Elect. Comm. in Probab. **9** (2004), 82–91

ELECTRONIC COMMUNICATIONS in PROBABILITY

POISSON STATISTICS FOR THE LARGEST EIGENVALUES OF WIGNER RANDOM MATRICES WITH HEAVY TAILS

ALEXANDER SOSHNIKOV[1]

Department of Mathematics, University of California at Davis, One Shields Ave, Davis, CA 95616, USA

email: soshniko@math.ucdavis.edu

Submitted 20 May 2004, *accepted in final form* 30 June 2004

AMS 2000 Subject classification: 15A52, 60B99, 60E07

Keywords: random matrices, largest eigenvalues, Poisson statistics

Abstract

We study large Wigner random matrices in the case when the marginal distributions of matrix entries have heavy tails. We prove that the largest eigenvalues of such matrices have Poisson statistics.

1 Introduction and Formulation of Results.

The main goal of this paper is to study the largest eigenvalues of Wigner real symmetric and Hermitian random matrices in the case when the matrix entries have heavy tails of distribution. We remind that a real symmetric Wigner random matrix is defined as a square symmetric $n \times n$ matrix with i.i.d. entries up from the diagonal

$$A = (a_{jk}), \quad a_{jk} = a_{kj}, \ 1 \leq j \leq k \leq n, \quad \{a_{jk}\}_{j \leq k} - \text{ i.i.d. } \textit{real} \text{ random variables.} \qquad (1)$$

A Hermitian Wigner random matrix is defined in a similar way, namely as a square $n \times n$ Hermitian matrix with i.i.d. entries up from the diagonal

$$A = (a_{jk}), \quad a_{jk} = \overline{a_{kj}}, \ 1 \leq j \leq k \leq n, \quad \{a_{jk}\}_{j \leq k} - \text{ i.i.d. complex random variables} \qquad (2)$$

(it is assumed quite often that the distribution of the diagonal matrix entries is different from the distribution of the off-diagonal entries). While our main results (with obvious modifications) are valid for the Hermitian random matrices with i.i.d. entries as well, we will restrict ourselves for the most of the paper to the real symmetric case.

The ensembles (1), (2) were introduced by E.Wigner in the fifties ([33], [34]), [35]) who considered the case of centered identically distributed (up from the diagonal) matrix entries with

[1]RESEARCH WAS SUPPORTED IN PART BY THE SLOAN RESEARCH FELLOWSHIP AND THE NSF GRANT DMS-0103948

the tail of the distribution decaying sufficiently fast so that all moments exist. Wigner proved that under the above conditions the mathematical expectation of the empirical distribution function of $n^{-1/2}A$ converges as $n \to \infty$ to the semicircle law, i.e.

$$\frac{1}{n}E\left(\#(\lambda_i \le x, \quad 1 \le i \le n)\right) \to F(x) = \int_{-\infty}^{x} f(t)dt, \tag{3}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $n^{-1/2}A$, the density of the semicircle law is given by $f(t) = \frac{1}{\pi\sigma^2}\sqrt{2\sigma^2 - x^2}$, for $t \in [-\sqrt{2}\sigma, \sqrt{2}\sigma]$ and σ^2 is the second moment of matrix entries. Wigner's result was subsequently strengthened by many mathematicians (see e.g. [17], [1], [24], [8]). In particular, Pastur ([24]) and Girko ([10]) proved that if $A^{(n)} = (a_{ij}^{(n)}), \ 1 \le i,j \le n,$ is an $n \times n$ random symmetric matrix with independent (not necessarily identicaly distributed) centered entries with the same second moment σ^2 , then the necessary and sufficient for the convergence of the empirical distribution function of the eigenvalues of $n^{-1/2}A^{(n)}$ to the semicircle law has the Lindeberg-Feller form: $\frac{1}{n^2}\sum_{1\le i\le j\le n}\int_{|x|>\tau\sqrt{n}} x^2 dF_{ij}(x) \to 0,$ where $F_{ij}(x)$ is the distribution function of $a_{ij}^{(n)}$.
The simplest case of (1) from the analytical viewpoint is when the entries $a_{ij}, \ 1 \le i \le j \le n$ are independent Gaussian $N(0, 1+\delta_{ij})$ random variables. Then the ensemble is called Gaussian Orthogonal Ensemble (GOE). We refer the reader to [21], chapter 6, for the discussion of the GOE ensemble. In a similar fashion the ensemble of Hermitian random matrices with independent and identically distributed (up from the diagonal) matrix entries is called Gaussian Unitary Ensemble (GUE, see [21], chapter 5) and the ensemble of Hermitian self-dual matrices with quaternion entries is called Gaussian Symplectic Ensemble (GSE, see [21], chapter 7). In the Gaussian ensembles one can study in detail the local statistical properties of the eigenvalues both in the bulk and at the edge of the spectrum. In the seminal papers ([31], [32]) Tracy and Widom proved that after the proper rescaling the distribution of the largest eigenvalues in the GOE, GUE and GSE ensembles converges to what is now called Tracy-Widom distribution. In particular they proved that for the largest eigenvalue of GOE

