2(k-l)} k^{2t}(t+1)^{6(k-l)} k^t \delta^{2k-2l} \left(\sqrt{n}\,\delta\right)^{-t} \\ &= \sum_{l=1}^{k} \frac{(2k)!}{l!\,l!\,(2k-2l)!} \sum_{t=0}^{2(k-l)} \left(\frac{k^3}{\sqrt{n}\,\delta}\right)^t (t+1)^{6(k-l)} \delta^{2(k-l)}. \end{aligned}$$

Consider the function $g(t) = a^t(t+1)^b$, $t \in [-1, \infty)$, where $0 < a < 1$ and $b > 0$ are constants. $g'(t) = a^t(t+1)^{b-1}(b + (t+1)\log a) \le 0$ when $t \ge b/|\log a| - 1$ and $g'(t) \ge 0$ for $-1 \le t \le b/|\log a| - 1$. Thus, in the interval $0 \le t \le b$,

$$g(t) \le a^{b/|\log a| - 1}\left(\frac{b}{|\log a|}\right)^b \le \left(\frac{b}{|\log a|}\right)^b, \quad \text{if } \frac{b}{|\log a|} - 1 \ge 0,$$

$$g(t) \le g(0) = 1, \quad \text{if } \frac{b}{|\log a|} - 1 < 0.$$

In any case we have, when $0 \le t \le b$,

$$g(t) \le \left(\frac{b}{|\log a|}\right)^b \vee 1 = \left(\frac{b}{|\log a|} \vee 1\right)^b.$$

Thus if n is large,

$$E\,\mathrm{tr}\left(\frac{1}{\sqrt{n}}W_n\right)^{2k} \le \sum_{l=1}^{k}\frac{(2k)!}{l!l!(2k-2l)!}\sum_{t=0}^{2(k-l)}\left(\frac{6(k-l)}{\log(\sqrt{n}\,\delta/k^3)}\delta^{1/3}\vee\delta^{1/3}\right)^{6(k-l)}$$

$$\le 2k\sum_{l=1}^{k}\frac{(2k)!}{l!l!(2k-2l)!}\left(\frac{18k\delta^{1/3}}{\log n}\vee\delta^{1/3}\right)^{6(k-l)}$$

$$\le 2k\left(1+1+\left(\frac{18k\delta^{1/3}}{\log n}\vee\delta^{1/3}\right)^{3}\right)^{2k} = 2k(2+\Delta_n)^{2k}.$$

We choose δ such that $\Delta_n = ((18k\delta^{1/3}/\log n)\vee\delta^{1/3})^3 \to 0$ as $n\to\infty$. Let $L_n = k/(\log n)$. Then $L_n \to \infty$ and

$$\sum_n \frac{E\left(\lambda_{\max}\left((1/\sqrt{n})W_n\right)\right)^{2k}}{z^{2k}} \le \sum_n 2k\frac{(2+\Delta_n)^{2k}}{z^{2k}} = \sum_n 2k e^{2k\log((2+\Delta_n)/z)}$$

$$= \sum_n 2kn^{2L_n\log((2+\Delta_n)/z)}.$$

The last series is evidently convergent, since $\Delta_n\to 0$, $z>2$, $L_n\to\infty$, and $k\le \log^2 n$, if we choose $\delta_n\to 0$ slower than $1/(\log n)$. □

4. The proof of the necessity part of Theorem A. We suppose $\lambda_{\max}((1/\sqrt{n})W_n)\to a$ as $n\to\infty$ a.s. Here a is a finite number. We are going to prove conditions (1.5)–(1.8) hold.

For the proof of (1.5), we suppose $E(X_{11}^+)^2 = +\infty$. We are going to show that $\limsup \lambda_{\max}((1/\sqrt{n})W_n) = +\infty$ a.s.

