

# On the local semicircular law for Wigner ensembles

FRIEDRICH GÖTZE<sup>1</sup>, ALEXEY NAUMOV<sup>2,3,4</sup>,  
ALEXANDER TIKHOMIROV<sup>5,6,\*</sup> and DMITRY TIMUSHEV<sup>5,\*\*</sup>

<sup>1</sup>*Faculty of Mathematics, Bielefeld University, P.O. Box 10 01 31, 33501, Bielefeld, Germany.  
E-mail: goetze@math.uni-bielefeld.de*

<sup>2</sup>*Skolkovo Institute of Science and Technology (Skoltech), Skolkovo Innovation Center, Building 3, 143026, Moscow, Russia. E-mail: a.naumov@skoltech.ru*

<sup>3</sup>*National Research University Higher School of Economics, 20 Myasnitskaya ulitsa, 101000, Moscow, Russia*

<sup>4</sup>*Institute for Information Transmission Problems RAS, Bolshoy Karetny per. 19, bld. 1, 127051, Moscow, Russia*

<sup>5</sup>*Department of Mathematics, Komi Science Center of Ural Division of RAS, Chernova 3a, 167000, Syktyvkar, Russia. E-mail: \* tikhomirov@dm.komisc.ru; \*\* timushev@dm.komisc.ru*

<sup>6</sup>*Syktyvkar State University, Oktyabrsky prospect, 55, 167001, Syktyvkar, Komi Republic, Russia*

We consider a random symmetric matrix  $\mathbf{X} = [X_{jk}]_{j,k=1}^n$  with upper triangular entries being i.i.d. random variables with mean zero and unit variance. We additionally suppose that  $\mathbb{E}|X_{11}|^{4+\delta} =: \mu_{4+\delta} < \infty$  for some  $\delta > 0$ . The aim of this paper is to significantly extend a recent result of the authors Götze, Naumov and Tikhomirov (2015) and show that with high probability the typical distance between the Stieltjes transform of the empirical spectral distribution (ESD) of the matrix  $n^{-\frac{1}{2}}\mathbf{X}$  and Wigner's semicircle law is of order  $(nv)^{-1} \log n$ , where  $v$  denotes the distance to the real line in the complex plane. We apply this result to the rate of convergence of the ESD to the distribution function of the semicircle law as well as to rigidity of eigenvalues and eigenvector delocalization significantly extending a recent result by Götze, Naumov and Tikhomirov (2015). The result on delocalization is optimal by comparison with GOE ensembles. Furthermore the techniques of this paper provide a new shorter proof for the optimal  $O(n^{-1})$  rate of convergence of the expected ESD to the semicircle law.

*Keywords:* delocalization; local semicircle law; mean spectral distribution; random matrices; rate of convergence; rigidity; Stieltjes transform

## 1. Introduction and main result

Let  $\mathbf{X} = [X_{jk}]_{j,k=1}^n$  be a random Hermitian matrix with upper triangular entries being independent random variables with mean zero and unit variance. Denote the  $n$  eigenvalues of the symmetric matrix  $\mathbf{W} := \frac{1}{\sqrt{n}}\mathbf{X}$  in the increasing order by

$$\lambda_1(\mathbf{W}) \leq \dots \leq \lambda_n(\mathbf{W})$$

and introduce the eigenvalue counting function  $N_I(\mathbf{W}) := |\{1 \leq k \leq n : \lambda_k(\mathbf{W}) \in I\}|$  for any interval  $I \subset \mathbb{R}$ , where  $|A|$  denotes the number of elements in the set  $A$ . The pioneering result

of [38] states that for any interval  $I \subset \mathbb{R}$  of fixed length and independent of  $n$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} N_I(\mathbf{W}) = \int_I g_{\text{sc}}(\lambda) d\lambda, \tag{1.1}$$

where  $g_{\text{sc}}(\lambda) := \frac{1}{2\pi} \sqrt{(4 - \lambda^2)_+}$  and  $(x)_+ := \max(x, 0)$ . Wigner considered the special case when all  $X_{jk}$  take only two values  $\pm 1$  with equal probabilities. Later on the result (1.1) was called *Wigner’s semicircle law* and has been extended in various aspects, see, for example, [2,15, 24,32,33] and [20]. For an extensive list of references, we refer to the monographs [1,4] and [35]. In what follows, we call *Wigner’s semicircle law or semicircle law* not only a result of type (1.1), but the limiting probability distribution as well.

If an interval  $I$  is of fixed length, independent of  $n$ , it typically contains a macroscopically large number of eigenvalues, which means a number of order  $n$ . In this case we may rewrite (1.1) in the following form

$$\frac{1}{n|I|} \mathbb{E} N_I(\mathbf{W}) = \frac{1}{|I|} \int_I g_{\text{sc}}(\lambda) d\lambda + o\left(\frac{1}{|I|}\right). \tag{1.2}$$

It is of interest to investigate the case of smaller intervals where the number of eigenvalues cease to be macroscopically large. In this case the second term on the right-hand side of (1.2) needs to be refined. An appropriate analytical tool for asymptotic approximations is the Stieltjes transform of the empirical spectral distribution function  $F_n(x) := \frac{1}{n} N_{(-\infty, x]}(\mathbf{W})$ , which is given by

$$m_n(z) := \int_{-\infty}^{\infty} \frac{dF_n(\lambda)}{\lambda - z} = \frac{1}{n} \text{Tr}(\mathbf{W} - z\mathbf{I})^{-1} = \frac{1}{n} \sum_{j=1}^n \frac{1}{\lambda_j(\mathbf{W}) - z},$$

where  $z = u + iv, v > 0$ . For the imaginary part of  $m_n(z)$ , we get

$$\text{Im} m_n(u + iv) = \int_{-\infty}^{\infty} \frac{v}{(\lambda - u)^2 + v^2} dF_n(\lambda) = \frac{1}{v} \int_{-\infty}^{\infty} \mathcal{P}\left(\frac{u - \lambda}{v}\right) dF_n(\lambda)$$

which is a kernel density estimator with a Poisson kernel  $\mathcal{P}(x)$  and bandwidth  $v$ . For a meaningful estimator of the spectral density, we cannot allow the distance  $v$  to the real line, that is the bandwidth of the kernel density estimator, to be smaller than the typical  $\frac{1}{n}$ -distance between eigenvalues. Hence, in what follows we shall be mostly interested in the situations when  $v \geq \frac{c}{n}, c > 0$ , where in some situations  $c$  may depend on  $n$ , growing for example like  $\log n$ .

Under rather general conditions, one can establish the convergence of  $m_n(z)$  for fixed  $v > 0$  to the Stieltjes transform of Wigner’s semicircle law which is given by

$$s(z) = \int_{-\infty}^{\infty} \frac{g_{\text{sc}}(\lambda) d\lambda}{\lambda - z} = -\frac{z}{2} + \sqrt{\frac{z^2}{4} - 1}.$$

It is much more difficult to establish the convergence in the region  $1 \geq v \geq \frac{c}{n}$ . Significant progress in that direction was recently made in a series of results by Erdős, Schlein and Yau [12–14], Erdős

et al. [10], showing that with high probability uniformly in  $u \in \mathbb{R}$

$$|m_n(u + iv) - s(u + iv)| \leq \frac{\log^\beta n}{nv}, \quad \beta > 0, \tag{1.3}$$

which they called *local semicircle law*. It means that the fluctuations of  $m_n(z)$  around  $s(z)$  are of order  $(nv)^{-1}$  (up to a logarithmic factor). The value of  $\beta$  depends on  $n$ , more precisely on  $\beta := \beta_n = c \log \log n$ , where  $c > 0$  denotes some constant. In in the papers [12–14] the inequality (1.3) has been shown assuming that the distribution of  $X_{jk}$  has sub-exponential tails for all  $1 \leq j, k \leq n$ . Moreover in [10] this assumption had been relaxed to requiring  $\mathbb{E}|X_{jk}|^p \leq \mu_p$  for all  $p \geq 1$ , where  $\mu_p$  are some constants.

Without loss of generality, we will assume in what follows that  $\mathbf{X}$  is a real symmetric matrix. We say that the conditions **(C0)** hold if:

- $X_{jk}, 1 \leq j \leq k \leq n$ , are i.i.d. real random variables;
- $\mathbb{E} X_{11} = 0, \mathbb{E} X_{11}^2 = 1$ ;
- $\mathbb{E}|X_{11}|^{4+\delta} =: \mu_{4+\delta} < \infty$  for some  $\delta > 0$ .

Our results proven below apply to the case of Hermitian matrices as well. Here we may additionally assume for simplicity that  $\text{Re } X_{jk}$  and  $\text{Im } X_{jk}$  are independent r.v. for all  $1 \leq j < k \leq n$ . Otherwise one needs to extend the moment inequalities for linear and quadratic forms in complex r.v. (see [18], Theorem A.1–A.2) to the case of dependent real and imaginary parts, the details of which we omit.

The result (1.3) under the conditions **(C0)** was proved in a series of papers [9,11,31] with an  $n$ -dependent value  $\beta = c \log \log n$ . In [18], we gave a self-contained proof based on the methods developed in [26,28] while at the same time reducing the power of  $\log n$  from  $\beta = c \log \log n$  to  $\beta = 2$ . Our work and some crucial bounds of our proof were motivated by the methods used in a recent paper of Cacciapuoti, Maltsev and Schlein [8], where the authors improved the log-factor dependence in (1.3) in the sub-Gaussian case. Let  $\kappa := \frac{\delta}{2(4+\delta)}$ . In [18], Theorem 1.1, it is shown, assuming conditions **(C0)**, that for any fixed  $V > 0$  there exist positive constants  $A_0, A_1$  and  $C$  depending on  $\delta$  and  $V$  such that

$$\mathbb{E}|m_n(z) - s(z)|^p \leq \left(\frac{Cp^2}{nv}\right)^p, \tag{1.4}$$

for all  $1 \leq p \leq A_1(nv)^\kappa, V \geq v \geq A_0n^{-1}$  and  $|u| \leq 2 + v$ . Note that when stating that some constant  $C$  depends on  $\delta$  we actually mean that it may depends on  $\mu_{4+\delta}$  as well, that is  $C = C(\delta, \mu_{4+\delta})$ . Applying Markov’s inequality, we may rewrite (1.4) in the following form

$$\mathbb{P}\left(|m_n(z) - s(z)| \geq \frac{K}{nv}\right) \leq \left(\frac{Cp^2}{K}\right)^p, \tag{1.5}$$

for all  $1 \leq p \leq A_1(nv)^\kappa, V \geq v \geq A_0n^{-1}$  and  $|u| \leq 2 + v$ . For applications, the range of  $v$ , such that (1.4) holds for fixed  $p$  is relevant. It is clear that  $V \geq v \geq Cp^{\frac{1}{\kappa}}n^{-1}$ . Since we are interested in error probabilities which are polynomially small only, it suffices to choose  $p$  to be of order  $\log n$ , which implies that  $V \geq v \geq Cn^{-1} \log^{\frac{1}{\kappa}} n$ . At the same time the constant  $K$  in (1.5) should be of order  $\log^2 n$ . Comparing with (1.3), we get  $\beta = 2$ .

In the region  $|u| > 2 + v$ , we can control the imaginary part of  $m_n(s)$  only. It was proved in [18], Theorem 1.1, that for any  $u_0 > 0$  there exist positive constants  $A_0, A_1$  and  $C$  depending on  $u_0, V$  and  $\delta$  such that

$$\mathbb{E}|\operatorname{Im} m_n(z) - \operatorname{Im} s(z)|^p \leq \left(\frac{Cp^2}{nv}\right)^p, \tag{1.6}$$

for all  $1 \leq p \leq A_1(nv)^\kappa, V \geq v \geq A_0n^{-1}$  and  $|u| \leq u_0$ .

As mentioned above, we are interested in the case when  $p$  is of order  $\log n$  which implies  $V \geq v \geq n^{-1} \log^{\frac{1}{\kappa}} n$ . This choice yields that in our applications all bounds will depend on  $\log^{\frac{1}{\kappa}} n$ . The power  $\kappa^{-1}$ , which is independent of  $n$ , may be rather large for  $\delta$  near zero. The aim of the current paper is to strengthen the results of [18,19] by proving bounds of type (1.3) with  $\beta = 1$  while at the same time showing that  $p$  and  $v$  may be taken of order  $\log n$  and  $n^{-1} \log n$ , respectively. This is done in Theorem 1.1 below. To this end, we combine our techniques from [18] with fruitful ideas from [9] and [31] in particular their moment matching technique. A crucial result in that direction is Lemma 3.1. See the sketch of proof below in Section 1.2. It still remains one challenging open problem, namely extending the bounds to weaken the moment condition to  $\delta = 0$ .

### 1.1. Main result

Let us introduce the following notation

$$\Lambda_n(z) := m_n(z) - s(z), \quad z = u + iv.$$

The main result of this paper is the following theorem, which estimates the fluctuations (1.3) and strengthens (1.4) and (1.6).

**Theorem 1.1.** *Assume that the conditions (C0) hold and let  $V > 0$  be some constant.*

(i) *There exist positive constants  $A_0, A_1$  and  $C$  depending on  $V$  and  $\delta$  such that*

$$\mathbb{E}|\Lambda_n(z)|^p \leq \left(\frac{Cp}{nv}\right)^p,$$

for all  $1 \leq p \leq A_1 \log n, V \geq v \geq A_0n^{-1} \log n$  and  $|u| \leq 2 + v$ .

(ii) *For any  $u_0 > 0$  there exist positive constants  $A_0, A_1$  and  $C$  depending on  $u_0, V$  and  $\delta$  such that*

$$\mathbb{E}|\operatorname{Im} \Lambda_n(z)|^p \leq \left(\frac{Cp}{nv}\right)^p,$$

for all  $1 \leq p \leq A_1 \log n, V \geq v \geq A_0n^{-1} \log n$  and  $|u| \leq u_0$ .

As a consequence of this result, we may show that similarly to (1.5) for all  $K > 0$

$$\mathbb{P}\left(|\Lambda_n(z)| \geq \frac{K}{nv}\right) \leq \left(\frac{Cp}{K}\right)^p, \tag{1.7}$$

valid for all  $1 \leq p \leq A_1 \log n$ ,  $V \geq v \geq A_0 n^{-1} \log n$  and  $|u| \leq 2 + v$ . Taking  $p$  and  $K$  of order  $\log n$ , we may guarantee that (1.7) is less than, for example,  $n^{-2}$ . Thus, comparing with (1.3), we get  $\beta = 1$ .

Applications of Theorem 1.1 outside the limit spectral interval, that is for  $|u| \geq 2$ , require stronger bounds on  $\text{Im } \Lambda_n$ . Let us denote

$$\gamma := \gamma(u) := ||u| - 2| \tag{1.8}$$

and introduce the following quantity

$$\alpha := \alpha(\delta) = \frac{2}{4 + \delta},$$

which will control the level of truncation of the matrix entries. We say that the set of conditions **(C1)** holds if:

- **(C0)** are satisfied;
- $|X_{jk}| \leq Dn^\alpha$ ,  $1 \leq j, k \leq n$ , where  $D := D(\alpha)$  denotes some positive constant depending on  $\alpha$  only.

By definition  $\kappa = \delta/(2(4 + \delta))$ , hence we may write  $\kappa = \frac{1-2\alpha}{2}$ .