$$\lim_{n\to\infty} \Pr\left(\lambda_{max} \le 2\sqrt{n} + \frac{s}{n^{1/6}}\right) \to F_1 = \exp\left(-\frac{1}{2}\int_s^{+\infty} q(x) + (x-s)q^2(x)dx\right), \tag{4}$$

and for the largest eigenvalue of the GUE ensemble

$$\lim_{n\to\infty} \Pr\left(\lambda_{max} \le 2\sqrt{n} + \frac{s}{n^{1/6}}\right) \to F_2 = \exp\left(-\int_s^{+\infty} (x-s)q^2(x)dx\right), \tag{5}$$

where $q(x)$ is the solution of the Painléve II differential equation $q''(x) = xq(x) + 2q^3(x)$ determined by the asymptotics $q(x) \sim Ai(x)$ at $x = +\infty$. It was also established (see [31], [32], [9]) that after the rescaling at the edge of the spectrumas the k-point correlation functions have a limit. In the GUE case the limiting k-point correlation functions have a determinantal form

$$\rho_k(x_1, \ldots, x_k) = \det\left(K(x_i, x_j)\right)_{1\le i,j\le k}, \tag{6}$$

where

$$K(x,y) = K_{Airy}(x,y) = \frac{Ai(x)Ai'(y) - Ai'(x)Ai(y)}{x-y} \tag{7}$$

is a so-called Airy kernel. In the GOE case the limiting k-point correlation functions have a pfaffian form and can be written as

$$\rho_k(x_1, \ldots, x_k) = \left(\det\left(K(x_i, x_j)\right)_{1\le i,j\le k}\right)^{1/2}, \tag{8}$$

where $K(x,y)$ is a 2×2 matrix kernel such that

$$
\begin{aligned}
K_{11}(x,y) &= K_{22}(y,x) = K_{Airy}(x,y) + \frac{1}{2}Ai(x)\int_{-\infty}^{y} Ai(t)dt, && (9)\\
K_{12}(x,y) &= -\frac{1}{2}Ai(x)Ai(y) - \frac{\partial}{\partial y}K_{Airy}(x,y), && (10)\\
K_{21}(x,y) &= \int_0^{+\infty}\left(\int_{x+u}^{+\infty} Ai(v)dv\right) Ai(x+u)du - \epsilon(x-y) + \frac{1}{2}\int_y^x Ai(u)du\\
&+ \frac{1}{2}\int_x^{+\infty} Ai(u)du \int_{-\infty}^{y} Ai(v)dv, && (11)
\end{aligned}
$$

where $\epsilon(z) = \frac{1}{2}sign(z)$. Similar formulas hold for the GSE.
Soshnikov ([28], see also [29]) proved that the Tracy-Widom law for the largest eigenvalues is universal provided the laws of the distribution of matrix entries are symmetric, all moments exist and do not grow too fast. To the best of our knowledge there are no results proving universality in the bulk of the spectrum for real symmetric Wigner matrices. In the Hermitian case Johansson ([13]) proved the universality in the bulk under the condition that matrix entries have a Gaussian component.
In this paper we will study the spectral properties of Wigner real symmetric and Hermitian random matrices in the case when the matrix entries have heavy tails of distribution. In other words, in addition to the assumption that a_{jk} are i.i.d. real (complex) random variables up from the diagonal ($1 \le j \le k \le n$) we also assume that the probability distribution function satisfies

$$G(x) = \Pr(|a_{jk}| > x) = \frac{h(x)}{x^{\alpha}}, \tag{12}$$

where $0 < \alpha < 2$ and $h(x)$ is a slowly varying function at infinity in a sense of Karamata ([14], [26]), in other words $h(x)$ is a positive function for all $x > 0$, such that $\lim_{x\to\infty}\frac{h(tx)}{h(x)} = 1$ for all $t > 0$. The condition (12) means that the distribution of $|a_{ij}|$ belongs to the domain of the attraction of a stable distribution with the exponent α (see e.g. [11], Theorem 2.6.1). The case $h(x) = const(1 + o(1))$ in (12) was considered on a physical level of rigor by Cizeau and Bouchaud in [4], who called such set of matrices "Lévy matrices" (see also [2] and [12] for some physical results on the unitary invariant Lévy ensembles). They argued that the typical eigenvalues of A should be of order of $n^{\frac{1}{\alpha}}$. Indeed, such normalization makes the euclidean norm of a typical row of A to be (with high probability) of the order of constant. Cizeau and Bouchaud suggested a formula for the limiting distribution of the (normalized by $n^{\frac{1}{\alpha}}$) eigenvalues. The formula for the limiting spectral distribution has a more complicated density then in the finite variance case. Namely, it was proposed that the density should be given by