Since $E(X_{11}^+)^2 = +\infty$, $P(X_{11}>0)>0$. Thus $\sum P(\max_{1\le i\le n}X_{ii}\le 0)<\infty$ and

$$\max_{1\le i\le n} X_{ii} > 0 \quad \text{for large } n \quad \text{a.s.}$$

Hence a.s.,

$$\lambda_{\max}\left(\frac{1}{\sqrt{n}}W_n\right) \ge \frac{1}{\sqrt{n}}\max_{1\le i\le n}X_{ii} = \frac{1}{\sqrt{n}}\max_{1\le i\le n}X_{ii}^+ \quad \text{for large } n.$$

But for any $K>0$,

$$\sum_i P\left((X_{ii}^+)^2 \ge iK\right) = +\infty.$$

So, $\limsup_{n\to\infty} X_{nn}^+/\sqrt{n} = +\infty$ a.s. and $\limsup_{n\to\infty}\lambda_{\max}((1/\sqrt{n})W_n) = +\infty$ a.s. follows.

For the proof of (1.6) we suppose $E(X_{11}^+)^2<\infty$ and $EX_{12}^4 = +\infty$. Let $M>0$ be such that $p = P(X_{11}\le -M) < \frac{1}{2}$. Define

$$\tilde{X}_{ii} = X_{ii}1_{(X_{ii}\le -M)}.$$

Then

$$X_{ii} - \tilde{X}_{ii} = X_{ii} 1_{(X_{ii} > -M)} = X_{ii}^{+} + X_{ii} 1_{(0 > X_{ii} > -M)}$$

and then

$$E\left(X_{ii} - \tilde{X}_{ii}\right)^2 < \infty, \qquad \sum_i P\left(\left(X_{ii} - \tilde{X}_{ii}\right)^2 \geq i\varepsilon\right) < \infty \qquad (\forall\, \varepsilon > 0),$$

which implies

$$\lim \frac{1}{\sqrt{n}} \max_{1 \leq i \leq n} |X_{ii} - \tilde{X}_{ii}| = 0 \quad \text{a.s.}$$

Now, let $\tilde{W}_n = W_n + \operatorname{diag}\{\tilde{X}_{11} - X_{11}, \dots, \tilde{X}_{nn} - X_{nn}\}$. Since

$$\begin{aligned} \left| x' \frac{1}{\sqrt{n}} \left(W_n - \tilde{W}_n\right) x \right| &= \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(X_{ii} - \tilde{X}_{ii}\right) x_i^2 \right| \\ &\leq \frac{1}{\sqrt{n}} \max_{1 \leq i \leq n} |X_{ii} - \tilde{X}_{ii}| \to 0 \quad \text{a.s.,} \end{aligned}$$

we have

$$\left| \lambda_{\max}\left(\frac{1}{\sqrt{n}} W_n\right) - \lambda_{\max}\left(\frac{1}{\sqrt{n}} \tilde{W}_n\right) \right| \leq \sup_{\|x\|=1} \left| x' \frac{1}{\sqrt{n}} \left(W_n - \tilde{W}_n\right) x \right| \to 0 \quad \text{a.s.}$$

So, we need only to show that $\limsup \lambda_{\max}((1/\sqrt{n})\tilde{W}_n) = +\infty$ a.s. or, in other words, we will assume $X_{ii} = X_{ii} 1_{(X_{ii} \leq -M)}$.

It is easy to see that

$$\lambda_{\max}\left(\frac{1}{\sqrt{n}} W_n\right) \geq \frac{1}{2\sqrt{n}} \max_{1 \leq i < j \leq n} \left(X_{ii} + X_{jj} + 2|X_{ij}|\right).$$

Thus, for $B > 0$,

$$\tag{4.1} \begin{aligned} & P\left(\lambda_{\max}\left(\frac{1}{\sqrt{n}} W_n\right) \geq B \text{ i.o.}\right) = 1 \\ & \Leftarrow P\left(\frac{1}{2} \frac{1}{\sqrt{n}} \max_{1 \leq i < j \leq n} \left(X_{ii} + X_{jj} + 2|X_{ij}|\right) \geq B \text{ i.o.}\right) = 1 \\ & \Leftarrow P\left(\frac{1}{2\sqrt{2^{k+1}}} \max_{2^k < i < j \leq 2^{k+1}} \left(X_{ii} + X_{jj} + 2|X_{ij}|\right) \geq B \text{ i.o.}\right) = 1 \\ & \Leftarrow \sum_k P\left(\frac{1}{2\sqrt{2^{k+1}}} \max_{1 \leq i < j \leq 2^k} \left(X_{ii} + X_{ij} + 2|X_{ij}|\right) \geq B\right) = +\infty. \end{aligned}$$