**Theorem 1.2.** *Assume that the conditions **(C1)** hold and  $u_0 > 2$  and  $V > 0$ . There exist positive constants  $A_0, A_1$  and  $C$  depending on  $\delta, u_0$  and  $V$  such that*

$$\mathbb{E}|\text{Im } \Lambda_n(z)|^p \leq \frac{C^p p^p}{n^p(\gamma + v)^p} + \frac{C^p p^{2p}}{(nv)^{2p}(\gamma + v)^{\frac{p}{2}}} + \frac{C^p}{n^p v^{\frac{p}{2}}(\gamma + v)^{\frac{p}{2}}} + \frac{C^p p^{\frac{p}{2}}}{(nv)^{\frac{3p}{2}}(\gamma + v)^{\frac{p}{4}}},$$

for all  $1 \leq p \leq A_1(nv)^\kappa$ ,  $V \geq v \geq A_0 n^{-1}$  and  $2 \leq |u| \leq u_0$ .

### 1.2. Sketch of the proof of Theorem 1.1 and Theorem 1.2

The proof of Theorem 1.1 is similar to the proof [18], Theorem 1.1. Applying Lemma B.1 in [18] see (2.9) below and [8], Proposition 2.2, it is shown in Section 2 that one may estimate  $\mathbb{E}|\Lambda_n(z)|^p$  or  $\mathbb{E}|\text{Im } \Lambda_n(z)|^p$  (depending on  $\text{Re}(z)$  being near or far from the spectral interval  $[-2, 2]$ ) by the moments  $\mathbb{E}|T_n(z)|^p$  (see definition (2.6)).

To estimate  $\mathbb{E}|T_n(z)|^p$ , we may repeat all the steps of the proof of Theorem 2.1 in [18] with one important modification. One of the crucial steps in the proof is the bound  $\mathbb{E}|\mathbf{R}_{jj}(z)|^p \leq C_0^p$ . It was shown in [18], Lemma 4.1, that this bound holds for all  $V \geq v \geq A_0 n^{-1}$  and  $1 \leq p \leq A_1(nv)^\kappa$ . Since we have to choose  $p$  of order  $\log n$  we need to prove bounds in the region  $v \geq \tilde{v} := n^{-1} \log^{\frac{1}{\kappa}} n$ . In order to close the gap in  $v$  from  $\tilde{v}$  to  $v_0 := n^{-1} \log n$  with  $p$  being still of order  $\log n$  we apply the following strategy. We start from the fact that  $\mathbb{E}|\mathbf{R}_{jj}(z)|^p \leq C_0^p \log^{(\frac{1}{\kappa}-1)p} n$  for all  $v \geq v_0$ , see Lemma 3.3. In order to remove the logarithmic factor from the r.h.s. of the last bound, we apply ideas motivated by moment matching techniques used in [9] and [31]. That is we construct a symmetric random matrix  $\mathbf{Y} := [Y_{jk}]_{jk}^n$  such that  $\mathbb{E}X_{jk}^s = \mathbb{E}Y_{jk}^s$ ,  $1 \leq s \leq 4$  and  $Y_{jk}$  are sub-Gaussian random variables, see Lemma 3.4 (see Lemma 5.1 [31]). Then we show

in Lemma 3.5 that the bound  $\mathbb{E} |\mathbf{R}_{jj}(z)|^p \leq C^p + \mathbb{E} |\mathbf{R}_{jj}^{\mathbf{Y}}(z)|^p$  still holds for all  $v \geq v_0$  and  $p$  of order  $\log n$ , where  $\mathbf{R}^{\mathbf{Y}}$  denotes  $\mathbf{R}$  with the  $\mathbf{X}$  matrix being replaced by the  $\mathbf{Y}$  matrix. Our technique in the proof of Lemma 3.5 is a Stein type method. Finally, using the sub-gaussian properties of  $\mathbf{Y}$  we show in Lemma 4.1 that  $\mathbb{E} |\mathbf{R}_{jj}^{\mathbf{Y}}(z)|^p \leq C_0^p$  for all  $v \geq v_0$  and  $1 \leq p \leq c \log n$ , for some constant  $c$ . All these arguments rely on the proof of Lemma 3.1. Note in contrast that in [31] the moment matching techniques were used to estimate the distance  $m_n(z) - s(z)$  directly combined with a combinatorial approach (see, for example, Lemma 5.1 and Lemma 5.2 in [31]).

The proof of Theorem 1.2 is based on the same arguments as the proof of Theorem 1.2 in [18]. Note that  $\mathbb{E} |T_n(z)|^p$  is bounded in terms of  $\mathbb{E} \text{Im}^p \mathbf{R}_{jj}$ . The same arguments as for  $\mathbb{E} |\mathbf{R}_{jj}(z)|^p$  will give us the bound  $\mathbb{E} \text{Im}^p \mathbf{R}_{jj} \leq H_0^p \text{Im}^p s(z) + H_0^p p^p (nv)^{-p}$  valid for some big constant  $H_0 > 0$  and  $v \geq v_0, 1 \leq p \leq c \log n$ . Since we can derive explicit bounds for  $\text{Im} s(z)$  inside as well as outside the limit spectrum we are able to control the size of  $\mathbb{E} |T_n(z)|^q$  as well as of  $\mathbb{E} |\text{Im} \Lambda_n(z)|^p$  on the whole real line in terms of the quantity  $\gamma$  (see (1.8)). This is a another key argument for the proof of Theorem 1.2.

### 1.3. Delocalization, rigidity and rate of convergence

In this section, we present results about delocalization of eigenvectors, rate of convergence of empirical spectral distribution function (ESD) to the semicircle law and rigidity of eigenvalues. These results strengthen the corresponding results in [19].

We start this section by showing delocalization of eigenvectors. This question has been intensively studied in many papers, for example, in [11,13,25] and [9].

Let us denote by  $\mathbf{u}_j := (u_{j1}, \dots, u_{jn})$  the eigenvectors of  $\mathbf{W}$  corresponding to the eigenvalue  $\lambda_j(\mathbf{W})$ .

**Theorem 1.3.** *Assume that conditions (C0) hold with  $\delta = 4$ . For any  $0 < \phi < 2$  there exist positive constants  $C$  and  $C_1$  depending on  $\phi$  and  $\mu_8$  such that for any*

$$\mathbb{P} \left( \max_{1 \leq j, k \leq n} |u_{jk}|^2 \geq \frac{C_1 \log n}{n} \right) \leq \frac{C}{n^{2-\phi}}.$$

We remark here that it is possible to relax the moment conditions to the case  $0 < \delta < 4$  as well. But here we may only conclude that there exists some constant  $c(\delta) > 0$  depending on  $\delta$  such that

$$\mathbb{P} \left( \max_{1 \leq j, k \leq n} |u_{jk}|^2 \geq \frac{C_1 \log n}{n} \right) \leq \frac{C}{n^{c(\delta)}}.$$

A comparison with a similar result for the GOE ensemble (see [1], Corollary 2.5.4) and the delocalization of eigenvectors of the unit sphere shows that this result is optimal with respect to the power of logarithm. It is not clear though whether it is still possible to strengthen the probability bounds above. The numerical calculations in Section 5 of [19] strongly suggest that the actual probability bounds should be very poor. The proof of this theorem is similar to the proof of Theorem 1.4 in [19], but since this result is optimal we present it here. It is given in Section 5. The proof is based on Lemma 3.1.

The results on rate of convergence and the rigidity of eigenvalues are based on Theorem 1.1 and Theorem 1.2. We first investigate the rate of convergence in probability of ESD to the distribution function of Wigner’s semicircle law defined as follows  $G_{sc}(x) := \int_{-\infty}^x g_{sc}(\lambda) d\lambda$ . To measure the distance between distribution functions, we introduce the uniform distance

$$\Delta_n^* := \sup_{x \in \mathbb{R}} |F_n(x) - G_{sc}(x)|.$$

At this point, we omit a detailed discussion of previous results and refer the reader instead to [19], which provides links to the related papers [3,22,25,26] and [36]. We prove the following theorem.

**Theorem 1.4.** *Assume that the conditions (C0) hold. Then there exist positive constants  $c$  and  $C$  depending on  $\delta$  such that for all  $1 \leq p \leq c \log n$*

$$\mathbb{P}(\Delta_n^* \geq K) \leq \frac{C^p \log^{2p} n}{K^p n^p}$$

for all  $K > 0$ .

Let  $N[x - \frac{\xi}{2n}, x + \frac{\xi}{2n}] := N_I(\mathbf{W})$  for  $I = [x - \frac{\xi}{2n}, x + \frac{\xi}{2n}]$ ,  $\xi > 0$ . The following result is the direct corollary of Theorem 1.4.

**Corollary 1.5.** *Assume that conditions (C0) hold. Then there exist positive constants  $c$  and  $C$  depending on  $\delta$  such that for all  $1 \leq p \leq c \log n$  and all  $\xi > 0, K > 0$*

$$\mathbb{P}\left(\left|\frac{N[x - \frac{\xi}{2n}, x + \frac{\xi}{2n}]}{\xi} - g_{sc}(x)\right| \geq \frac{K}{\xi}\right) \leq \frac{C^p \log^{2p} n}{K^p n^p}. \tag{1.9}$$

Taking  $K = C_1 n^{-1} \log^2 n$  with  $C_1 = C e^{2/c}$  (one may of course take a larger constant), we get that

$$\mathbb{P}\left(\Delta_n^* \geq \frac{C_1 \log^2 n}{n}\right) \leq \frac{1}{n^2}. \tag{1.10}$$

The proof of Theorem 1.4 will be given in Section 6. We believe that it is still possible to reduce the power of log in (1.9) from  $2p$  to  $p$  or even  $\frac{p}{2}$ , which would be optimal due to the result of Gustavsson [29] for the *Gaussian Unitary Ensembles* (GUE).

Instead of  $\Delta_n^*$  one may study the following distance of the *mean* spectral distribution to its limit

$$\Delta_n := \sup_{x \in \mathbb{R}} |\mathbb{E} F_n(x) - G_{sc}(x)|.$$

The first estimate of  $\Delta_n$  was obtained by Z. Bai [5], who showed the bound  $\Delta_n = O(n^{-\frac{1}{4}})$  assuming  $\mu_4 < \infty$ . Already in 1998, Girko [16] published an error bound order  $O(n^{-\frac{1}{2}})$  under the same moment conditions. In 2002, [17] he closed gaps in his proof. The same result was independently obtained by Bai, Miao and Tsay [6] and Götze and Tikhomirov [22], Girko [16] claimed that the actual rate of convergence of the expected spectral distribution function to the

semi-circle law is  $O(n^{-1/2})$  even in the Gaussian case. In 2002, Götze and Tikhomirov [21] showed that for the GUE the rate of convergence to the semi-circle law is  $O(n^{-2/3})$ . In 2005, [23] they improved this bound to  $O(n^{-1})$ . In 2007 in [37] the bound  $\Delta_n = O(n^{-1})$  was obtained for the *Gaussian Orthogonal Ensemble* (GOE) as well. Up to this point it was not clear what the optimal rate of convergence to the semi-circular law under weak moment conditions only should be. It follows from [7] that if the distributions of the matrix entries satisfy a Poincaré type inequality then  $\Delta_n = O(n^{-\frac{2}{3}})$ . Recently Götze and Tikhomirov [28] proved the bound  $\Delta_n = O(n^{-1})$  assuming that  $\mu_8 < \infty$  or  $\mu_4 < \infty$  combined with the assumption  $|X_{jk}| \leq Cn^{\frac{1}{4}}$  a.s. Finally in [27], the following theorem was proved.

**Theorem 1.6.** *Assume that the conditions (C0) hold. There exist a positive constant  $C(\delta)$  depending on  $\delta$  such that*

$$\Delta_n \leq \frac{C(\delta)}{n}.$$

In Section 6, we give an alternative proof of this theorem, based on the methods developed in this paper.

Another application of Theorem 1.1 is the following result which shows the rigidity of eigenvalues. Let us define the quantile position of the  $j$ th eigenvalue by

$$\gamma_j : \int_{-\infty}^{\gamma_j} g_{sc}(\lambda) d\lambda = \frac{j}{n}, \quad 1 \leq j \leq n.$$

We will prove the following theorem.

**Theorem 1.7.** *Assume that the conditions (C0) hold and let  $K > 0$  be an integer. Then:*

(i) *For all  $j \in [K, n - K + 1]$  there exist constants  $c$  and  $C, C_1$  depending on  $\delta$  such that for all  $1 \leq p \leq c \log n$  we have*

$$\mathbb{P}(|\lambda_j - \gamma_j| \geq C_1 K [\min(j, n - j + 1)]^{-\frac{1}{3}} n^{-\frac{2}{3}}) \leq \frac{C^p \log^{2p} n}{K^p}.$$

(ii) *Assume that  $\delta = 4$ . For any  $0 < \phi < 2$  and all  $j \leq K$  or  $j \geq n - K + 1$  there exist constants  $c$  and  $C, C_1$  depending on  $\phi$  and  $\mu_8$  such that for  $5 \leq p \leq c \log n$*

$$\mathbb{P}(|\lambda_j - \gamma_j| \geq C_1 K [\min(j, n - j + 1)]^{-\frac{1}{3}} n^{-\frac{2}{3}}) \leq \frac{C}{n^{2-\phi}} + \frac{C^p \log^{12p} n}{K^p}.$$

For comparisons, we refer the interested reader to relevant results of Gustavsson [29] for the Gaussian case, as well as to the results in the papers [10], Theorem 7.6, [11], Theorem 2.13, [25], Remark 1.2, [31], Theorem 3.6, and [8], Theorem 4, already mentioned above. The proof of Theorem 1.7 is similar to the proof of Theorem 1.3 in [19] up to some small improvements due to an improved bound in Theorem 1.4. For proving the part (ii), we use the result of Theorem 1.2, using ideas from [8], Lemma 8.1, and [10], Theorem 7.6. Note here that our techniques allow to treat the case  $0 < \delta < 4$  in (ii) as well, but with weaker probability bounds in (ii) of order  $n^{-c(\delta)}$ , where  $c(\delta) > 0$  depends on moment exponent  $\delta$ . We omit the details and the proof of Theorem 1.7.

### 1.4. Notations

Throughout the paper, we will use the following notations. We assume that all random variables are defined on common probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $\mathbb{E}$  be the mathematical expectation with respect to  $\mathbb{P}$ . We denote by  $\mathbb{1}[A]$  the indicator function of the set  $A$ .

We denote by  $\mathbb{R}$  and  $\mathbb{C}$  the set of all real and complex numbers. We also define  $\mathbb{C}^+ := \{z \in \mathbb{C} : \text{Im } z \geq 0\}$ . Let  $\mathbb{T} = [1, \dots, n]$  denotes the set of the first  $n$  positive integers. For any  $\mathbb{J} \subset \mathbb{T}$  introduce  $\mathbb{T}_{\mathbb{J}} := \mathbb{T} \setminus \mathbb{J}$ . To simplify all notations, we will write  $\mathbb{T}_j, \mathbb{T}_{\mathbb{J},j}$  instead of  $\mathbb{T}_{\{j\}}$  and  $\mathbb{T}_{\mathbb{J} \cup \{j\}}$  respectively.

For any matrix  $\mathbf{W}$  together with its resolvent  $\mathbf{R}$  and Stieltjes transform  $m_n$  we shall systematically use the corresponding notations  $\mathbf{W}^{(\mathbb{J})}, \mathbf{R}^{(\mathbb{J})}, m_n^{(\mathbb{J})}$ , respectively, for the sub-matrix of  $\mathbf{W}$  with entries  $X_{jk}, j, k \in \mathbb{T} \setminus \mathbb{J}$ . For simplicity, we write  $\mathbf{W}^{(j)}, \mathbf{W}^{(\mathbb{J},j)}$  instead of  $\mathbf{W}^{(\{j\})}, \mathbf{W}^{(\mathbb{J} \cup \{j\})}$ . The same is applies to  $\mathbf{R}, m_n$  etc.