$$f(x) = L_{\alpha/2}^{C(x),\beta(x)}(x), \tag{13}$$

where $L_{\alpha}^{C,\beta}$ is a density of a centered Lévy stable distribution defined through its Fourier transform $\hat{L}(k)$:

$$L_{\alpha}^{C,\beta} = \frac{1}{2\pi}\int dk\hat{L}(k)e^{ikx}, \tag{14}$$

$$\ln \hat{L}(k) = -C|k|^{\alpha}\big(1 + i\beta sgn(k)\tan(\pi\alpha/2)\big), \tag{15}$$

and $C(x),\ \beta(x)$ satisfy a system of integral equations

$$C(x) = \int_{-\infty}^{+\infty} |y|^{\frac{\alpha}{2}-2} L_{\alpha/2}^{C(y),\beta(y)}\big(x - \frac{1}{y}\big)dy, \tag{16}$$

$$\beta(x) = \int_{x}^{+\infty} L_{\alpha/2}^{C(y),\beta(y)}\big(x - \frac{1}{y}\big)dy. \tag{17}$$

Note that the density $f(x)$ in (13) is not a Lévy density itself, since $C(x),\ \beta(x)$ are functions of x. It was also argued in [4] that the density $f(x)$ should decay like $\frac{1}{x^{1+\alpha}}$ at infinity, which suggests that the largest eigenvalues of A (in the case (12), $h(x) = const$) should be of order $n^{\frac{2}{\alpha}}$, and not $n^{\frac{1}{\alpha}}$ as the typical eigenvalues.
Recently, Soshnikov and Fyodorov ([30]) used the method of determinants to study the largest eigenvalues of a sample covariance matrix A^tA in the case when the entries of a rectangular $m \times n$ matrix A are i.i.d. Cauchy random variables. The main result of [30] states that the largest eigenvalues of A^tA are of the order m^2n^2 and

$$\lim_{n\to\infty} E\left(\det(1 + \frac{z}{m^2n^2}A^tA)\right)^{-1/2} = \lim_{n\to\infty} E\prod_{i=1}^{n}(1 + z\tilde{\lambda_i})^{-1/2} = \exp\left(-\frac{2}{\pi}\sqrt{z}\right) \tag{18}$$

$$= \mathbf{E}\prod_{i=1}^{\infty}(1 + zx_i)^{-1/2}, \tag{19}$$

where $\tilde{\lambda_i},\ i = 1, \ldots, n$ are the eigenvalues of $\frac{1}{m^2n^2}A^tA,\ \ z$ is a complex number with a positive real part, the branch of $\sqrt{z}$ on $D = \{z : \Re z > 0\}$ is such that $\sqrt{1} = 1$. E denotes the mathematical expectation with respect to the random matrix ensemble of sample covariance matrices and $\mathbf{E}$ denotes the mathematical expectation with respect to the inhomogeneous Poisson random point process on the positive half-axis with the intensity $\frac{1}{\pi x^{3/2}}$. The convergence is uniform inside D (i.e. it is unform on the compact subsets of D).
The goal of the paper is to study the distribution of the largest eigenvalues of real symmetric (1) and Hermitian (2) Wigner random matrices in the case of the heavy tails (12). Let us introduce the normalization coefficient b_n defined so that

$$\frac{n^2}{2}G(b_nx) \to \frac{1}{x^\alpha}, \tag{20}$$

for all positive $x > 0$, where G is defined in (12) above. This can be achieved by selecting b_n to be the infimum of all t for which $G(t-0) \geq \frac{2}{n^2} \geq G(t+0)$. It is clear that

$$n^{\frac{2}{\alpha}-\delta} \ll b_n \ll n^{\frac{2}{\alpha}+\delta} \tag{21}$$

for any $\delta > 0$ as $n \to \infty$ and $\frac{n^2h(b_n)}{b_n^\alpha} \to 2$ as $n \to \infty$. Let us denote the eigenvalues of $b_n^{-1}A$ by $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n$. As we will see below this normalization is chosen in such a way that the largest normalized eigenvalues are of the order of constant and the vast majority of the eigenvalues go to zero. The main result of our paper is formulated in the next two theorems.