But

$$P\left(\frac{1}{2\sqrt{2^{k+1}}}\max_{1\le i<j\le 2^k}\left(X_{ii}+X_{jj}+2|X_{ij}|\right)\ge B\right)$$

$$\ge \sum_{l=2^{k-1}}^{2^k}\sum_{1\le i_1<\cdots<i_l\le 2^k} P\Big(\text{just } X_{i_1i_1},\dots,X_{i_li_l} \text{ are the vanishing diagonal entries and } \max_{1\le u<v\le l}|X_{i_ui_v}|\ge B2^{(k+1)/2)}\Big)$$

(4.2)

$$= \sum_{l=2^{k-1}}^{2^k}\binom{2^k}{l}\left[P(X_{11}=0)\right]^l\left[P(X_{11}\ne 0)\right]^{2^k-l} \times P\Big(\max_{1\le i<j\le l}|X_{ij}|\ge B2^{(k+1)/2}\Big)$$

$$\ge \frac{1}{2}P\Big(\max_{1\le i<j\le 2^{k-1}}|X_{ij}|\ge B2^{(k+1)/2}\Big),$$

since $P(X_{11}=0)>\frac{1}{2}$. Evidently,

$$\sum_k P\Big(\max_{1\le i<j\le 2^{k-1}}|X_{ij}|\ge B2^{(k+1)/2}\Big)=+\infty$$

(4.3)

$$\Leftrightarrow \prod_k\left(P\big(|X_{12}|<B2^{(k+1)/2}\big)\right)^{(2^{k-1}(2^{k-1}-1))/2}=0$$

$$\Leftrightarrow \sum \frac{2^{k-1}(2^{k-1}-1)}{2}P\big(|X_{12}|\ge B2^{(k+1)/2}\big)=+\infty$$

$$\Leftarrow EX_{12}^4=+\infty.$$

Combining (4.1)–(4.3) we get

$$EX_{12}^4=+\infty \Rightarrow P\left(\limsup_{n\to\infty}\lambda_{\max}\left(\frac{1}{\sqrt{n}}W_n\right)=+\infty\right)=1.$$

Now suppose $E(X_{11}^+)^2<\infty$ and $EX_{12}^4<\infty$, but $EX_{12}=\alpha>0$. We are going to show that $P\big(\limsup_{n\to\infty}\lambda_{\max}\big((1/\sqrt{n})W_n\big)=+\infty\big)=1$. As before we may assume $X_{ii}=X_{ii}1_{(X_{ii}<-M)}$ and $P(X_{11}<-M)<\frac{1}{2}$.

Because $\lambda_{\max}((1/\sqrt{n})W_n)\ge x'(1/\sqrt{n})W_nx$ for any unit vector x, we can see that for any integers j, $i_1,\dots,i_j$ with $1\le j\le n$, $1\le i_1<\cdots<i_j\le n$, by appropriately choosing x, we have

$$\lambda_{\max}\left(\frac{1}{\sqrt{n}}W_n\right)\ge \frac{1}{j}\frac{1}{\sqrt{n}}\sum_{r=1}^{j}\sum_{s=1}^{j}X_{i_ri_s}.$$

Thus, for any $B > 0$,

$$P\left(\lambda_{\max}\left(\frac{1}{\sqrt{n}} W_n\right) \geq B \text{ i.o.}\right) = 1$$

$$\Leftarrow P\left(\bigcup_{1 \leq j \leq 2^k} \bigcup_{2^k < i_1 < \cdots < i_j \leq 2^{k+1}} \left(\frac{1}{j} 2^{-(k+1)/2} \sum_{r=1}^{j} \sum_{s=1}^{j} X_{i_r i_s} \geq B\right) \text{ i.o.}\right) = 1 \tag{4.4}$$

$$\Leftrightarrow \sum_k P\left(\bigcup_{1 \leq j \leq 2^k} \bigcup_{2^k < i_1 < \cdots < i_j \leq 2^{k+1}} \left(\frac{1}{j} 2^{-(k+1)/2} \sum_{r=1}^{j} \sum_{s=1}^{j} X_{i_r i_s} \geq B\right)\right) = \infty.$$