By  $C$  and  $c$  we denote some positive constants. If we write that  $C$  depends on  $\delta$ , we mean that  $C = C(\delta, \mu_{4+\delta})$ .

For an arbitrary matrix  $\mathbf{A}$  taking values in  $\mathbb{C}^{n \times n}$  we define the operator norm by  $\|\mathbf{A}\| := \sup_{x \in \mathbb{R}^n: \|x\|=1} \|\mathbf{A}x\|_2$ , where  $\|x\|_2 := (\sum_{j=1}^n |x_j|^2)^{\frac{1}{2}}$ . We also define the Hilbert–Schmidt norm by  $\|\mathbf{A}\|_2 := \text{Tr}^{\frac{1}{2}} \mathbf{A} \mathbf{A}^* = (\sum_{j,k=1}^n |\mathbf{A}_{jk}|^2)^{\frac{1}{2}}$ .

## 2. Proof of the main result

The proofs of Theorem 1.1 and Theorem 1.2 repeat the arguments of [18], but for the readers convenience we provide the main steps here. We start with the recursive representation of the diagonal entries  $\mathbf{R}_{jj}(z) = [(\mathbf{W} - z\mathbf{I})^{-1}]_{jj}$  of the resolvent. As noted before, we shall systematically use for any matrix  $\mathbf{W}$  together with its resolvent  $\mathbf{R}$ , Stieltjes transform  $m_n$  and etc. the corresponding quantities  $\mathbf{W}^{(\mathbb{J})}, \mathbf{R}^{(\mathbb{J})}, m_n^{(\mathbb{J})}$  and etc. for the corresponding sub matrix with entries  $X_{jk}, j, k \in \mathbb{T} \setminus \mathbb{J}$ . We will often omit the argument  $z$  from  $\mathbf{R}(z)$  and write  $\mathbf{R}$  instead. We may express  $\mathbf{R}_{jj}$  in the following way

$$\mathbf{R}_{jj} = \frac{1}{-z + \frac{X_{jj}}{\sqrt{n}} - \frac{1}{n} \sum_{l,k \in \mathbb{T}_j} X_{jk} X_{jl} \mathbf{R}_{kl}^{(j)}}. \tag{2.1}$$

Let  $\varepsilon_j := \varepsilon_{1j} + \varepsilon_{2j} + \varepsilon_{3j} + \varepsilon_{4j}$ , where

$$\varepsilon_{1j} = \frac{1}{\sqrt{n}} X_{jj}, \quad \varepsilon_{2j} = -\frac{1}{n} \sum_{l \neq k \in \mathbb{T}_j} X_{jk} X_{jl} \mathbf{R}_{kl}^{(j)}, \tag{2.2}$$

$$\varepsilon_{3j} = -\frac{1}{n} \sum_{k \in \mathbb{T}_j} (X_{jk}^2 - 1) \mathbf{R}_{kk}^{(j)}, \quad \varepsilon_{4j} = \frac{1}{n} (\text{Tr } \mathbf{R} - \text{Tr } \mathbf{R}^{(j)}). \tag{2.3}$$

Using these notations, we may rewrite (2.1) as follows

$$\mathbf{R}_{jj} = -\frac{1}{z + m_n(z)} + \frac{1}{z + m_n(z)} \varepsilon_j \mathbf{R}_{jj}. \tag{2.4}$$

Summing last equations for  $j = 1, \dots, n$ , we obtain

$$1 + zm_n(z) + m_n^2(z) = T_n, \tag{2.5}$$

where

$$T_n := \frac{1}{n} \sum_{j=1}^n \varepsilon_j \mathbf{R}_{jj}, \tag{2.6}$$

It is well known that  $s(z)$  satisfies the following quadratic equation

$$1 + zs(z) + s^2(z) = 0. \tag{2.7}$$

From (2.5) and (2.7), we conclude that

$$\Lambda_n = \frac{T_n}{z + m_n(z) + s(z)} = \frac{T_n}{b_n(z)},$$

where

$$b_n(z) := b(z) + \Lambda_n, \quad b(z) := z + 2s(z). \tag{2.8}$$

From Lemma B.1 in [18] (see also [8], Proposition 2.2) it follows that for all  $v > 0$  and  $|u| \leq 2 + v$  (using the quantities (2.8))

$$|\Lambda_n| \leq C \min \left\{ \frac{|T_n|}{|b(z)|}, \sqrt{|T_n|} \right\}. \tag{2.9}$$

Moreover, let  $u_0$  be an arbitrary fixed positive number. Then for all  $v > 0$  and  $|u| \leq u_0$

$$|\operatorname{Im} \Lambda_n| \leq C \min \left\{ \frac{|T_n|}{|b(z)|}, \sqrt{|T_n|} \right\}. \tag{2.10}$$

This means that in order to bound  $\mathbb{E} |\Lambda_n|^p$  (or  $\mathbb{E} |\operatorname{Im} \Lambda|^p$ , respectively) it is enough to estimate  $\mathbb{E} |T_n|^p$ .

Let  $V$  be an arbitrary fixed positive real number and  $A_0$  is some large constant defined below. We introduce the following region in the complex plane:

$$\mathbb{D} := \{z = u + iv \in \mathbb{C} : |u| \leq u_0, V \geq v \geq v_0 := A_0 n^{-1} \log n\}. \tag{2.11}$$

The following theorem provides a general bound for  $\mathbb{E} |T_n|^p$  for all  $z \in \mathbb{D}$  in terms of diagonal resolvent entries. To formulate the result of the theorem, we need to introduce additional notations. Let

$$\mathcal{A}(q) := \max_{\mathbb{J} \leq 1} \max_{j \in \mathbb{T}_{\mathbb{J}}} \mathbb{E}^{\frac{1}{q}} \operatorname{Im}^q \mathbf{R}_{jj}^{(\mathbb{J})}, \tag{2.12}$$

where  $\mathbb{J}$  may be an empty set or one point set. We also denote

$$\mathcal{E}_p := \frac{p^p \mathcal{A}^p(\kappa p)}{(nv)^p} + \frac{p^{2p}}{(nv)^{2p}} + \frac{|b(z)|^{\frac{p}{2}} \mathcal{A}^{\frac{p}{2}}(\kappa p)}{(nv)^p},$$

where  $\kappa = \frac{16}{1-2\alpha}$ .

**Theorem 2.1.** *Assume that the conditions (C1) hold and  $u_0 > 2$  and  $V > 0$ . There exist positive constants  $A_0, A_1$  and  $C$  depending on  $\alpha, u_0$  and  $V$  such that for all  $z \in \mathbb{D}$  we have*

$$\mathbb{E} |T_n|^p \leq C^p \mathcal{E}_p,$$

where  $1 \leq p \leq A_1 \log n$ .

**Proof.** See [18], Theorem 2.1. □

**Proof of Theorem 1.1.** By standard truncation arguments (see [18], Lemmas D.1–D.3) in what follows we may assume that

$$|X_{jk}| \leq Dn^\alpha \quad \text{for all } 1 \leq j, k \leq n$$

and some  $D := D(\alpha) > 0$ .

Applying Theorem 2.1, we will show in Section 3, Lemma 3.1, that there exist constants  $H_0$  depending on  $u_0, V$  and  $A_0, A_1$  depending on  $\alpha$  and  $H_0$  such that

$$\mathcal{A}^p(\kappa p) \leq H_0^p \operatorname{Im}^p s(z) + \frac{H_0^p p^p}{(nv)^p}$$

for all  $1 \leq p \leq A_1 \log n$  and  $z \in \mathbb{D}$ . This inequality and Theorem 2.1 together imply that

$$\mathbb{E} |T_n|^p \leq \frac{C^p p^p \operatorname{Im}^p s(z)}{(nv)^p} + \frac{C^p p^{2p}}{(nv)^{2p}} + \frac{C^p |b(z)|^{\frac{p}{2}} \operatorname{Im}^{\frac{p}{2}} s(z)}{(nv)^p} + \frac{C^p |b(z)|^{\frac{p}{2}} p^{\frac{p}{2}}}{(nv)^{\frac{3p}{2}}} \quad (2.13)$$

with some new constant  $C$  which depends on  $H_0$ . To estimate  $\mathbb{E} |\operatorname{Im} \Lambda_n|^p$ , we may choose one of the bounds (2.10), depending on  $z$  being near the edge of the limiting spectral interval or not. If  $|b(z)|^p \geq \frac{C^p p^p}{(nv)^p}$ , then we may use the bound

$$\mathbb{E} |\operatorname{Im} \Lambda_n|^p \leq \frac{C^p \mathbb{E} |T_n|^p}{|b(z)|^p}.$$

The r.h.s. of the last inequality may be estimated applying (2.13). We get

$$\mathbb{E} |\operatorname{Im} \Lambda_n|^p \leq \frac{C^p p^p \operatorname{Im}^p s(z)}{(nv)^p |b(z)|^p} + \frac{C^p p^{2p}}{(nv)^{2p} |b(z)|^p} + \frac{C^p \operatorname{Im}^{\frac{p}{2}} s(z)}{(nv)^p |b(z)|^{\frac{p}{2}}} + \frac{C^p p^{\frac{p}{2}}}{(nv)^{\frac{3p}{2}} |b(z)|^{\frac{p}{2}}}.$$

Since  $|b(z)|^p \geq \frac{C^p p^p}{(nv)^p}$  the last inequality may be rewritten in the following way

$$\mathbb{E} |\operatorname{Im} \Lambda_n|^p \leq \frac{C^p p^p \operatorname{Im}^p s(z)}{(nv)^p |b(z)|^p} + \frac{C^p \operatorname{Im}^{\frac{p}{2}} s(z)}{(nv)^p |b(z)|^{\frac{p}{2}}} + \frac{C^p p^p}{(nv)^p}.$$

It remains to estimate the imaginary part of  $s(z)$ . Since

$$\operatorname{Im}^p s(z) \leq c^p |b(z)|^p \quad \text{for } |u| \leq 2 \quad \text{and} \quad \operatorname{Im}^p s(z) \leq \frac{c^p v^p}{|b(z)|^p} \quad \text{otherwise}$$

both inequalities combined yield

$$\mathbb{E} |\operatorname{Im} \Lambda_n|^p \leq \left( \frac{Cp}{nv} \right)^p,$$

where we have used as well the fact that  $c\sqrt{\gamma+v} \leq |b(z)| \leq C\sqrt{\gamma+v}$  for all  $|u| \leq u_0$ ,  $0 < v \leq v_1$ . We assume now that  $|b(z)|^p \leq \frac{C^p p^p}{(nv)^p}$ . Then  $\operatorname{Im}^p s(z) \leq \frac{C^p p^p}{(nv)^p}$  and we obtain a bound proportional to  $|T_n|^{\frac{1}{2}}$ . Hence,

$$\mathbb{E} |\operatorname{Im} \Lambda_n|^p \leq C^p \mathbb{E} |T_n|^{\frac{p}{2}} \leq \left( \frac{Cp}{nv} \right)^p.$$

Similar arguments apply to  $\mathbb{E} |\Lambda_n|^p$ . □

**Proof of Theorem 1.2.** From Theorem 2.1, we may conclude that

$$\mathbb{E} |T_n|^p \leq \frac{C^p p^p \operatorname{Im}^p s(z)}{(nv)^p} + \frac{C^p p^{2p}}{(nv)^{3p}} + \frac{C^p |b(z)|^{\frac{p}{2}} \operatorname{Im}^{\frac{p}{2}} s(z)}{(nv)^p} + \frac{C^p |b(z)|^{\frac{p}{2}} p^{\frac{p}{2}}}{(nv)^{\frac{3p}{2}}}.$$

Applying (2.10), we get

$$\mathbb{E} |\operatorname{Im} \Lambda_n|^p \leq \frac{C^p \mathbb{E} |T_n|^p}{|b(z)|^p}.$$

This inequality together with (2.13) leads to

$$\mathbb{E} |\operatorname{Im} \Lambda_n|^p \leq \frac{C^p p^p \operatorname{Im}^p s(z)}{(nv)^p |b(z)|^p} + \frac{C^p p^{2p}}{(nv)^{2p} |b(z)|^p} + \frac{C^p \operatorname{Im}^{\frac{p}{2}} s(z)}{(nv)^p |b(z)|^{\frac{p}{2}}} + \frac{C^p p^{\frac{p}{2}}}{(nv)^{\frac{3p}{2}} |b(z)|^{\frac{p}{2}}}.$$

Since  $c\sqrt{\gamma+v} \leq |b(z)| \leq C\sqrt{\gamma+v}$  for all  $|u| \leq u_0$ ,  $0 < v \leq v_1$  and

$$\frac{cv}{\sqrt{\gamma+v}} \leq \operatorname{Im} s(z) \leq \frac{cv}{\sqrt{\gamma+v}} \quad \text{for all } 2 \leq |u| \leq u_0, 0 < v \leq v_1,$$

we finally get

$$\mathbb{E} |\operatorname{Im} \Lambda_n|^p \leq \frac{C^p p^p}{n^p (\gamma+v)^p} + \frac{C^p p^{2p}}{(nv)^{2p} (\gamma+v)^{\frac{p}{2}}} + \frac{C^p}{n^p v^{\frac{p}{2}} (\gamma+v)^{\frac{p}{2}}} + \frac{C^p p^{\frac{p}{2}}}{(nv)^{\frac{3p}{2}} (\gamma+v)^{\frac{p}{4}}}.$$

This bound concludes the proof of the theorem. □

### 3. A moment bound for diagonal entries of the resolvent

In this section, we prove bounds for the diagonal entries of the resolvent. As mentioned before in the Introduction and Section 1.2, these bounds will play a crucial role in the proof of Theorem 1.1 and Theorem 1.2. To formulate the main result, we shall introduce additional notations. We denote

$$\Psi(z) := \operatorname{Im} s(z) + \frac{p}{nv}, \tag{3.1}$$

and recall definition of the region  $\mathbb{D}$ ,

$$\mathbb{D} := \{z = u + iv \in \mathbb{C} : |u| \leq u_0, V \geq v \geq v_0 := A_0 n^{-1} \log n\},$$

where  $u_0, V > 0$  are any fixed real numbers and  $A_0$  is some large constant determined below. The main result of this section is the following lemma.

**Lemma 3.1.** *Assuming the conditions (C1) there exist positive constants  $C_0, H_0$  depending on  $u_0, V$  and positive constants  $A_0, A_1$  depending on  $C_0, H_0, \delta$  and  $K$  (see the next section for definition of  $K$ ) such that for all  $z \in \mathbb{D}$  and  $1 \leq p \leq A_1 \log n$  we have*

$$\max_{j \in \mathbb{T}} \mathbb{E} |\mathbf{R}_{jj}(z)|^p \leq C_0^p, \tag{3.2}$$

$$\mathbb{E} \frac{1}{|z + m_n(z)|^p} \leq C_0^p, \tag{3.3}$$

$$\max_{j \in \mathbb{T}} \mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}(z) \leq H_0^p \Psi^p(z). \tag{3.4}$$

Note that the region  $\mathbb{D}$  and  $p$  depend on  $\alpha$  via some constants  $A_0$  and  $A_1$  only, but the power of the logarithmic factor in the definition of  $\mathbb{D}$  is independent of  $\alpha$ . We split the proof of this lemma into two subsections. In the first subsection we prove (3.2) and (3.3). In the second subsection, we prove (3.4). The proofs are very similar and we shall give the proof of (3.2) in full detail while only sketching the proofs of (3.3) and (3.4). Note that in the subsequent sections we keep the notations for the constants  $A_0, A_1, C_0$  and  $H_0$  introduced above, but the particular values may depend on the location.