Theorem 1 *Let A be a Wigner real symmetric (1) or Hermitian (2) random matrix with the heavy tails (12) and λ_1 the largest eigenvalue of $b_n^{-1}A$. Then*

$$\lim_{n\to\infty} \Pr(\lambda_1 \leq x) = \exp(-x^{-\alpha}). \tag{22}$$

Remark 1
It is easy to see (and will be shown later) that the r.h.s. of (22) also gives the limiting distribution of the maximum of the (properly normalized) matrix entries, i.e.

$$\lim_{n\to\infty} \Pr(\max_{1\le i,j\le n} b_n^{-1}|a_{ij}| \le x) = \exp(-x^{-\alpha}). \tag{23}$$

Similarly to Theorem 1.1 one can study the distribution of the second, third, fourth, etc largest eigenvalue of A. The following general result holds.

Theorem 2 *Let A be a Wigner real symmetric (1) or Hermitian (2) random matrix with the heavy tail of the distribution of matrix entries (12). Then the random point configuration composed of the positive eigenvalues of $b_n^{-1}A$ converges in distribution on the cylinder sets to the inhomogeneous Poisson random point process on $(0, +\infty)$ with the intensity $\rho(x) = \frac{\alpha}{x^{1+\alpha}}$.*

Remark 2 As we will see below the result of the Theorem 1.2 can be reformulated in such a way that for any finite integer $k \ge 1$ the joint distribution of the k largest eigenvalues of $b_n^{-1}A$ coincides in the limit $n \to \infty$ with the joint distribution of the first k order statistics of the absolute values of the normalized matrix entries $\{b_n^{-1}|a_{ij}|, \ 1 \le i \le j \le n\}$. This reformulation is also indicative of the method of proof of the results of Theorems 1.1 and 1.2. Of course a similar result about the negative eigenvalues of $b_n^{-1}A$ holds true as well.
We remind the reader that a Poisson random point process on the real line with the locally integrable intensity function $\rho(x)$ is defined in such a way that the counting functions (e.g. numbers of particles) for the disjoint intervals $I_1, \dots, I_k$ are independent Poison random variables with the parameters $\int_{I_j} \rho(x)dx, \ \ j = 1, \dots, k$. Equivalently, one can define the Poisson random point by requiring that the k-point correlations functions are given by the products of the one-point correlation functions (intensities), i.e. $\rho_k(x_1, \dots, x_k) = \prod_{j=1}^{k} \rho(x_j)$.
Remark 3 The arguments of the proof are quite soft and can be extended without any difficulty to banded random matrices and some other classes of Hermitian and real symmetric random matrices with independent entries with heavy tails of distribution.
Remark 4 For random Schrödinger operators the Poisson statistics of the eigenvalues in the localization regime was first proved by Molchanov in [23] (see also the paper by Minami [22]). There is a vast literature on the Poisson statistics of the energy levels of quantum systems in the case of the regular underlying dynamics (see e.g. [3], [27], [16], [25], [5], [18], [19], [20]).
The rest of the paper is organized as follows. We prove Theorem 1.1 in section 2. The proof of Theorem 1.2 is presented in section 3.

2 Proof of Theorem 1.1

We will prove Theorems 1.1 and 1.2 in the real symmetric case. In the Hermitian case the arguments are essentially the same. We start by considering the distribution of the largest matrix entry. It follows from (20) that

$$\Pr(\max_{1\le i\le j\le n} b_n^{-1}|a_{ij}| \le x) = (1 - G(b_n x))^{n(n+1)/2} = \left(1 - \frac{2}{n^2 x^\alpha}\right)^{n(n+1)/2} (1+o(1)) = \exp(-x^{-\alpha})(1+o(1)), \tag{24}$$

(see e.g. [15], Theorem 1.5.1).
Let us order the $N = n(n+1)/2$ i.i.d. random variables $|a_{ij}|, \ \ 1 \le i \le j \le n$, i.e. $|a_{i_1 j_1}| \ge |a_{i_2 j_2}| \ge |a_{i_3 j_3}| \ge \dots \ge |a_{i_N j_N}|$, where (i_l, j_l) are the indices of the l-th largest (in

absolute value) matrix entry. We will use the notation $a^{(l)} = b_n^{-1}|a_{i_l j_l}|, \quad l = 1, \ldots, N$ for the normalized l-th largest absolute value of matrix entries.
In this paper only the statistical properties of a finite number of the largest order statistics will be of importance. In particular, for any finite k the inequalities between the first k order statistics are strict with probability going to 1 as $n \to \infty$. We start with a standard proposition that generalizes (24).