But the general term of the last series is greater than or equal to

$$P\left(\bigcup_{1 \leq j \leq 2^k} \bigcup_{2^k < i_1 < \cdots < i_j \leq 2^{k+1}} \left(X_{i_1 i_1} = 0, \dots, X_{i_j i_j} = 0 \text{ only}\right.\right.$$

$$\left.\left. \text{and } \frac{1}{j} 2^{-(k+2)/2} 2 \sum_{1 \leq r < s \leq j} X_{i_r i_s} \geq B\right)\right)$$

$$= \sum_{j=1}^{2^k} \binom{2^k}{j} [P(X_{11} = 0)]^j [P(X_{11} \neq 0)]^{2^k - j}$$

$$\times P\left(\frac{1}{j} 2^{-(k+1)/2} 2 \sum_{1 \leq r < s \leq j} X_{rs} \geq B\right)$$

$$\geq \sum_{j=2^{k-1}}^{2^k} \binom{2^k}{j} [P(X_{11} = 0)]^j [P(X_{11} \neq 0)]^{2^k - j}$$

$$\times P\left(\frac{\sum_{1 \leq r < s \leq j} X_{rs}}{j(j-1)/2} \geq \frac{(Bj2^{(k+1)/2})/2}{j(j-1)/2}\right).$$

But

$$\frac{(Bj2^{(k+1)/2})/2}{j(j-1)/2} \leq \frac{B2^{(k+1)/2}}{2^{k-1} - 1} \to 0 \quad \text{as } k \to \infty,$$

$$\frac{\sum_{1 \leq r < s \leq j} X_{rs}}{j(j-1)/2} \to EX_{12} = \alpha > 0 \quad \text{as } j \geq 2^{k-1} \to \infty \text{ a.s.}$$

So as k is large,

$$P\left(\sum_{1 \leq r < s \leq j} X_{rs} / \frac{j(j-1)}{2} \geq \frac{(Bj2^{(k+1)/2})/2}{j(j-1)/2}\right) > \frac{1}{2}$$

and since $P(X_{11}=0) > \frac{1}{2}$,

$$\sum_{j=2^{k-1}}^{2^k} \binom{2^k}{j} [P(X_{11}=0)]^j [P(X_{11}\neq 0)]^{2^k-j} \geq \tfrac{1}{2},$$

therefore, the series in (4.4) is divergent and $\limsup \lambda_{\max}((1/\sqrt{n})W_n) = +\infty$ a.s.

5. The sufficiency of the conditions. Suppose $E(X_{11}^+)^2 < \infty$, $EX_{12}^4 < +\infty$ and $EX_{12} \leq 0$. By $E(X_{11}^+)^2 < \infty$, we can assume that $X_{ii} = X_{ii}1(X_{ii} \leq -M)$ for some $M > 0$ with $P(X_{11} \leq -M) < \frac{1}{2}$.

Let $\tilde{X}_{ij} = X_{ij} - EX_{ij}$ for $i \neq j$, $\tilde{X}_{ii} \equiv 0$. Then if $x'x = 1$

$$\begin{aligned}
x'\frac{1}{\sqrt{n}}W_n x &= x'\frac{1}{\sqrt{n}}\tilde{W}_n x + x'\frac{1}{\sqrt{n}}\begin{pmatrix} X_{11} & EX_{12} & \cdots & EX_{12} \\ EX_{12} & X_{22} & \cdots & EX_{12} \\ \cdots & & & \\ EX_{12} & EX_{12} & \cdots & X_{nn}\end{pmatrix} x \\
&= x'\frac{1}{\sqrt{n}}\tilde{W}_n x + 2\frac{1}{\sqrt{n}}EX_{12}\sum_{1\leq i<j\leq n} x_i x_j + \frac{1}{\sqrt{n}}\sum X_{ii}x_i^2 \\
&\leq x'\frac{1}{\sqrt{n}}\tilde{W}_n x + \frac{1}{\sqrt{n}}EX_{12}\Big(\sum x_i\Big)^2 + \frac{1}{\sqrt{n}}|EX_{12}| \\
&\leq x'\frac{1}{\sqrt{n}}\tilde{W}_n x + \frac{|EX_{12}|}{\sqrt{n}},
\end{aligned}$$

since $EX_{12} \leq 0$. Therefore, $\lambda_{\max}((1/\sqrt{n})W_n) \leq \lambda_{\max}((1/\sqrt{n})\tilde{W}_n) + |EX_{12}|/\sqrt{n}$.