#### 3.1. Moment bounds for the absolute value of resolvent entries

We start this section with the following lemma which was proved in [18], Lemma 4.1.

**Lemma 3.2.** *Assuming the conditions (C1) there exist a positive constant  $C_0$  depending on  $u_0, V$  and positive constants  $A_0, A_1$  depending on  $C_0, \delta$  such that for all  $z \in \mathbb{D}$  and  $1 \leq p \leq A_1 (nv)^\kappa$  we have*

$$\max_{j, k \in \mathbb{T}} \mathbb{E} |\mathbf{R}_{jk}(z)|^p \leq C_0^p \tag{3.5}$$

and

$$\mathbb{E} \frac{1}{|z + m_n(z)|^p} \leq C_0^p.$$

**Proof.** See [18], Lemma 4.1. □

**Proof of (3.2) and (3.3).** Since  $u$  is fixed and  $|u| \leq u_0$  we shall omit  $u$  from the notation of the resolvent and denote  $\mathbf{R}(v) := \mathbf{R}(z)$ . Sometimes in order to simplify notations we shall also omit the argument  $v$  in  $\mathbf{R}(v)$  and just write  $\mathbf{R}$ . The same applies to the Stieltjes transform  $m_n(z)$ .

Let  $\tilde{v} = n^{-1} \log^{\frac{1}{\kappa}} n$ , then  $1 \leq p \leq A_1(n\tilde{v})^\kappa = A_1 \log n$ . Without loss of generality, we may consider  $p = A_1 \log n$  only (otherwise one may apply Lyapunov’s inequality for moments). It follows from (3.5) that

$$\max_{j \in \mathbb{T}} \mathbb{E} |\mathbf{R}_{jj}(v)|^p \leq C_0^p$$

for all  $V \geq v \geq \tilde{v}$ . To prove (3.2), we need to descent from  $\tilde{v}$  to  $v_0$  while keeping  $p = A_1 \log n$ . Applying Lemma 3.3 below with  $s_0 := \log^{\frac{1}{\kappa}-1} n$  we may show that for all  $v \geq v_0$

$$\max_{j \in \mathbb{T}} \mathbb{E} |\mathbf{R}_{jj}(v)|^p \leq C_0^p \log^{(\frac{1}{\kappa}-1)p} n.$$

It remains to remove the log factor on the right hand side of the previous inequality. To this aim, we shall adopt the moment matching technique which has been successfully used recently by Lee and in Yin in [31] (see Lemma 5.2 and Lemma 5.3). We denote by  $Y_{jk}, 1 \leq j \leq k \leq n$  a triangular set of random variables such that  $|Y_{jk}| \leq D$ , for some  $D$  chosen later, and

$$\mathbb{E} X_{jk}^s = \mathbb{E} Y_{jk}^s \quad \text{for } s = 1, \dots, 4.$$

It follows from Lemma 3.4 below that such a set of random variables exists. Let us denote  $\mathbf{W}^y := \frac{1}{\sqrt{n}} \mathbf{Y}, \mathbf{R}^y := (\mathbf{W}^y - z\mathbf{I})^{-1}$  and  $m_n^y(z) := \frac{1}{n} \text{Tr} \mathbf{R}^y(z)$ . Then, in Lemma 3.5 we show that for all  $v \geq v_0$  and  $5 \leq p \leq A_1 \log n$  there exist positive constants  $C_1, C_2$  such that

$$\mathbb{E} |\mathbf{R}_{jj}(v)|^p \leq C_1^p + C_2 \mathbb{E} |\mathbf{R}_{jj}^y(v)|^p.$$

It is easy to see that  $Y_{jk}$  are sub-Gaussian random variables. It follows from Lemma 4.1 of the following section that

$$\mathbb{E} |\mathbf{R}_{jj}^y(v)|^p \leq C_0^p.$$

This fact concludes the proof of (3.2). Similar arguments yield the estimate (3.3). □

**Lemma 3.3.** *Let  $C_0$  be some constant. Assume that for all  $v' \geq \tilde{v}$  and all  $1 \leq p \leq A_1(nv')^\kappa$  we have*

$$\max_{j,k \in \mathbb{T}} \mathbb{E} |\mathbf{R}_{jk}(v')|^p \leq C_0^p \quad \text{and} \quad \mathbb{E} \frac{1}{|z + m_n(v')|^p} \leq C_0^p. \tag{3.6}$$

Then for all  $s_0 \geq 1$ ,  $v \geq \tilde{v}/s_0$  we have

$$\max_{j,k \in \mathbb{T}} \mathbb{E} |\mathbf{R}_{jk}(v)|^p \leq 2^p (1 + s_0)^p C_0^p \quad \text{and} \quad \mathbb{E} \frac{1}{|z + m_n(v)|^p} \leq s_0^p C_0^p.$$

**Proof.** We first consider the diagonal entries. Let us fix  $s_0 \geq 1$  and  $v \geq \tilde{v}/s_0$  and denote  $v' = s_0 v \geq \tilde{v}$ . Applying Lemma A.1 and the main assumption (3.6), we get

$$|\mathbf{R}_{jj}(v)|^p \leq s_0^p |\mathbf{R}_{jj}(s_0 v)| \leq (C_0 s_0)^p. \tag{3.7}$$

Applying the resolvent equality

$$|\mathbf{R}_{jk}(v) - \mathbf{R}_{jk}(s_0 v)| \leq v(s_0 - 1) |[\mathbf{R}(v)\mathbf{R}(s_0 v)]_{jk}|.$$

The Cauchy–Schwarz inequality and Lemma A.3 together imply that

$$|\mathbf{R}_{jk}(v) - \mathbf{R}_{jk}(s_0 v)| \leq \sqrt{s_0} \sqrt{|\mathbf{R}_{jj}(v)| |\mathbf{R}_{kk}(s_0 v)|}.$$

It follows that

$$\mathbb{E} |\mathbf{R}_{jk}(v)|^p \leq 2^p \mathbb{E} |\mathbf{R}_{jk}(s_0 v)|^p + 2^p s_0^{\frac{p}{2}} \mathbb{E}^{\frac{1}{2}} |\mathbf{R}_{jj}(v)|^p \mathbb{E}^{\frac{1}{2}} |\mathbf{R}_{kk}(s_0 v)|^p.$$

Applying (3.7), we get

$$\mathbb{E} |\mathbf{R}_{jk}(v)|^p \leq 2^p C_0^p + 2^p s_0^p C_0^p \leq 2^p (1 + s_0)^p C_0^p.$$

Similarly, applying Lemma A.1, (Inequality (A.1)), we arrive at a bound for  $\mathbb{E} \frac{1}{|z + m_n(v)|^p}$ . We omit the details. □

**Lemma 3.4.** For any  $A, B \in \mathbb{R}$  with  $B \geq A^2 + 1$ , there exists a random variable  $X$  such that

$$\mathbb{E} X = 0, \quad \mathbb{E} X^2 = 1, \quad \mathbb{E} X^3 = A, \quad \mathbb{E} X^4 = B$$

and

$$\text{supp}(X) \in [-DB, DB],$$

for some  $D$  independent of  $B$ .

**Proof.** See [31], Lemma 5.2. □

**Lemma 3.5.** For all  $v \geq v_0$  and  $5 \leq p \leq \log n$  there exist positive constants  $C_1, C_2$  such that

$$\mathbb{E} |\mathbf{R}_{jk}(v)|^p \leq C_1^p + C_2 \mathbb{E} |\tilde{\mathbf{R}}_{jk}(v)|^p \tag{3.8}$$

and

$$\mathbb{E} \frac{1}{|z + m_n(v)|^p} \leq C_1^p + C_2 \mathbb{E} \frac{1}{|z + \tilde{m}_n(v)|^p}. \tag{3.9}$$

**Proof.** We first prove (3.8). The method is based on the following replacement scheme, which has been used in recent results [9] and [31]. We replace the  $\frac{n(n+1)}{2}$  matrix entries  $X_{ab}$  by  $Y_{ab}$ ,  $1 \leq a \leq b \leq n$ , thus replacing the corresponding resolvent entries  $\mathbf{R}_{jk}$  by  $\mathbf{R}_{jk}^y$  for every pair of  $(j, k)$ . Let  $\mathbb{J}, \mathbb{K} \subset \mathbb{T}$ . Denote by  $\mathbf{W}^{(\mathbb{J}, \mathbb{K})}$  the random matrix  $\mathbf{W}$  with all entries in the positions  $(\mu, \nu)$ ,  $\mu \in \mathbb{J}, \nu \in \mathbb{K}$  replaced by  $\frac{1}{\sqrt{n}}Y_{\mu\nu}$ . Assume that we have already exchanged all entries in positions  $(\mu, \nu)$ ,  $\mu \in \mathbb{J}, \nu \in \mathbb{K}$  and are going to replace an additional entry in the position  $(a, b)$ ,  $a \in \mathbb{T} \setminus \mathbb{J}, b \in \mathbb{T} \setminus \mathbb{K}$ . Without loss of generality, we may assume that  $\mathbb{J} = \emptyset, \mathbb{K} = \emptyset$  (hence  $\mathbf{W}^{(\mathbb{J}, \mathbb{K})} = \mathbf{W}$ ) and then denote  $\mathbf{V} := \mathbf{W}^{(a, \{b\})}$ . The following additional notations will be needed

$$\mathbf{E}^{(a,b)} = \begin{cases} \mathbf{e}_a \mathbf{e}_b^\top + \mathbf{e}_b \mathbf{e}_a^\top, & 1 \leq a < b \leq n, \\ \mathbf{e}_a \mathbf{e}_a^\top, & a = b \end{cases}$$

and  $\mathbf{U} := \mathbf{W} - \frac{X_{ab}}{\sqrt{n}}\mathbf{E}^{(a,b)}$ , where  $\mathbf{e}_j$  denotes a unit column-vector with all zeros except  $j$ th position. In these notations, we may write

$$\mathbf{W} = \mathbf{U} + \frac{1}{\sqrt{n}}X_{ab}\mathbf{E}^{(a,b)}, \quad \mathbf{V} = \mathbf{U} + \frac{1}{\sqrt{n}}Y_{ab}\mathbf{E}^{(a,b)}.$$

Recall that  $\mathbf{R} := (\mathbf{W} - z\mathbf{I})^{-1}$  and denote  $\mathbf{S} := (\mathbf{V} - z\mathbf{I})^{-1}$  and  $\mathbf{T} := (\mathbf{U} - z\mathbf{I})^{-1}$ . Let us assume that we have already proved the following fact

$$\mathbb{E} |\mathbf{R}_{jk}|^p = \mathcal{I}(p) + \frac{\theta_1 C^p}{n^2} + \frac{\theta_1 \mathbb{E} |\mathbf{R}_{jk}|^p}{n^2}, \tag{3.10}$$

where  $\mathcal{I}(p)$  is some quantity depending on  $p, n$  (see (3.24) below for precise definition) and  $|\theta_1| \leq 1, C > 0$  are some numbers. Similarly,

$$\mathbb{E} |\mathbf{S}_{jk}|^p = \mathcal{I}(p) + \frac{\theta_2 C^p}{n^2} + \frac{\theta_2 \mathbb{E} |\mathbf{S}_{jk}|^p}{n^2}, \tag{3.11}$$

where  $|\theta_2| \leq 1$ . It follows from (3.10) and (3.11) that

$$\left(1 - \frac{\theta_1}{n^2}\right) \mathbb{E} |\mathbf{R}_{jk}|^p \leq \left(1 - \frac{\theta_2}{n^2}\right) \mathbb{E} |\mathbf{S}_{jk}|^p + \frac{2C^p}{n^2}. \tag{3.12}$$

Let us denote  $\rho := (1 - \frac{\theta_2}{n^2})(1 - \frac{\theta_1}{n^2})^{-1}$ . We get

$$\mathbb{E} |\mathbf{R}_{jk}|^p \leq \rho \mathbb{E} |\mathbf{S}_{jk}|^p + \frac{C_1^p}{n^2}, \tag{3.13}$$

with some positive constant  $C_1$ . Repeating (3.13) recursively  $n(n+1)/2$  times, we arrive at the following bound

$$\mathbb{E} |\mathbf{R}_{jk}|^p \leq \rho^{\frac{n(n+1)}{2}} \mathbb{E} |\mathbf{R}_{jk}^y|^p + \frac{C_1^p}{n^2} (1 + \rho_1 + \dots + \rho_1^{\frac{n(n+1)}{2}-1}). \tag{3.14}$$

It is easy to see from the definition of  $\rho$  that for some  $\theta$ , say  $|\theta| < 4$ , we have

$$\rho \leq 1 + \frac{|\theta|}{n^2}.$$

From this inequality and (3.14), we deduce that

$$\mathbb{E} |\mathbf{R}_{jk}|^p \leq C_2 \mathbb{E} |\mathbf{R}_{jk}^y|^p + C_3^p, \tag{3.15}$$

with some positive constants  $C_2$  and  $C_3$ . From the last inequality, we may conclude the statement of the lemma. It remains to prove (3.10) (resp. (3.11)). Applying the resolvent equation, we get for  $m \geq 0$

$$\mathbf{R} = \mathbf{T} + \sum_{\mu=1}^m \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} X_{ab}^\mu (\mathbf{TE}^{(a,b)})^\mu \mathbf{T} + \frac{(-1)^{m+1}}{n^{\frac{m+1}{2}}} X_{ab}^{m+1} (\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}. \tag{3.16}$$

The same identity holds for  $\mathbf{S}$

$$\mathbf{S} = \mathbf{T} + \sum_{\mu=1}^m \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} Y_{ab}^\mu (\mathbf{TE}^{(a,b)})^\mu \mathbf{T} + \frac{(-1)^{m+1}}{n^{\frac{m+1}{2}}} Y_{ab}^{m+1} (\mathbf{TE}^{(a,b)})^{m+1} \mathbf{S}. \tag{3.17}$$

We investigate (3.16). In order handle arbitrary high moments of  $\mathbf{R}_{jk}$  we apply a Stein type technique which we have used in previous papers [18] and [26]. Let us introduce the following function  $\varphi(z) := \bar{z}|z|^{p-2}$  and write

$$\mathbb{E} |\mathbf{R}_{jk}|^p = \mathbb{E} \mathbf{R}_{jk} \varphi(\mathbf{R}_{jk}).$$

Applying (3.16), we get

$$\begin{aligned} \mathbb{E} |\mathbf{R}_{jk}|^p &= \sum_{\mu=0}^4 \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E} X_{ab}^\mu [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} \varphi(\mathbf{R}_{jk}) \\ &\quad + \sum_{\mu=5}^m \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E} X_{ab}^\mu [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} \varphi(\mathbf{R}_{jk}) \\ &\quad + \frac{1}{n^{\frac{m+1}{2}}} \mathbb{E} X_{ab}^{m+1} [(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jk} \varphi(\mathbf{R}_{jk}) \\ &=: \mathcal{A}_0 + \mathcal{A}_1 + \mathcal{A}_2. \end{aligned} \tag{3.18}$$