Proposition 3 *Let $0 < c_1 < d_1 < c_2 < d_2 < \ldots c_k < d_k \leq +\infty$, and $I_l = (c_l, d_l), \quad l = 1, \ldots k$, be disjoint intervals on the positive half-line. Then the counting functions $\#(I_l) = \#(1 \leq i \leq j \leq n : \quad b_n^{-1}|a_{ij}| \in I_l), \quad l = 1, \ldots, k$, are independent in the limit $n \to \infty$ and have Poisson distribution with the parameters $\mu_l = \int_{I_l} \rho(x)dx$, where $\rho(x) = \frac{\alpha}{x^{1+\alpha}}$.*

In other words, for any $\epsilon > 0$ the restriction of the point configuration $\{b_n^{-1}|a_{ij}|\}$ to the interval $[\epsilon, +\infty)$ converges in distribution on the cylinder sets to the inhomogeneous Poisson random point process with the intensity $\rho(x) = \frac{\alpha}{x^{1+\alpha}}$ (we refer the reader to [6], sections 2 and 3 for the definition and elementary properties of Poisson random point processes). In particular,

$$\Pr(a^{(k)} \leq x) \to \exp(-x^{-\alpha}) \sum_{l=1}^{k-1} \frac{x^{-l\alpha}}{l!}. \tag{25}$$

The proof of the proposition is a straightforward generalization of the calculations in (24) (see e.g. [15] Theorem 2.3.1).
Now we proceed to the proof of Theorem 1.1. The result of the theorem follows from (24) if we can show that $\frac{\lambda_1}{a^{(1)}} \to 1$ in probability as $n \to \infty$, where as before λ_1 is the largest eigenvalue of $b_n^{-1}A$ and $a^{(1)} = b_n^{-1} \max_{1 \leq i \leq j \leq n} |a_{ij}|$.
We start with an elementary lemma.

Lemma 4 *a) With probability going to 1 there are no diagonal entries greater (in absolute value) than $b_n^{11/20}$.*
b) With probability going to 1 there is no pair $(i, j), \quad 1 \leq i < j \leq n$, such that $|a_{ij}| > b_n^{99/100}$ and $|a_{ii}| + |a_{jj}| > b_n^{1/10}$.
c) For any positive constant $\delta > 0$ with probability going to 1 there is no row $1 \leq i \leq n$ that has at least two entries greater in absolute value than $b_n^{\frac{1}{2}+\delta}$.

For the convinience of the reader we sketch the proof of the lemma. We start with part a). We first remind the reader that by a classical result by Karamata ([14], see also [26] for a nice exposition of the subject) a slowly varying function at infinity can be respresented on the interval $[B, +\infty)$, where B is sufficiently large, as

$$h(x) = \exp\left(\eta(x) + \int_B^{+\infty} \frac{\varepsilon(t)}{t} dt\right), \tag{26}$$

where $\eta(x)$ is a bounded measurable function which has a limit at infinity and $\varepsilon(x)$ is a continuous function on $[B, +\infty)$ such that $\lim_{x\to\infty} \varepsilon(x) = 0$. Using (12), (21) and (26) we obtain

$$\begin{aligned}\Pr\left(\max_{1\leq i\leq n} |a_{ii}| \leq b_n^{11/20}\right) &= \left(1 - G(b_n^{11/20})\right)^n = \left(1 - \frac{h(b_n^{11/20})}{b_n^{\frac{11\alpha}{20}}}\right)^n \geq \\ &\geq \exp(-n^{-\frac{1}{40}})(1 + o(1)).\end{aligned} \tag{27}$$

In a similar fashion, in order to prove the statement b) we have to show that $\frac{n(n-1)}{2}G(b_n^{99/100}) \times G(b_n^{1/10}) \leq n^{-1/20}(1+O(1))$, which again follows from (12), (21) and (26). The proof of part c) is very similar and, therefore, left to the reader.
Let (i_1, j_1) be the indices of the maximal (in absolute value) matrix element. It follows from (24) and part a) of Lemma 1 that with probability going to 1 one has that $|a_{i_1 j_1}| \geq b_n^{99/100}$ and $i_1 \neq j_1$. Let f_1 be a unit vector in $\mathbb{R}^n$ such that all its coordinates except the i_1-th and j_1-th are zero, the i_1-th coordinate is 1 and the j_1-th coordinate is $+1$ if $a_{i_1 j_1}$ is nonnegative and -1 otherwise. Then one can easily calculate tha value of the quadratic form $(A \times f_1, f_1) = |a_{i_1 j_1}| + \frac{1}{2}a_{i_1 i_1} + \frac{1}{2}a_{j_1 j_1}$, and it follows from (24) and Lemma 1 b) that with probability going to 1 the sum of the last two terms is much smaller than the first term, and, in particular, $(Af_1, f_1) = |a_{i_1 j_1}|(1 + o(b_n^{-4/5}))$. Since f_1 is a unit vector, we clearly have $\lambda_1 \geq (Af_1, f_1)$. To show that $|a_{i_1 j_1}|(1 + o(1))$ is also an upper bound on the largest eigenvalue we need one more lemma.