By Theorem C, we get $\limsup_{n\to\infty} \lambda_{\max}((1/\sqrt{n})W_n) \leq 2$ a.s.

Now let k be the number of diagonal entries of W_n which are not 0. k is a random variable with binomial distribution $B(n, p)$, $p = P(X_{11} \neq 0) < \frac{1}{2}$.

Let V be the set of all $n \times 1$ vectors $x = (x_1, \dots, x_n)'$ with the properties: $\sum_1^n x_i = 0$ and $X_{ii} \neq 0 \Rightarrow x_i = 0$. Then V is a vector space of dimension $n - k - 1$. Then we have

$$\begin{aligned}
\lambda_{\max}\left(\frac{1}{\sqrt{n}}W_n\right) &\geq \sup_{\substack{x\in V\\ \|x\|=1}} x'\left(\frac{1}{\sqrt{n}}W_n\right)x \\
&= \sup_{\substack{x\in V\\ \|x\|=1}} \left(x'\left(\frac{1}{\sqrt{n}}\tilde{W}_n\right)x + \frac{1}{\sqrt{n}}\sum_{i=1}^n (X_{ii} - EX_{12})x_i^2\right) \\
&\geq \sup_{\substack{x\in V\\ \|x\|=1}} x'\left(\frac{1}{\sqrt{n}}\tilde{W}_n\right)x \geq \lambda_{k+2}\left(\frac{1}{\sqrt{n}}\tilde{W}_n\right).
\end{aligned}$$

By a lemma in Bai and Yin (1988),

$$P\left(\frac{k}{n} - p \geq p\right) \leq e^{-(n/4)p} \quad \text{for all } n.$$

Thus

$$P(k \geq 2np \text{ i.o.}) = 0.$$

Therefore,

$$\lambda_{\max}\left(\frac{1}{\sqrt{n}} W_n\right) \geq \lambda_{[2pn]+2}\left(\frac{1}{\sqrt{n}} \tilde{W}_n\right) \quad \text{a.s.}$$

for large n. But as $n \to \infty$,

$$\lambda_{[2pn]+2}\left(\frac{1}{\sqrt{n}} \tilde{W}_n\right) \to F^{-1}(1-2p) = \nu_{2p} \quad \text{a.s.},$$

where $F(x) = (1/2\pi)\int_{-2}^{x}\sqrt{4-u^2}\,du$. Thus

$$\liminf_{n\to\infty} \lambda_{\max}\left(\frac{1}{\sqrt{n}} W_n\right) \geq \nu_{2p}$$

a.s. for any $p < \frac{1}{2}$. Letting $p \to 0$, $\nu_{2p} \to 2$.

Therefore, $\lim_{n\to\infty} \lambda_{\max}((1/\sqrt{n})W_n) = 2$ a.s.

Acknowledgments. The authors are grateful to Ms. Sarah Oordt and Ms. Susan Rollins for their patience and excellence in typing this paper.

REFERENCES

ARNOLD, O. L. (1971). On Wigner's semicircle law of eigenvalues of random matrices. *Z. Wahrsch. verw. Gebiete* **19** 191–198.

BAI, Z. D. and YIN, Y. Q. (1988). Convergence to the semicircle law. *Ann. Probab.* **16** 863–875.

FÜREDI, Z. and KOMLÓS, J. (1981). The eigenvalues of random symmetric matrices. *Combinatorica* **1** 233–241.

GIRKO, V. L. (1975). *Random Matrices*. Kiev. (In Russian).

JUHÁSZ, F. (1981). On the spectrum of a random graph. *Algebraic Methods in Graph Theory. Colloq. Math. Soc. János Bolyai* **25** 313–316. North-Holland, Amsterdam.

WIGNER, E. P. (1958). On the distribution of the roots of certain symmetric matrices. *Ann. of Math.* **67** 325–327.

CENTER FOR MULTIVARIATE ANALYSIS
UNIVERSITY OF PITTSBURGH
PITTSBURGH, PENNSYLVANIA 15260

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF LOWELL
LOWELL, MASSACHUSETTS 01854