The bound for  $\mathcal{A}_2$  is easy. It is straightforward to check that  $[(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jk}$  is the sum of  $2^{m+1}$  terms of the following type

$$\mathbf{T}_{ji_1} \mathbf{T}_{i_1 i_2} \cdots \mathbf{T}_{i_m i_m} \mathbf{R}_{i_m k},$$

where  $i_l = a$  or  $i_l = b$  for  $l = 1, \dots, m$ . We may estimate  $|\mathbf{R}_{ak}|$  or  $|\mathbf{R}_{bk}|$  by  $n$ , since  $v > n^{-1}$ . Taking absolute values, applying Hölder's inequality and the bound  $\mathbb{E}^{\frac{1}{c^p}} |\mathbf{T}_{ii+1}|^{cp} < C_0 \log^{\frac{1}{\kappa}-1} n$  for some  $c > 0$  (see Lemma 3.3 with  $s_0 := \log^{\frac{1}{\kappa}-1} n$ ) we get

$$\begin{aligned} |\mathcal{A}_2| &\leq \frac{1}{n^{\frac{m+1}{2}}} \mathbb{E} |X_{ab}|^{m+1} [(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jk} |\mathbf{R}_{jk}|^{p-1} \\ &\leq \frac{n \log^{c(\kappa)} n}{n^{\frac{m+1}{2}}} \mathbb{E}^{\frac{1}{p}} |X_{ab}|^{(m+1)p} \mathbb{E}^{\frac{p-1}{p}} |\mathbf{R}_{jk}|^p, \end{aligned}$$

where  $c(\kappa) > 0$  – some constant depending on  $\kappa$ . Using **(C1)** we may assume without loss of generality that  $|X_{jk}| \leq Dn^{\frac{1}{2}-\phi}$  for some  $\phi > 0$  depending on  $\delta$ . Indeed, one may choose  $\phi := \frac{\delta}{2(4+\delta)}$ . Applying this fact and the last inequality, we may write

$$|\mathcal{A}_2| \leq \frac{Cn \log^{c(\kappa)} n}{n^{\phi(m+1)}} \mathbb{E}^{\frac{p-1}{p}} |\mathbf{R}_{jk}|^p.$$

We now choose  $m$  such that  $\phi(m+1) = 4$ . Applying  $2 = \frac{2}{p} + \frac{2(p-1)}{p}$  and Young's inequality, we obtain

$$|\mathcal{A}_2| \leq \frac{C}{n^{\frac{2}{p} + \frac{2(p-1)}{p}}} \mathbb{E}^{\frac{p-1}{p}} |\mathbf{R}_{jk}|^p \leq \frac{C^p}{pn^2} + \frac{\mathbb{E} |\mathbf{R}_{jk}|^p}{n^2}. \quad (3.19)$$

Let us consider the term  $\mathcal{A}_1 = \mathcal{A}_{1,5} + \dots + \mathcal{A}_{1,m}$ , where

$$\mathcal{A}_{1,\mu} := \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E} X_{ab}^\mu [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} \varphi(\mathbf{R}_{jk}), \quad \mu = 5, \dots, m.$$

We fix some  $\mu \geq 5$  and bound  $\mathcal{A}_{1,\mu}$ . Let us introduce the following quantity

$$\tilde{\mathbf{T}}_{jk} := \mathbb{E}(\mathbf{R}_{jk} | \mathfrak{M}^{(a,b)}), \quad (3.20)$$

where  $\mathfrak{M}^{(a,b)} := \sigma\{X_{st}, (s,t) \neq (a,b)\}$  –  $\sigma$ -algebra generated by  $X_{st}$ ,  $1 \leq s \leq t \leq n$ ,  $(s,t) \neq (a,b)$ . We may split  $\mathcal{A}_{1,\mu}$  into a sum of two terms  $\mathcal{A}_{1,\mu} = \mathcal{B}_1 + \mathcal{B}_2$ , where

$$\begin{aligned} \mathcal{B}_1 &:= \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E} X_{ab}^\mu [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} \varphi(\tilde{\mathbf{T}}_{jk}), \\ \mathcal{B}_2 &:= \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E} X_{ab}^\mu [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} [\varphi(\mathbf{R}_{jk}) - \varphi(\tilde{\mathbf{T}}_{jk})]. \end{aligned}$$

For the first term, we may write

$$\begin{aligned} |\mathcal{B}_1| &\leq \frac{1}{n^{\frac{\mu}{2}}} \mathbb{E} |X_{ab}|^\mu \mathbb{E} |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk}| |\tilde{\mathbf{T}}_{jk}|^{p-1} \leq \frac{C \log^{c(\kappa)} n}{n^{\frac{\mu}{2}}} \mathbb{E} |X_{ab}|^\mu \mathbb{E}^{\frac{p-1}{p}} |\mathbf{R}_{jk}|^p \\ &\leq \frac{C \log^{c(\kappa)} n}{n^{2+\phi(\mu-4)}} \mathbb{E}^{\frac{p-1}{p}} |\mathbf{R}_{jk}|^p \leq \frac{C}{n^{\frac{2}{p} + \frac{2(p-1)}{p}}} \mathbb{E}^{\frac{p-1}{p}} |\mathbf{R}_{jk}|^p \leq \frac{C^p}{n^2} + \frac{\mathbb{E} |\mathbf{R}_{jk}|^p}{n^2}. \end{aligned}$$

Let us introduce the following function

$$f(t) := \varphi(\tilde{\mathbf{T}}_{jk} + t(\mathbf{R}_{jk} - \tilde{\mathbf{T}}_{jk})), \quad t \in [0, 1]. \tag{3.21}$$

It is easy to see that

$$f(1) = \varphi(\mathbf{R}_{jk}), \quad f(0) = \varphi(\tilde{\mathbf{T}}_{jk}).$$

Moreover, by Taylor’s formula, for all  $l \geq 1$ ,

$$f(t) = \sum_{k=0}^{l-1} \frac{f^{(k)}(0)}{k!} t^k + \frac{t^l}{(l-1)!} \mathbb{E}_\theta (1-\theta)^{l-1} f^{(l)}(\theta t), \tag{3.22}$$

where  $\theta$  is uniformly distributed on  $[0, 1]$  and independent of all other r.v. It is straightforward to check that there exists some positive constant  $K_l$  depending on  $l$  such that

$$|f^{(l)}(t)| \leq K_l p^l |t|^l |\tilde{\mathbf{T}}_{jk} + t(\mathbf{R}_{jk} - \tilde{\mathbf{T}}_{jk})|^{p-l-1}.$$

Applying (3.22) with  $t = 1, l = 1$  we estimate  $\mathcal{B}_2$  as follows

$$|\mathcal{B}_2| \leq \frac{p}{n^{\frac{\mu}{2}}} \mathbb{E} |X_{ab}|^\mu |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} | |\mathbf{R}_{jk} - \tilde{\mathbf{T}}_{jk}| |\tilde{\mathbf{T}}_{jk} + \theta(\mathbf{R}_{jk} - \tilde{\mathbf{T}}_{jk})|^{p-2}.$$

Using triangular inequalities, we may bound (up to some constant) the r.h.s. of the previous inequality by the sum  $|\mathcal{B}_2| \leq \mathcal{I}_1 + \mathcal{I}_2$ , where

$$\begin{aligned} \mathcal{I}_1 &= \frac{p}{n^{\frac{\mu}{2}}} \mathbb{E} |X_{ab}|^\mu |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} | |\mathbf{R}_{jk} - \tilde{\mathbf{T}}_{jk}| |\tilde{\mathbf{T}}_{jk}|^{p-2}, \\ \mathcal{I}_2 &= \frac{p^p}{n^{\frac{\mu}{2}}} \mathbb{E} |X_{ab}|^\mu |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} | |\mathbf{R}_{jk} - \tilde{\mathbf{T}}_{jk}|^{p-1}. \end{aligned}$$

Applying again (3.16), we obtain  $\mathcal{I}_1 \leq \mathcal{I}_{11} + \dots + \mathcal{I}_{14}$ , where we denoted

$$\begin{aligned} \mathcal{I}_{11} &:= p \sum_{l=1}^m \frac{1}{n^{\frac{\mu+l}{2}}} \mathbb{E} |X_{ab}|^{\mu+l} |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} [(\mathbf{TE}^{(a,b)})^l \mathbf{T}]_{jk} | |\tilde{\mathbf{T}}_{jk}|^{p-2}, \\ \mathcal{I}_{12} &:= p \sum_{l=1}^m \frac{1}{n^{\frac{\mu+l}{2}}} \mathbb{E} |X_{ab}|^l \mathbb{E} |X_{ab}|^\mu \mathbb{E} |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} [(\mathbf{TE}^{(a,b)})^l \mathbf{T}]_{jk} | |\tilde{\mathbf{T}}_{jk}|^{p-2}, \\ \mathcal{I}_{13} &:= \frac{p}{n^{\frac{\mu+m+1}{2}}} \mathbb{E} |X_{ab}|^{\mu+m+1} |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} [(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jk} | |\tilde{\mathbf{T}}_{jk}|^{p-2}, \\ \mathcal{I}_{14} &:= \frac{p}{n^{\frac{\mu+m+1}{2}}} \mathbb{E} |X_{ab}|^\mu |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} | \mathbb{E} (|X_{ab}|^{m+1} |[(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jk} | |\mathfrak{M}^{(a,b)}|) |\tilde{\mathbf{T}}_{jk}|^{p-2}. \end{aligned}$$

It is straightforward to check that

$$\mathcal{I}_{11} \leq \frac{Cm}{n^{\frac{4}{p} + \frac{2(p-2)}{p}}} \mathbb{E}^{\frac{p-2}{p}} |\mathbf{R}_{jk}|^p \leq \frac{C^p}{n^2} + \frac{\mathbb{E} |\mathbf{R}_{jk}|^p}{n^2}.$$

The term  $\mathcal{I}_{12}$  may be estimated in a similar way. The terms  $\mathcal{I}_{13}$  and  $\mathcal{I}_{14}$  can be estimated as before choosing  $m$  sufficiently large. Let us consider the term  $\mathcal{I}_2$ . Similarly, it may be rewritten as follows

$$\mathcal{I}_{21} := \sum_{l=1}^m \frac{p^p}{n^{\frac{\mu+l(p-1)}{2}}} \mathbb{E} |X_{ab}|^{\mu+l(p-1)} \mathbb{E} |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} [(\mathbf{TE}^{(a,b)})^l \mathbf{T}]_{jk}|^{p-1},$$

$$\mathcal{I}_{22} := \sum_{l=1}^m \frac{p^p}{n^{\frac{\mu+l(p-1)}{2}}} \mathbb{E}^{p-1} |X_{ab}|^l \mathbb{E} |X_{ab}|^\mu \mathbb{E} |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} [(\mathbf{TE}^{(a,b)})^l \mathbf{T}]_{jk}|^{p-1},$$

$$\mathcal{I}_{23} := \frac{p^p}{n^{\frac{\mu+(m+l)(p-1)}{2}}} \mathbb{E} |X_{ab}|^{\mu+(m+l)(p-1)} |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} [(\mathbf{TE}^{(a,b)})^{m+l} \mathbf{R}]_{jk}|^{p-1},$$

$$\mathcal{I}_{24} := \frac{p^p}{n^{\frac{\mu+(m+l)(p-1)}{2}}} \mathbb{E} |X_{ab}|^\mu |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk}| [\mathbb{E} (|X_{ab}|^{m+1} |[(\mathbf{TE}^{(a,b)})^{m+l} \mathbf{R}]_{jk}| |\mathfrak{M}^{(a,b)}|)]^{p-1}.$$

We investigate  $\mathcal{I}_{21}$ . We obtain

$$\mathcal{I}_{21} \leq \sum_{l=1}^m \frac{C^{cp} p^p \log^{c(\kappa)p} n}{n^{2+\phi l(p-1)}} \leq \frac{C^p}{n^2}.$$

The same estimate is valid for  $\mathcal{I}_{22}$ . The terms  $\mathcal{I}_{23}$ ,  $\mathcal{I}_{24}$  may be estimated as before choosing  $m$  sufficiently large. Finally, we conclude that

$$|\mathcal{A}_1| \leq \frac{C^p}{n^2} + \frac{\mathbb{E} |\mathbf{R}_{jk}|^p}{n^2}. \quad (3.23)$$

Let us consider now the term  $\mathcal{A}_0$ . Applying (3.22) with  $t = 1$ ,  $l = 5 - \mu$  we rewrite  $\mathcal{A}_0$  as follows

$$\begin{aligned} \mathcal{A}_0 &= \sum_{\mu=0}^4 \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E} X_{ab}^\mu \mathbb{E} [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} \varphi(\mathbf{T}_{jk}) \\ &\quad + \sum_{\mu=0}^4 \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \sum_{l=1}^{4-\mu} \frac{1}{l!} \mathbb{E} X_{ab}^\mu [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} (\mathbf{R}_{jk} - \mathbf{T}_{jk})^l \varphi^{(l)}(\mathbf{T}_{jk}) \\ &\quad + \sum_{\mu=0}^4 \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \frac{1}{(4-\mu)!} \mathbb{E} X_{ab}^\mu [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} (\mathbf{R}_{jk} - \mathbf{T}_{jk})^{5-k} (1-\theta)^{4-k} \\ &\quad \times \varphi^{(5-\mu)}(\mathbf{T}_{jk} + \theta(\mathbf{R}_{jk} - \mathbf{T}_{jk})) \\ &=: \mathcal{A}_{00} + \mathcal{A}_{01} + \mathcal{A}_{02}. \end{aligned}$$

Let us study the term  $\mathcal{A}_{01}$ . We may write  $\mathcal{A}_{01} = \sum_{\mu=0}^4 \sum_{l=1}^{4-\mu} (-1)^\mu (l!)^{-1} \mathcal{B}_{\mu l}$ . We fix some  $\mu = 0, \dots, 4$  and some  $l = 1, \dots, 4 - \mu$ . We may apply now (3.16) with  $m \geq 4$  and get  $\mathcal{B}_{\mu l} =$

$\mathcal{B}_{\mu l}^{(0)} + \mathcal{B}_{\mu l}^{(1)} + \mathcal{B}_{\mu l}^{(2)}$ , where

$$\begin{aligned} \mathcal{B}_{\mu l}^{(0)} &:= \sum_{\substack{\mu_1 + \dots + \mu_m = l \\ \mu + \mu_1 + 2\mu_2 + \dots + m\mu_m \leq 4}} \frac{C_{\mu_1, \dots, \mu_m}^l}{n^{\frac{\mu}{2} + \frac{\mu_1}{2} + \frac{2\mu_2}{2} + \dots + \frac{m\mu_m}{2}}} \mathbb{E} X_{ab}^{\mu + \mu_1 + 2\mu_2 + \dots + m\mu_m} \\ &\quad \times \mathbb{E} [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} [(\mathbf{TE}^{(a,b)}) \mathbf{T}]_{jk}^{\mu_1} \dots [(\mathbf{TE}^{(a,b)})^m \mathbf{T}]_{jk}^{\mu_m} \varphi^{(l)}(\mathbf{T}_{jk}), \\ \mathcal{B}_{\mu l}^{(1)} &:= \sum_{\substack{\mu_1 + \dots + \mu_m = l \\ \mu + \mu_1 + 2\mu_2 + \dots + m\mu_m > 4}} \frac{C_{\mu_1, \dots, \mu_m}^l}{n^{\frac{\mu}{2} + \frac{\mu_1}{2} + \frac{2\mu_2}{2} + \dots + \frac{m\mu_m}{2}}} \mathbb{E} X_{ab}^{\mu + \mu_1 + 2\mu_2 + \dots + m\mu_m} \\ &\quad \times \mathbb{E} [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} [(\mathbf{TE}^{(a,b)}) \mathbf{T}]_{jk}^{\mu_1} \dots [(\mathbf{TE}^{(a,b)})^m \mathbf{T}]_{jk}^{\mu_m} \varphi^{(l)}(\mathbf{T}_{jk}), \\ \mathcal{B}_{\mu l}^{(2)} &:= \sum_{\mu_{m+1}=1}^l \sum_{\mu_1 + \dots + \mu_m = l - \mu_{m+1}} \frac{C_{\mu_1, \dots, \mu_{m+1}}^l}{n^{\frac{\mu}{2} + \frac{\mu_1}{2} + \frac{2\mu_2}{2} + \dots + \frac{(m+1)\mu_{m+1}}{2}}} \\ &\quad \times \mathbb{E} X_{ab}^{\mu + \mu_1 + 2\mu_2 + \dots + (m+1)\mu_{m+1}} \varphi^{(l)}(\mathbf{T}_{jk}) \\ &\quad \times [(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jk} [(\mathbf{TE}^{(a,b)}) \mathbf{T}]_{jk}^{\mu_1} \dots [(\mathbf{TE}^{(a,b)})^m \mathbf{T}]_{jk}^{\mu_m} [(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jk}^{\mu_{m+1}}, \end{aligned}$$

where  $C_{\mu_1, \dots, \mu_{m+1}}^l$  are multinomial coefficients. We now introduce in (3.10)

$$\mathcal{I}(p) := \mathcal{A}_{00} + \sum_{\mu=0}^4 \sum_{l=1}^{4-\mu} \frac{(-1)^\mu}{l!} \mathcal{B}_{\mu l}^{(0)}. \tag{3.24}$$