Lemma 5 *With probability going to 1 a Wigner matrix (1), (12) does not have a row such that both the maximum of the absolute values of the matrix elements in that row and the sum of the absolute values of the remaining elements in the row are at least $b_n^{\frac{1}{2}+\frac{\alpha}{4}}$. In other words,*

$$\Pr\{\exists i, \quad 1 \leq i \leq n: \quad \max_{1\leq j\leq n} |a_{ij}| > b_n^{\frac{1}{2}+\frac{\alpha}{4}}, \quad \left(\sum_{1\leq j\leq n} |a_{ij}|\right) - \max_{1\leq j\leq n} |a_{ij}| > b_n^{\frac{1}{2}+\frac{\alpha}{4}}\} \to 0 \quad (28)$$

as $n \to \infty$.

We will treat separately the two cases $1 < \alpha < 2$ and $0 < \alpha \leq 1$.
Let us start with the first case, i.e. $1 < \alpha < 2$. We choose $\epsilon = \frac{1}{2M+1}$, where M is some integer, so that $\epsilon < \frac{1}{4} - \frac{\alpha}{8}$. We claim that there exists sufficiently small positive constant γ $(\gamma = \frac{1}{4M+2}$ would suffice) that for any row $1 \leq i \leq n$ and for all integers $0 \leq k \leq M$

$$E\left(\#\left(1 \leq j \leq n : |a_{ij}| \geq b_n^{k\epsilon}\right)\right) = nG(b_n^{k\epsilon}), \quad (29)$$
$$\Pr\left(\#\left(1 \leq j \leq n : |a_{ij}| \geq b_n^{k\epsilon}\right) \geq 2nG(b_n^{k\epsilon})\right) \leq \exp(-n^{\gamma}) \quad (30)$$

for sufficiently large n. Indeed, (29) immediately follows from the definition of G in (12), while (30) follows from the Chernoff's inequality $\Pr(X \geq EX + t) \leq \exp(-\frac{t^2}{2np} + \frac{t^3}{6(np)^2})$ for a binomial random variable $X \sim Bin(n,p)$ and (12), (21). It then follows from (30) that with probability $1 - O(n\exp(-n^{\gamma}))$ we have that for all rows $1 \leq i \leq n$ of the matrix A

$$\sum_{j:|a_{ij}|\leq b_n^{\frac{M+1}{2M+1}}} |a_{ij}| \leq \sum_{k=0}^{M} \#\left(1 \leq j \leq n : |a_{ij}| \geq b_n^{k\epsilon}\right) b_n^{(k+1)\epsilon} \leq$$
$$\sum_{k=0}^{M} b_n^{(k+1)\epsilon} 2nG(b_n^{k\epsilon}) = \sum_{k=0}^{M} b_n^{(k+1)\epsilon} 2nh(b_n^{k\epsilon}) b_n^{-\alpha k\epsilon} \leq n b_n^{\frac{1}{4}-\frac{\alpha}{8}} \leq b_n^{\frac{1}{2}+\frac{\alpha}{4}} \quad (31)$$

for sufficiently large n. Since by part c) of Lemma 1 with probability going to 1 each row has at most one entry greater in absolute value than $b_n^{\frac{M+1}{2M+1}}$ the bound (31) implies the statement of the Lemma 1.2. in the case $1 < \alpha < 2$.

In the case $0<\alpha<1$ we again choose $\epsilon = \frac{1}{2M+1}$, where M is some integer in such a way that $\epsilon < \frac{\alpha}{8}$ and $\gamma = \frac{1}{4M+2}$ so that with probability $1-O(n\exp(-n^{\gamma}))$ we have that for all rows $1 \le i \le n$ of the matrix A

$$\sum_{j:|a_{ij}|\le b_n^{\frac{M+1}{2M+1}}} |a_{ij}| \le \sum_{k=0}^{M} \#\left(1\le j\le n : |a_{ij}| \ge b_n^{k\epsilon}\right) b_n^{(k+1)\epsilon} \le$$
$$\sum_{k=0}^{M} b_n^{(k+1)\epsilon} 2nG(b_n^{k\epsilon}) = \sum_{k=0}^{M} b_n^{(k+1)\epsilon} 2nh(b_n^{k\epsilon}) b_n^{-\alpha k\epsilon} \le b_n^{\epsilon} 2n b_n^{\frac{1-\alpha}{2}} \le b_n^{\frac{1}{2}+\frac{\alpha}{4}} \tag{32}$$

for sufficiently large n. Lemma 2 is proven.
It follows from (24) and Lemma 2 (in particular (31), (32) that

$$\|A\|_{\infty} := \max_{1\le i\le n} \sum_{1\le j\le n} |a_{ij}| = \max_{1\le i\le j\le n} |a_{ij}|(1+o(1)) \tag{33}$$

with probability going to 1. The Theorem 1.1 now follows from the fact that the matrix norm $\|A\|_{\infty}$ is an upper bound for the largest eigenvalue of the matrix A.

3 Proof of Theorem 1.2

Let as in section 2 $a^{(l)} = b_n^{-1}|a_{i_l j_l}|$, $l=1,\ldots,N=n(n+1)/2$ be the values $b_n^{-1}\times |a_{ij}|$, $1\le i\le j\le n$ put in the decreasing order. In particular, $a^{(1)} = \max_{ij} b_n^{-1}|a_{ij}|$. Similarly to section 2 we construct the unit vectors $f_l \in \mathbb{R}^n$, $l=1,2,\ldots,$ such that all coordinates of f_l except the i_l-th and j_l-th are zero, the i_l-th coordinate is 1 and the j_l-th coordinate is $+1$ if $a_{i_l j_l}$ is nonnegative and -1 otherwise. Then $Af_l = |a_{i_l j_l}| f_l + \frac{1}{\sqrt{2}} \times \sum_{m\ne i,j} (a_{i_l m} + a_{j_l m}) e_m$, where e_m are the standard basis vectors in $\mathbb{R}^n$, i.e. all coordinates of e_m are zero, except for the m-th one which is 1. Since $\left(\sum_{m\ne i,j} a_{i_l m}^2\right)^{1/2} \le \sum_{m\ne i,j} |a_{i_l m}|$, it follows from Proposition 1 and Lemma 2 that for any finite number of the largest values $a^{(l)}$, $1\le l\le k$ one has that $b_n^{-1}Af_l = a^{(l)}f_l + r_l$, where the euclidian norms of r_l are o(1). Therefore by an elementary argument in perturbation theory for Hermitian operators we obtain that $b_n^{-1}A$ has eigenvalues $a^{(l)}(1+o(1))$, $1\le l\le k$, for any finite k. The last thing we need to show is that (with probability going to 1) these eigenvalues are exactly the k largest eigenvalues of $b_n^{-1}A$. In other words, so far we proved that with probability going to 1 one has $\lambda_l \ge a^{(l)}(1+o(1))$, $l=1,\ldots,k$. Our goal now is to prove the reverse inequalities $\lambda_l \le a^{(l)}(1+o(1))$, $l=1,\ldots,k$. To simplify the exposition we will prove the result for the second eigenvalue. The reasoning in the general case is very similar. Let, as before, (i_1, j_1) be the indices of the largest in absolute value matrix element. Consider an $(n-1)\times(n-1)$ submatrix obtained from A by deleting the i_1-th row and the i_1-th column. Let us denote the submatrix by B. It follows from Lemma 1 that $i_2 \ne i_1$, $j_2 \ne j_1$ with probability going to 1. By the same reasoning as in Theorem 1 we have that the largest eigenvalue of $b_n^{-1}B$ is $a^{(2)}(1+o(1))$, but by the interlacing property of the eigenvalues of A and B we have that the largest eigenvalue of B is not smaller than the second largest eigenvalue of A. Theorem 1.2 is proven.

References

[1] L.Arnold, On the asymptotic distribution of the eigenvalues of random matrices, *J. Math. Anal. Appl.*, **20**, 262-268, (1967).

[2] Z.Burda et. al., Free random Lévy matrices, *Phys Rev E*, **65**, 021106, (2002).

[3] M.V.Berry, M. Tabor, Level clustering in the regular spectrum, *Proc. R. Soc. London Ser. A*, **356**, 375-394, (1977).

[4] P.Cizeau, J.P.Bouchaud, Theory of Lévy matrices, *Phys Rev E*, **50**, 1810-1822, (1994).

[5] Z.Cheng, J.L.Lebowitz and P.Major, On the number of lattice points between two enlarged and randomly shifted copies of an oval, *Prob. Theo. Rel. Fields*, **100**, 253-268, (1994).

[6] D.J. Daley, D.Vere-Jones, **An Introduction to the Theory of Point Processes**, vol.I, 2nd edition, Springer, New York, 2003.