The estimate of  $\mathcal{B}_{\mu l}^{(1)}$  is similar to  $\mathcal{A}_1$ , see (3.23). Here, one has to expand  $\varphi^{(l)}(\mathbf{T}_{jk})$  at the point  $\tilde{\mathbf{T}}_{jk}$  and apply the same arguments as before. To estimate  $\mathcal{B}_{\mu l}^{(2)}$  we again expand  $\varphi^{(l)}(\mathbf{T}_{jk})$  at the point  $\tilde{\mathbf{T}}_{jk}$ . From here on, we may apply the same arguments as for  $\mathcal{A}_1$  and  $\mathcal{A}_2$  (see (3.23) and (3.19)) by taking  $m$  large enough and bound  $|\mathbf{R}_{ak}|$  or  $|\mathbf{R}_{bk}|$  by  $n$ . The same procedure applies to  $\mathcal{A}_{02}$ . We finally get

$$\mathbb{E} |\mathbf{R}_{jk}|^p = \mathcal{I}(p) + r_n(p),$$

where

$$|r_n(p)| \leq \frac{C^p}{n^2} + \frac{\mathbb{E} |\mathbf{R}_{jk}|^p}{n^2}.$$

The proof of (3.9) is similar. Let us denote by  $m_n^{(a,b)}(z)$  the Stieltjes transform of  $\mathbf{W}^{(\{a\}, \{b\})}$ . It is easy to see that

$$\frac{1}{z + m_n(z)} = \frac{1}{z + m_n^{(a,b)}(z)} + \frac{m_n(z) - m_n^{(a,b)}(z)}{(z + m_n(z))(z + m_n^{(a,b)}(z))}.$$

We may use this formula recursively together with (3.16) and get (3.9). We omit the details.  $\square$

### 3.2. Moments of the imaginary part of diagonal entries of the resolvent

**Lemma 3.6.** *Assuming conditions (C1) there exist positive constants  $H_0$  depending on  $u_0$ ,  $V$  and positive constants  $A_0, A_1$  depending on  $H_0, \delta$  such that for all  $1 \leq p \leq A_1(nv)^\kappa$  and  $z \in \mathbb{D}$  we get*

$$\max_{j \in \mathbb{T}} \mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}(z) \leq H_0^p \operatorname{Im}^p s(z) + \frac{H_0^p p^{2p}}{(nv)^p}.$$

**Proof.** See [18], Lemma 5.1. □

**Proof of (3.4).** Let  $\tilde{v} = n^{-1} \log^{\frac{1}{\kappa}} n$ , then  $1 \leq p \leq A_1(n\tilde{v}) = A_1 \log n$ . Without loss of generality we may consider  $p = A_1 \log n$  only. Using the notation introduced in (3.1) the statement of Lemma 3.6 may be rewritten (up to constant) in the following way

$$\max_{j \in \mathbb{T}} \mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}(z) \leq H_0^p p^p \Psi^p(z) \leq H_0^p \log^p n \Psi^p(z) \tag{3.25}$$

for all  $V \geq v \geq \tilde{v}$ . To prove (3.2), we need to descent from  $\tilde{v}$  to  $v_0$  while maintaining that  $p = A_1 \log n$ . Applying Lemma 3.7 below with  $s_0 := \log^{\frac{1}{\kappa}-1} n$  we may show that for all  $v \geq v_0$

$$\max_{j \in \mathbb{T}} \mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}(v) \leq H_0^p \log^{(\frac{2}{\kappa}-1)p} n \Psi^p(z).$$

It remains to remove the log factor from the r.h.s. on the previous equation. We may proceed as before and arrive at the following bound for any  $j \in \mathbb{T}$

$$\mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}(v) \leq C_1^p \Psi^p(z) + C_2 \mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}^y(v).$$

See Lemma 3.8 below. It follows from Lemma 4.2 in the next section that

$$\max_{j \in \mathbb{T}} \mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}^y(v) \leq H_0^p \Psi^p(z). \tag{3.25}$$

**Lemma 3.7.** *Let  $C_0$  be some constant. Assume that for all  $v' \geq \tilde{v}$  and all  $1 \leq p \leq A_1(nv')^\kappa$  we have*

$$\max_{j \in \mathbb{T}} \mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}(v') \leq H_0^p \log^p n \Psi^p(z). \tag{3.26}$$

Then for all  $s_0 \geq 1, v \geq \tilde{v}/s_0$  we have

$$\max_{j, k \in \mathbb{T}} \mathbb{E} \operatorname{Im}^p \mathbf{R}_{jk}(v) \leq s_0^{2p} H_0^p \log^p n \Psi^p(z).$$

**Proof.** The proof is similar to the proof of Lemma 3.3. Applying Lemma A.2 we get

$$\mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}(v) \leq s_0^p \mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}(s_0 v) \leq s_0^p H_0^p \Psi^q(s_0 v) \leq s_0^{2p} H_0^p \log^p n \Psi^q(v).$$

We consider now the case  $j \neq k$ . Then

$$\operatorname{Im} \mathbf{R}_{jk} = v[\mathbf{R}\mathbf{R}^*]_{jk} = v \sum_{l=1}^n \mathbf{R}_{jl} \mathbf{R}_{lk}^* \leq v \left( \sum_{l=1}^n |\mathbf{R}_{jl}|^2 \right)^{\frac{1}{2}} \left( \sum_{l=1}^n |\mathbf{R}_{lk}|^2 \right)^{\frac{1}{2}}.$$

Applying Lemma A.3, we get

$$\operatorname{Im} \mathbf{R}_{jk} \leq \sqrt{\operatorname{Im} \mathbf{R}_{jj} \operatorname{Im} \mathbf{R}_{kk}}.$$

It follows that

$$\mathbb{E} \operatorname{Im}^p \mathbf{R}_{jk}(v) \leq s_0^{2p} H_0^p \log^p n \Psi^q(v). \quad \square$$

**Lemma 3.8.** For all  $v \geq v_0$  and  $5 \leq p \leq \log n$  there exists a constant  $C_0 > 0$  such that

$$\mathbb{E} |\operatorname{Im} \mathbf{R}_{jj}(v)|^p \leq C_0^p \Psi^p(z) + \mathbb{E} |\operatorname{Im} \mathbf{R}_{jj}^y(v)|^p. \quad (3.27)$$

**Proof.** The proof of this lemma is similar to the proof of Lemma 3.5. We shall omit the details, but emphasize some important points. Applying (3.16) with some  $m \geq 0$  we get

$$\begin{aligned} \operatorname{Im} \mathbf{R}_{jj} &= \operatorname{Im} \mathbf{T}_{jj} + \sum_{\mu=1}^m \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} X_{ab}^\mu \operatorname{Im}[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj} \\ &\quad + \frac{(-1)^{m+1}}{n^{\frac{m+1}{2}}} \mathbf{X}_{ab}^{m+1} \operatorname{Im}[(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jj}. \end{aligned}$$

Introduce the function  $\psi(x) := x^{p-1}$  and write  $\mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj} = \mathbb{E} \operatorname{Im} \mathbf{R}_{jj} \psi(\operatorname{Im} \mathbf{R}_{jj})$ . Similarly to (3.18), we get

$$\begin{aligned} \mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj} &= \sum_{\mu=0}^4 \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E} X_{ab}^\mu \operatorname{Im}[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj} \psi(\operatorname{Im} \mathbf{R}_{jj}) \\ &\quad + \sum_{\mu=5}^m \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E} X_{ab}^\mu \operatorname{Im}[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj} \psi(\operatorname{Im} \mathbf{R}_{jj}) \\ &\quad + \frac{1}{n^{\frac{m+1}{2}}} \mathbb{E} \mathbf{X}_{ab}^{m+1} \operatorname{Im}[(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jj} \psi(\operatorname{Im} \mathbf{R}_{jj}) \\ &=: \mathcal{A}_0 + \mathcal{A}_1 + \mathcal{A}_2. \end{aligned} \quad (3.28)$$

We shall keep the same notations as in the proof of Lemma 3.5. Let us consider the term  $\mathcal{A}_2$ . Repeating the same arguments as in the proof of Lemma 3.5 for the corresponding terms and

applying  $\|\mathbf{R}\| \leq v^{-1}$ , we get

$$\begin{aligned} |\mathcal{A}_2| &\leq \frac{1}{n^{\frac{m+1}{2}}} \mathbb{E} |X_{ab}|^{m+1} \operatorname{Im}[(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jj} \operatorname{Im}^{p-1} \mathbf{R}_{jj} \\ &\leq \frac{\log^c n}{n^{\frac{m+1}{2}} v} \mathbb{E}^{\frac{1}{p}} |\mathbf{X}_{ab}|^{(m+1)p} \mathbb{E}^{\frac{p-1}{p}} \operatorname{Im}^p \mathbf{R}_{jj}. \end{aligned}$$

Since  $|X_{jk}| \leq Dn^{\frac{1}{2}-\phi}$  we obtain that

$$|\mathcal{A}_2| \leq \frac{C \log^{c(\kappa)} n}{n^{\phi(m+1)} v} \mathbb{E}^{\frac{p-1}{p}} \operatorname{Im}^p \mathbf{R}_{jj}.$$

We may choose  $m$  such that  $\phi(m+1) = 4$ . Applying  $2 = \frac{2}{p} + \frac{2(p-1)}{p}$ , Young's inequality and  $(nv)^{-1} \leq \Psi(z)$ , we obtain

$$|\mathcal{A}_2| \leq \frac{C\Psi(z)}{n^{\frac{2}{p} + \frac{2(p-1)}{p}}} \mathbb{E}^{\frac{p-1}{p}} \operatorname{Im}^p \mathbf{R}_{jj} \leq \frac{C^p \Psi^p(z)}{n^2} + \frac{\mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}}{n^2}. \tag{3.29}$$

Let us consider  $\mathcal{A}_1$  and split it into the sum  $\mathcal{A}_1 = \mathcal{A}_{1,5} + \dots + \mathcal{A}_{1,m}$ . For an arbitrary  $5 \leq \mu \leq m$ , we get  $\mathcal{A}_{1,\mu} = \mathcal{B}_1 + \mathcal{B}_2$ , where

$$\begin{aligned} \mathcal{B}_1 &:= \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E} X_{ab}^\mu \operatorname{Im}[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj} \psi(\operatorname{Im} \tilde{\mathbf{T}}_{jj}), \\ \mathcal{B}_2 &:= \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E} X_{ab}^\mu \operatorname{Im}[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj} [\psi(\operatorname{Im} \mathbf{R}_{jj}) - \psi(\operatorname{Im} \tilde{\mathbf{T}}_{jj})] \end{aligned}$$

and  $\tilde{\mathbf{T}}_{jj}$  are defined in (3.20). The term  $[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj}$  is a sum of  $2^\mu$  terms of the following type

$$\mathbf{T}_{ji_1} \mathbf{T}_{i_1 i_2} \cdots \mathbf{T}_{i_{\mu-1} i_\mu} \mathbf{T}_{i_\mu j},$$

where  $i_l = a$  or  $i_l = b$  for  $l = 1, \dots, \mu$ . The imaginary part of such a product may be bounded from above by a product where at least one factor is  $|\operatorname{Im} \mathbf{T}_{i_l i_{l+1}}|$ . All other factors may be bounded by their absolute values due to the first statement (3.2) of Lemma 3.1. Applying Hölder's inequality to this product and Lemma 3.7, we get

$$|\mathcal{B}_1| \leq \frac{C^p \Psi^p(z)}{n^2} + \frac{\mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}}{n^2}.$$

By the same arguments we get similar bounds for  $\mathcal{B}_2$  and  $\mathcal{A}_0$ . We omit the details. □

### 4. Bounds for moments of diagonal entries of the resolvent in the sub-Gaussian case

As mentioned in the previous section, we have to bound the moments of the diagonal entries of the resolvent in the sub-Gaussian case. We denote

$$\tilde{\mathbb{D}} := \{z = u + iv \in \mathbb{C} : |u| \leq u_0, V \geq v \geq \tilde{v}_0 := A_0 n^{-1}\},$$

where  $u_0, V > 0$  are any fixed real numbers and  $A_0$  is some large constant determined below. Comparing with  $\mathbb{D}$  we allow to descent to  $A_0 n^{-1}$  along  $v$ .

We say that the conditions **(CG)** are satisfied if  $X_{jk}$  satisfies the conditions **(C0)** and have a sub-Gaussian distribution. It is well known that the random variables  $\xi$  are sub-Gaussian if and only if  $\mathbb{E} |\xi|^p = O(p^{\frac{p}{2}})$  as  $p \rightarrow \infty$ . We define the sub-Gaussian norm of  $\xi$  as

$$\|\xi\|_{\psi_2} := \sup_{p \geq 1} p^{-\frac{1}{2}} \mathbb{E}^{\frac{1}{p}} |\xi|^p.$$

In what follows, we assume that  $K := \|X_{jk}\|_{\psi_2}$ .

**Lemma 4.1.** *Assuming the conditions **(CG)** there exist a positive constant  $C_0$  depending on  $u_0, V$  and positive constants  $A_0, A_1$  depending on  $C_0$  and  $K$  such that for all  $z \in \tilde{\mathbb{D}}$  and  $1 \leq p \leq A_1 n v$  we have*

$$\max_{j,k \in \mathbb{T}} \mathbb{E} |\mathbf{R}_{jk}(z)|^p \leq C_0^p \tag{4.1}$$

and

$$\mathbb{E} \frac{1}{|z + m_n(z)|^p} \leq C_0^p.$$

The proof of Lemma 4.1 is based on several auxiliary results and will be given at the end of this section. In this proof, will shall use ideas from [26] and [28]. One of main ingredients of the proof is the descent method for  $\mathbf{R}_{jj}$  which is based on Lemma 4.3 below and Lemma A.1 in the Appendix, which in this form appeared in [8]. Comparing the result of Lemma 4.1 with [8], Lemma 3.4, that in the latter the power  $p$  is bounded from above by  $(nv)^{\frac{1}{4}}$ , which is non-optimal.

Since  $u$  is fixed and  $|u| \leq u_0$  we shall omit  $u$  from the notation of the resolvent and denote  $\mathbf{R}(v) := \mathbf{R}(z)$ . Sometimes in order to simplify notations we shall also omit the argument  $v$  in  $\mathbf{R}(v)$  and just write  $\mathbf{R}$ .

We also estimate the moments of the imaginary part of the diagonal entries of the resolvent. Recall that (see definition (3.1))

$$\Psi(z) := \text{Im } s(z) + \frac{p}{nv}.$$

To simplify notations, we will often write  $\Psi(v)$  and  $\Psi$  instead of  $\Psi(z)$ .

**Lemma 4.2.** *Assuming conditions (CG) there exist a positive constant  $H_0$  depending on  $u_0, V$  and positive constants  $A_0, A_1$  depending on  $H_0$  and  $K$  such that for all  $1 \leq p \leq A_1 n v$  and  $z \in \widetilde{\mathbb{D}}$  we get*

$$\max_{j \in \mathbb{T}} \mathbb{E} |\operatorname{Im} \mathbf{R}_{jj}(z)|^p \leq H_0^p \Psi^p(z).$$

Note that the values of  $A_0$  and  $A_1$  in this lemma are different from the values of corresponding quantities in Lemma 4.1, but for simplicity we shall use the same notations. Applying both lemmas, we shall restrict the upper limit of the moment of order  $p$  to the minimum of the two  $A_1$ 's and the lower end of the range of  $v$  to the maximum of the two  $A_0$ 's via  $v \geq A_0 n^{-1}$ .

For any  $j \in \mathbb{T}_{\mathbb{J}}$ , we may express  $\mathbf{R}_{jj}^{(\mathbb{J})}$  in the following way

$$\mathbf{R}_{jj}^{(\mathbb{J})} = \frac{1}{-z + \frac{X_{jj}}{\sqrt{n}} - \frac{1}{n} \sum_{l, k \in \mathbb{T}_{\mathbb{J}, j}} X_{jk} X_{jl} \mathbf{R}_{lk}^{(\mathbb{J}, j)}}. \tag{4.2}$$

Let  $\varepsilon_j^{(\mathbb{J})} := \varepsilon_{1j}^{(\mathbb{J})} + \varepsilon_{2j}^{(\mathbb{J})} + \varepsilon_{3j}^{(\mathbb{J})} + \varepsilon_{4j}^{(\mathbb{J})}$ , where

$$\begin{aligned} \varepsilon_{1j}^{(\mathbb{J})} &:= \frac{1}{\sqrt{n}} X_{jj}, & \varepsilon_{2j}^{(\mathbb{J})} &:= -\frac{1}{n} \sum_{l \neq k \in \mathbb{T}_{\mathbb{J}, j}} X_{jk} X_{jl} \mathbf{R}_{kl}^{(\mathbb{J}, j)}, \\ \varepsilon_{3j}^{(\mathbb{J})} &:= -\frac{1}{n} \sum_{k \in \mathbb{T}_{\mathbb{J}, j}} (X_{jk}^2 - 1) \mathbf{R}_{kk}^{(\mathbb{J})}(z), & \varepsilon_{4j}^{(\mathbb{J})} &:= \frac{1}{n} (\operatorname{Tr} \mathbf{R}^{(\mathbb{J})} - \operatorname{Tr} \mathbf{R}^{(\mathbb{J}, j)}(z)). \end{aligned}$$

We also introduce the quantities  $\Lambda_n^{(\mathbb{J})}(z) := m_n^{(\mathbb{J})}(z) - s(z)$  and

$$T_n^{(\mathbb{J})} := \frac{1}{n} \sum_{j \in \mathbb{T}_{\mathbb{J}}} \varepsilon_j^{(\mathbb{J})} \mathbf{R}_{jj}^{(\mathbb{J})}.$$

The following lemma, Lemma 4.3, allows to recursively estimate the moments of the diagonal entries of the resolvent. The proof of the first part of this lemma may be found in [8] and it is included here for the readers convenience.

**Lemma 4.3.** *For an arbitrary set  $\mathbb{J} \subset \mathbb{T}$  and all  $j \in \mathbb{T}_{\mathbb{J}}$  there exist a positive constant  $c_0$  depending on  $u_0, V$  only such that for all  $z = u + i v$  with  $V \geq v > 0$  and  $|u| \leq u_0$  we have*

$$|\mathbf{R}_{jj}^{(\mathbb{J})}| \leq c_0 (1 + |T_n^{(\mathbb{J})}|^{\frac{1}{2}} |\mathbf{R}_{jj}^{(\mathbb{J})}| + |\varepsilon_j^{(\mathbb{J})}| |\mathbf{R}_{jj}^{(\mathbb{J})}|) \tag{4.3}$$

and

$$\frac{1}{|z + m_n^{(\mathbb{J})}(z)|} \leq c_0 \left( 1 + \frac{|T_n^{(\mathbb{J})}|^{\frac{1}{2}}}{|z + m_n^{(\mathbb{J})}(z)|} \right). \tag{4.4}$$

**Proof.** See [18], Lemma 4.2. □

**Lemma 4.4.** *Assume that the conditions (CG) hold. Let  $C_0$  and  $s_0$  be arbitrary numbers such that  $H_0 \geq \max(1/V, 6c_0)$ ,  $s_0 \geq 2$ . There exist a sufficiently large constant  $A_0$  and small constant  $A_1$  depending on  $C_0, s_0, V$  only such that the following statement holds. Fix some  $\tilde{v} : \tilde{v}_0 s_0 \leq \tilde{v} \leq V$ . Suppose that for some integer  $L > 0$ , all  $u, v', q$  such that  $\tilde{v} \leq v' \leq V, |u| \leq u_0, 1 \leq q \leq A_1(nv')$*

$$\max_{\mathbb{J}:|\mathbb{J}|\leq L} \max_{l,k \in \mathbb{T}_{\mathbb{J}}} \mathbb{E}|\mathbf{R}_{lk}^{(\mathbb{J})}(v')|^q \leq C_0^q. \tag{4.5}$$

Then for all  $u, v, q$  such that  $\tilde{v}/s_0 \leq v \leq V, |u| \leq u_0, 1 \leq q \leq A_1(nv)$

$$\max_{\mathbb{J}:|\mathbb{J}|\leq L-1} \max_{l,k \in \mathbb{T}_{\mathbb{J}}} \mathbb{E}|\mathbf{R}_{lk}^{(\mathbb{J})}(v)|^q \leq C_0^q.$$

**Proof.** Let us fix an arbitrary  $s_0 \geq 2$  and  $v \geq \tilde{v}/s_0, \mathbb{J} \subset \mathbb{T}$  such that  $|\mathbb{J}| \leq L - 1$ . In the following let  $j, k \in \mathbb{T}_{\mathbb{J}}$ . By an obvious inequality we have

$$\mathbb{E}|\varepsilon_j^{(\mathbb{J})}|^{2q} \leq 3^{2q} (\mathbb{E}|\varepsilon_{1j}^{(\mathbb{J})}|^{2q} + \mathbb{E}|\varepsilon_{2j}^{(\mathbb{J})} + \varepsilon_{3j}^{(\mathbb{J})}|^{2q} + \mathbb{E}|\varepsilon_{4j}^{(\mathbb{J})}|^{2q}).$$

From (CG) and Lemmas A.5–A.6 we may conclude that

$$\mathbb{E}|\varepsilon_j^{(\mathbb{J})}|^{2q} \leq \frac{C^q q^q}{n^q} + \frac{C^q q^q}{(nv)^q} \mathbb{E} \text{Im}^q m_n^{(\mathbb{J},j)}(z) + \frac{C^q q^{2q}}{(nv)^{2q}}.$$

Applying Lemma A.1 and (4.5), we obtain the following estimate

$$\mathbb{E}[\text{Im} m_n^{(\mathbb{J},j)}(z)]^q \leq s_0^q C_0^q.$$

In view of these inequalities, we may write

$$\mathbb{E}|\varepsilon_j^{(\mathbb{J})}|^{2q} \leq \frac{(CC_0 s_0)^q q^q}{(nv)^q} + \frac{C^q q^{2q}}{(nv)^{2q}}. \tag{4.6}$$

Similarly, we can estimate

$$\begin{aligned} \mathbb{E}|T_n^{(\mathbb{J})}|^q &\leq \left( \frac{1}{n} \sum_{j \in \mathbb{T}_{\mathbb{J}}} \mathbb{E}|\varepsilon_j^{(\mathbb{J})}|^{2q} \right)^{1/2} \left( \frac{1}{n} \sum_{j \in \mathbb{T}_{\mathbb{J}}} \mathbb{E}|\mathbf{R}_{jj}^{(\mathbb{J})}(v)|^{2q} \right)^{1/2} \\ &\leq \frac{(CC_0^{\frac{3}{2}} s_0^{\frac{3}{2}})^q q^{\frac{q}{2}}}{(nv)^{\frac{q}{2}}} + \frac{(CC_0 s_0)^q q^q}{(nv)^q}. \end{aligned} \tag{4.7}$$

Applying Hölder’s inequality and Lemma 4.3, we write

$$\mathbb{E}|\mathbf{R}_{jj}^{(\mathbb{J})}(v)|^q \leq (3c)^q \left( 1 + \mathbb{E}^{\frac{1}{2}} |T_n^{(\mathbb{J})}|^q \mathbb{E}^{\frac{1}{2}} |\mathbf{R}_{jj}^{(\mathbb{J})}(v)|^{2q} + \mathbb{E}^{\frac{1}{2}} |\varepsilon_j^{(\mathbb{J})}|^{2q} \mathbb{E}^{\frac{1}{2}} |\mathbf{R}_{jj}^{(\mathbb{J})}(v)|^{2q} \right).$$

Lemma A.1 and assumption (4.5) imply  $\mathbb{E}|\mathbf{R}_{jj}^{(\mathbb{J})}(v)|^{2q} \leq C_0^{2q} s_0^{2q}$ . Hence, we may use (4.6) and (4.7) to show that

$$\mathbb{E}|\mathbf{R}_{jj}^{(\mathbb{J})}(v)|^q \leq (3c)^q \left( 1 + \frac{(CC_0s_0)^{2q} q^{\frac{q}{4}}}{(nv)^{\frac{q}{4}}} + \frac{(CC_0s_0)^{\frac{3q}{2}} q^{\frac{q}{2}}}{(nv)^{\frac{q}{2}}} + \frac{(CC_0s_0)^q q^q}{(nv)^q} \right). \tag{4.8}$$

The off-diagonal entries  $\mathbf{R}_{jk}^{(\mathbb{J})}$  may be expressed as follows

$$\mathbf{R}_{jk}^{(\mathbb{J})} = -\frac{1}{\sqrt{n}} \sum_{l \in \mathbb{T}_{\mathbb{J},j}} X_{jl} \mathbf{R}_{lk}^{(\mathbb{J},j)} \mathbf{R}_{jj}^{(\mathbb{J})}.$$

Applying Hölder’s inequality, Lemma A.1 and assumption (4.5) we obtain

$$\mathbb{E}|\mathbf{R}_{jk}^{(\mathbb{J})}|^q \leq n^{-\frac{q}{2}} \mathbb{E}^{\frac{1}{2}} \left| \sum_{l \in \mathbb{T}_{\mathbb{J},j}} X_{jl} \mathbf{R}_{lk}^{(\mathbb{J},j)} \right|^{2q} \mathbb{E}^{\frac{1}{2}} |\mathbf{R}_{jj}^{(\mathbb{J})}|^{2q} \leq (s_0 C_0)^q n^{-\frac{q}{2}} \mathbb{E}^{\frac{1}{2}} \left| \sum_{l \in \mathbb{T}_{\mathbb{J},j}} X_{jl} \mathbf{R}_{lk}^{(\mathbb{J},j)} \right|^{2q}.$$

Khinchine’s inequality for sub-Gaussian random variables, Lemmas A.3, A.1 and assumption (4.5) together imply

$$\mathbb{E}^{\frac{1}{2}} \left| \sum_{l \in \mathbb{T}_{\mathbb{J},j}} X_{jl} \mathbf{R}_{lk}^{(\mathbb{J},j)} \right|^{2q} \leq C^q q^{\frac{q}{2}} \mathbb{E}^{\frac{1}{2}} \left| \sum_{l \in \mathbb{T}_{\mathbb{J},j}} |\mathbf{R}_{lk}^{(\mathbb{J},j)}|^2 \right|^q \leq \frac{(Cs_0 C_0 q)^{\frac{q}{2}}}{v^{\frac{q}{2}}}.$$

From the last two inequalities, we conclude the following bound

$$\mathbb{E}|\mathbf{R}_{jk}^{(\mathbb{J})}|^q \leq \frac{(CC_0^{\frac{3}{2}} s_0^{\frac{3}{2}})^q q^{\frac{q}{2}}}{(nv)^{\frac{q}{2}}}. \tag{4.9}$$

We may choose now the constants  $A_0$  sufficiently large, respectively  $A_1$  sufficiently small such that (4.8)–(4.9) result in

$$\max_{j,k \in \mathbb{T}_{\mathbb{J}}} \mathbb{E}|\mathbf{R}_{jk}^{(\mathbb{J})}(v)|^q \leq C_0^q$$

for  $1 \leq q \leq A_1(n\tilde{v}/s_0)$ ,  $v \geq \tilde{v}/s_0$ . □

**Proof of Lemma 4.1.** We first prove (4.1). Let us choose some sufficiently large constant  $C_0 > \max(1/V, 6c_0)$  and fix  $s_0 := 2$ . Here  $c_0$  is defined in Lemma 4.3. We also choose  $A_0$  and  $A_1$  as in Lemma 4.4. Let  $L := \lceil \log_{s_0} V/\tilde{v}_0 \rceil + 1$ . Since  $\|\mathbf{R}^{(\mathbb{J})}(V)\| \leq V^{-1}$  we may write

$$\max_{\mathbb{J}:|\mathbb{J}|\leq L} \max_{l,k \in \mathbb{T}_{\mathbb{J}}} \mathbb{E}|\mathbf{R}_{lk}^{(\mathbb{J})}(V)|^p \leq C_0^p$$

for all  $u, p$  such that  $|u| < 2$  and  $1 \leq p \leq A_1(nV)$ . Fix arbitrary  $v : V/s_0 \leq v \leq V$  and  $p : 1 \leq p \leq A_1(nv)$ . Lemma 4.4 yields that

$$\max_{\mathbb{J}:|\mathbb{J}|\leq L-1} \max_{l,k \in \mathbb{T}_{\mathbb{J}}} \mathbb{E}|\mathbf{R}_{lk}^{(\mathbb{J})}(v)|^p \leq C_0^p$$

for  $1 \leq p \leq A_1(nV/s_0)$ ,  $v \geq V/s_0$ . We may repeat this procedure  $L$  times and finally obtain

$$\max_{l,k \in \mathbb{T}} \mathbb{E} |\mathbf{R}_{lk}(v)|^p \leq C_0^p$$

for  $1 \leq p \leq A_1(nV/s_0^L) \leq A_1(n\tilde{v}_0)$  and  $v \geq V/s_0^L = \tilde{v}_0$ . Thus, we proved (4.1). Similarly one may prove (4.4). □

The following lemma is the analogue of Lemma 4.3 and provides a recurrence relation for  $\text{Im } \mathbf{R}_{jj}$ .