[7] P.Deift **Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach**, Courant Lecture Notes in Mathematics, Vol. 3, New York, 1999.

[8] Z. Füredi and J. Komlós, The eigenvalues of random symmetric matrices, *Combinatorica*, **1, No. 3**, 233-241, (1981).

[9] P. Forrester, The spectral edge of random matrix ensembles, *Nucl. Phys. B*, **402**, 709-728, (1994).

[10] V.L. Girko, **Spectral Theory of Random Matrices**, (in Russian), Moscow, Nauka, 1988.

[11] I.A.Ibragimov, Yu.V.Linnik, **Independent and Stationary Sequences of Random Variables**, translation from the Russian edited by J.F.C.Kingman, Wolters-Noordhoff Publishing, Groningen, 1971.

[12] R.A.Janik, New multicritical random matrix ensembles, *Nuclear Phys B*, **635**, 492-504 (2002).

[13] K.Johansson, Universality of the Local Spacing Distribution in Certain Ensembles of Hermitian Wigner Matrices, *Commun. Math. Phys.*, **215**, 683-705, (2001).

[14] J. Karamata, Sur un mode de croissance régulière des fonctions, *Mathematica (Cluj)*, **4**, 38-53, (1930).

[15] M.R. Leadbetter, G.Lindgren and H. Rootzén, **Extremes and Related Properties of Random Sequences and Processes**, Springer-Verlag, New York, 1983.

[16] P.Major, Poisson law for the number of lattice points in a random strip with finite area, *Prob. Theo. Rel. Fields*, **92**, 423-464, (1992).

[17] V.A. Marchenko, L.A. Pastur, Distribution of some sets of random matrices, *Math. USSR-Sb.* **1**, 457-483, (1967).

[18] J. Marklof, Pair correlation densities of inhomogeneous quadratic forms, *Annals of Mathematics*, **158**, 419-471, (2003).

[19] J. Marklof, Pair correlation densities of inhomogeneous quadratic forms II, *Duke Mathematical Journal*, **115**, 409-434, (2002).

[20] J. Marklof, The Berry-Tabor conjecture, *Proceedings of the 3rd European Congress of Mathematics, Barcelona 2000, Progress in Mathematics 202*, 421-427, (2001).

[21] M.L.Mehta, **Random Matrices**, Academic Press, New York, 1991.

[22] N.Minami, Local fluctuation of the spectrum of a multidimensional Anderson tight binding model, *Commun. Math. Phys.*, **177**, 709-725, (1996).

[23] S.A.Molchanov, The local structure of the spectrum of the one-dimensional Schrödinger operator, *Commun. Math. Phys.* **78**, 429-446, (1981).

[24] L.A. Pastur, On the spectrum of random matrices, *Teor. Mat. Fiz.*, **10**, 102-112, (1972).

[25] P.Sarnak, Values at integers of binary quadratic forms, *Harmonic Analysis and Number Theory* (Montreal, PQ, 1996), 181-203, CMS Conf. Proc. **21**, Amer. Math. Soc., Providence, RI, 1997.

[26] E.Seneta, **Regularly Varying Functions**, Lecture Notes in Mathematics **508** (eds. A.Dold and B.Eckmann), Springer, New York, 1976.

[27] Ya.Sinai, Poisson distribution in a geometric problem, *Adv. Sov. Math., AMS Publ.*, **3**, 199-215, (1991).

[28] A.Soshnikov, Universality at the edge of the spectrum in Wigner random matrices, *Commun. Math. Phys.*, **207**, 697-733, (1999).

[29] A.Soshnikov, A Note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices, *J. Stat. Phys.*, **108, Nos. 5/6**, 1033-1056, (2002), a special issue dedicated to 65-th birthdays of David Ruelle and Yakov Sinai.

[30] A.Soshnikov, Y.Fyodorov, On the singular values of random matrices with independent Cauchy entries, arXiv preprint math.PR/0403425.

[31] C.A.Tracy, H.Widom, Level-spacing distribution and the Airy kernel, *Commun. Math. Phys.*, **159**, 151-174, (1994).

[32] C.A.Tracy, H.Widom, On orthogonal and symplectic random matrix ensembles, *Commun. Math. Phys.*, **177**, 724-754, (1996).

[33] E.Wigner, Characteristic vectors of bordered matrices with infinite dimensions, *Ann. of Math.*, **62**, 548-564, (1955).

[34] E.Wigner, On the distribution of the roots of certain symmetric matrices, *Ann. of Math.*, **67**, 325-328, (1958).

[35] E.Wigner, Random matrix theory in physics, *SIAM Rev.*, **9**, 1-23, (1967)