**Lemma 4.5.** *For any set  $\mathbb{J}$  and  $j \in \mathbb{T}_{\mathbb{J}}$  there exists a positive constant  $C_1$  depending on  $u_0, V$  such that for all  $z = u + iv$  with  $V \geq v > 0$  and  $|u| \leq u_0$  we have*

$$\begin{aligned} \text{Im } \mathbf{R}_{jj}^{(\mathbb{J})}(z) &\leq C_1 [\text{Im } s(z) (1 + (|\varepsilon_j^{(\mathbb{J})}| + |T_n^{(\mathbb{J})}|^{\frac{1}{2}}) |\mathbf{R}_{jj}^{(\mathbb{J})}(z)|) + |\text{Im } \varepsilon_j^{(\mathbb{J})} + \text{Im } \Lambda_n^{(\mathbb{J})}| |\mathbf{R}_{jj}^{(\mathbb{J})}(z)| \\ &\quad + (|\varepsilon_j^{(\mathbb{J})}| + |T_n^{(\mathbb{J})}|^{\frac{1}{2}}) \text{Im } \mathbf{R}_{jj}^{(\mathbb{J})}(z)]. \end{aligned}$$

**Proof.** The proof is similar to the proof of Lemma 4.3 and by this reason it is omitted. □

**Lemma 4.6.** *Assume that the conditions (CG) hold. Let  $H_0$  be sufficiently large positive constant and  $s_0$  be an arbitrary number such that  $s_0 \geq 2$ . There exist sufficiently large  $A_0$  and small  $A_1$  depending on  $H_0, s_0, V$  only such that the following holds. Fix some  $\tilde{v} : \tilde{v}_0 s_0 \leq \tilde{v} \leq V$ . Suppose that for some integer  $L > 0$ , all  $u, v', q$  such that  $\tilde{v} \leq v' \leq V$ ,  $|u| \leq u_0$ ,  $1 \leq q \leq A_1(nv')$*

$$\max_{\mathbb{J}: |\mathbb{J}| \leq L} \max_{l \in \mathbb{T}_{\mathbb{J}}} \mathbb{E} \text{Im}^q \mathbf{R}_{ll}^{(\mathbb{J})}(v') \leq H_0^q \Psi^q(v'). \tag{4.10}$$

Then for all  $u, v, q$  such that  $\tilde{v}/s_0 \leq v \leq V$ ,  $|u| \leq u_0$ ,  $1 \leq q \leq A_1(nv)$

$$\max_{\mathbb{J}: |\mathbb{J}| \leq L-1} \max_{l \in \mathbb{T}_{\mathbb{J}}} \mathbb{E} \text{Im}^q \mathbf{R}_{ll}^{(\mathbb{J})}(v) \leq H_0^q \Psi^q(v).$$

**Proof.** From Lemma 4.5 it follows that

$$\begin{aligned} \mathbb{E} \text{Im}^q \mathbf{R}_{jj}^{(\mathbb{J})} &\leq (CC_0)^q \text{Im}^q s(z) \mathbb{E}^{\frac{1}{2}} (1 + (|\varepsilon_j^{(\mathbb{J})}| + |T_n^{(\mathbb{J})}|^{\frac{1}{2}})^{2q} \\ &\quad + (CC_0)^q \mathbb{E}^{\frac{1}{2}} |\text{Im } \varepsilon_j^{(\mathbb{J})} + \text{Im } \Lambda_n^{(\mathbb{J})}|^{2q} \\ &\quad + C^q \mathbb{E}^{\frac{1}{2}} (|\varepsilon_j^{(\mathbb{J})}| + |T_n^{(\mathbb{J})}|^{\frac{1}{2}})^{2q} \mathbb{E}^{\frac{1}{2}} \text{Im}^{2q} \mathbf{R}_{jj}^{(\mathbb{J})}. \end{aligned}$$

To estimate  $\mathbb{E} |\varepsilon_j^{(\mathbb{J})}|^{2q}$  and  $\mathbb{E} |T_n^{(\mathbb{J})}|^q$  we may proceed as in Lemma 4.4. We obtain the following inequalities

$$\mathbb{E} |\varepsilon_j^{(\mathbb{J})}|^{2q} \leq 3^{2q} \left[ \frac{C^q q^q}{n^q} + \frac{(C)^q q^q}{(nv)^q} + \frac{C^q q^{2q}}{(nv)^{2q}} + \frac{1}{(nv)^{2q}} \right] \leq \frac{C^q q^q}{(nv)^q} \tag{4.11}$$

and

$$\mathbb{E}|T_n^{(\mathbb{J})}|^q \leq C_0^q \left( \frac{1}{n} \sum_{j \in \mathbb{T}_j} \mathbb{E}|\varepsilon_j^{(\mathbb{J})}|^{2q} \right)^{1/2} \leq \frac{C^q q^{\frac{q}{2}}}{(nv)^{\frac{q}{2}}}. \tag{4.12}$$

Choosing  $v' := s_0 v \geq v_1$  we may show that  $2q \leq A_1 n v'$ . Applying Lemma A.2 and using the assumption (4.10) we get

$$\mathbb{E} \operatorname{Im}^{2q} \mathbf{R}_{jj}^{(\mathbb{J})}(v) \leq s_0^{2q} \mathbb{E} \operatorname{Im}^{2q} \mathbf{R}_{jj}^{(\mathbb{J})}(s_0 v) \leq s_0^{2q} H_0^{2q} \Psi^{2q}(s_0 v).$$

Since we need an estimate involving  $\Psi^{2q}(v)$  instead of  $\Psi^{2q}(s_0 v)$  on the r.h.s. of the previous inequality we need to perform a descent along the imaginary line from  $s_0 v$  to  $v$ . Hence, we again need to apply Lemma A.2. Choosing suitable constants  $A_0$  and  $A_1$  in (4.11) and (4.12) one may show that

$$\mathbb{E} \operatorname{Im}^q \mathbf{R}_{jj}^{(\mathbb{J})} \leq (CC_0)^q \mathbb{E}^{\frac{1}{2}} |\operatorname{Im} \varepsilon_j^{(\mathbb{J})}| + \operatorname{Im} \Lambda_n^{(\mathbb{J})}|^{2q} + \frac{H_0^q}{2} \Psi^q. \tag{4.13}$$

Applying Lemmas A.7 and A.6 we obtain

$$\mathbb{E} |\operatorname{Im} \varepsilon_j^{(\mathbb{J})}|^{2q} \leq \frac{C^q q^q}{(nv)^q} \mathbb{E} \operatorname{Im}^q m_n^{(\mathbb{J},j)}(z) + \frac{C^q q^{2q}}{(nv)^{2q}}$$

which may be rewritten as follows

$$\mathbb{E} |\operatorname{Im} \varepsilon_j^{(\mathbb{J})}|^{2q} \leq \frac{(Cs_0)^{2q} q^q H^q}{(nv)^q} \Psi^q(z) + \frac{C^q q^{2q}}{(nv)^{2q}}. \tag{4.14}$$

To estimate  $\mathbb{E} |\operatorname{Im} \Lambda_n^{(\mathbb{J})}|^q$  we may proceed as in the proof of Theorem 1.1. We will apply Theorem 2.1 (one has to replace in the definition of (2.12) the maximum over  $|\mathbb{J}| \leq 1$  by the maximum over  $|\mathbb{J}| \leq L$ ) and assumption (4.10). Hence,

$$\mathbb{E} |\operatorname{Im} \Lambda_n^{(\mathbb{J})}|^{2q} \leq \frac{(Cs_0)^{2q} q^q H^q}{(nv)^q} \Psi^q(z) + \frac{C^q q^{2q}}{(nv)^{2q}}. \tag{4.15}$$

Combining the estimates (4.14) and (4.15), we may choose constants  $A_0$  and  $A_1$  (correcting the previous choice if needed) such that

$$(CC_0)^q \mathbb{E}^{\frac{1}{2}} |\operatorname{Im} \varepsilon_j^{(\mathbb{J})}| \leq \frac{H_0^q}{2} \Psi^q(z).$$

The last two inequalities and (4.13) together imply the desired bound

$$\mathbb{E} \operatorname{Im}^q \mathbf{R}_{jj}^{(\mathbb{J})} \leq H_0^q \Psi^q. \tag{□}$$

**Proof of Lemma 4.2.** Let us take any  $u_0 > 0$  and any  $\hat{v} \geq 2 + u_0$ ,  $|u| \leq u_0$ . Furthermore we fix an arbitrary  $\mathbb{J} \subset \mathbb{T}$ . We claim that

$$\operatorname{Im} s(u + i\hat{v}) \geq \frac{1}{2} \operatorname{Im} \mathbf{R}_{jj}^{(\mathbb{J})}(u + i\hat{v}). \tag{4.16}$$

Indeed, note first that for all  $u$  (and  $|u| \leq u_0$  as well)

$$\operatorname{Im} \mathbf{R}_{jj}^{(\mathbb{J})}(u + i\hat{v}) \leq \frac{1}{\hat{v}}. \tag{4.17}$$

For all  $|u| \leq u_0$  and  $|x| \leq 2$ , we obtain

$$\frac{\hat{v}}{(x - u)^2 + \hat{v}^2} \geq \frac{\hat{v}}{(2 + u_0)^2 + \hat{v}^2} \geq \frac{1}{2\hat{v}}.$$

It follows from the last inequality that

$$\operatorname{Im} s(u + i\hat{v}) = \frac{1}{2\pi} \int_{-2}^2 \frac{\hat{v}}{(u - x)^2 + \hat{v}^2} \sqrt{4 - x^2} dx \geq \frac{1}{2\hat{v}}. \tag{4.18}$$

Comparing (4.17) and (4.18), we arrive at (4.16).

We now take  $v \geq \max(\hat{v}, V)$ . Let  $H_0$  be some large constant,  $H_0 \geq \max(C', C'')$ . We choose  $s_0, A_0$  and  $A_1$  as in the previous Lemma 4.6 obtaining

$$\max_{\mathbb{J}: |\mathbb{J}| \leq L} \max_{j \in \mathbb{T}_{\mathbb{J}}} \mathbb{E} \operatorname{Im}^q \mathbf{R}_{jj}^{(\mathbb{J})}(z) \leq H_0^q \Psi^q(z)$$

with  $L = \lceil -\log_{s_0} \tilde{v}_0 \rceil + 1$ . We may now proceed recursively in  $L$  steps and arrive at

$$\max_{j \in \mathbb{T}} \mathbb{E} \operatorname{Im}^q \mathbf{R}_{jj}(z) \leq H_0^q \Psi^q(z)$$

for  $v \geq \tilde{v}_0$  and  $1 \leq q \leq A_1 n v$ . □

## 5. Delocalization of eigenvectors

In this section, we prove Theorem 1.3. The ideas of the proof are similar to [19], Theorem 1.4, but for completeness we provide the details below. Note that the proof is essentially based on Lemma 3.1.

**Proof of Theorem 1.3.** Let us introduce the following distribution function

$$F_{nj}(x) := \sum_{k=1}^n |u_{jk}|^2 \mathbb{1}[\lambda_k(\mathbf{W}) \leq x].$$

Using the eigenvalue decomposition of  $\mathbf{W}$ , it is easy to see that

$$\mathbf{R}_{jj}(z) = \sum_{k=1}^n \frac{|u_{jk}|^2}{\lambda_k(\mathbf{W}) - z} = \int_{-\infty}^{\infty} \frac{1}{x - z} dF_{nj}(x),$$

which means that  $\mathbf{R}_{jj}(z)$  is the Stieltjes transform of  $F_{nj}(x)$ . For any  $\lambda > 0$ , we have

$$\max_{1 \leq k \leq n} |u_{jk}|^2 \leq \sup_x (F_{nj}(x + \lambda) - F_{nj}(x)) \leq 2 \sup_u \lambda \operatorname{Im} \mathbf{R}_{jj}(u + i\lambda). \tag{5.1}$$

To finish the proof, we need to show that with high probability the r.h.s. of (5.1) is bounded by  $n^{-1} \log n$ . Let us recall the following notations. We chose an arbitrary  $0 < \phi' < \frac{1}{4}$ . Let  $\hat{X}_{jk} := X_{jk} \mathbb{1}[|X_{jk}| \leq Dn^{\frac{1}{2} - \phi'}]$ ,  $\tilde{X}_{jk} := X_{jk} \mathbb{1}[|X_{jk}| \leq Dn^{\frac{1}{2} - \phi'}] - \mathbb{E} X_{jk} \mathbb{1}[|X_{jk}| \leq Dn^{\frac{1}{2} - \phi'}]$  and finally  $\check{X}_{jk} := \tilde{X}_{jk} \sigma^{-1}$ , where  $\sigma^2 := \mathbb{E} |\check{X}_{11}|^2$ . Let  $\hat{\mathbf{X}}, \tilde{\mathbf{X}}$  and  $\check{\mathbf{X}}$  denote symmetric random matrices with entries  $\hat{X}_{jk}, \tilde{X}_{jk}$  and  $\check{X}_{jk}$ , respectively. Similarly we denote the resolvent matrices by  $\hat{\mathbf{R}}, \tilde{\mathbf{R}}$  and  $\check{\mathbf{R}}$ . In this case we have

$$\mathbb{P}(\mathbf{W} \neq \hat{\mathbf{W}}) \leq \frac{C}{n^{2-\phi}},$$

where  $\phi := 8\phi'$ . Let  $u_0 > 0$  denote a large constant, whose exact value will be chosen later. Applying [19], Lemmas A.1, A.2, it follows that

$$\mathbb{P}(\|\mathbf{W}\| \geq u_0) \leq \frac{C}{n^{2-\phi}}.$$

In what follows, we may assume that  $\|\mathbf{W}\| \leq u_0$  and  $\mathbf{W} = \hat{\mathbf{W}}$ . Then for  $|u| \geq 2u_0$  and  $v > 0$  we get

$$|\mathbf{R}_{jj}(u + iv)| \leq \int_{-u_0}^{u_0} \frac{1}{\sqrt{(x - u)^2 + v^2}} dF_{nj}(x) \leq \frac{1}{u_0} \leq C,$$

where  $C$  is some large positive constant which will be chosen later. It remains to estimate  $|\mathbf{R}_{jj}(u + iv)|$  for all  $-2u_0 \leq u \leq 2u_0$ . Denote this interval by  $\mathcal{U}_0 := [-2u_0, 2u_0]$ . By the triangular inequality, we may write  $|\mathbf{R}_{jj}| = |\hat{\mathbf{R}}_{jj}| \leq |\tilde{\mathbf{R}}_{jj}| + |\hat{\mathbf{R}}_{jj} - \tilde{\mathbf{R}}_{jj}|$ . Using the simple identity

$$\hat{\mathbf{R}}_{jj} - \tilde{\mathbf{R}}_{jj} = [\hat{\mathbf{R}}(\hat{\mathbf{W}} - \tilde{\mathbf{W}})\tilde{\mathbf{R}}]_{jj}$$

we get

$$|\hat{\mathbf{R}}_{jj} - \tilde{\mathbf{R}}_{jj}| \leq \|\hat{\mathbf{W}} - \tilde{\mathbf{W}}\| \|\mathbf{e}_j^T \hat{\mathbf{R}}\|_2 \|\tilde{\mathbf{R}} \mathbf{e}_j\|_2,$$

where  $\mathbf{e}_j$  is a unit column-vector with all entries zero except for an entry one at the position  $j$ . Using Lemma A.3 in the Appendix, we conclude that

$$|\hat{\mathbf{R}}_{jj}| \leq |\tilde{\mathbf{R}}_{jj}| + \frac{1}{v} \|\hat{\mathbf{W}} - \tilde{\mathbf{W}}\| \sqrt{|\hat{\mathbf{R}}_{jj}| |\tilde{\mathbf{R}}_{jj}|}.$$

It is easy to see that

$$\|\widehat{\mathbf{W}} - \widetilde{\mathbf{W}}\|_2^2 = \frac{1}{n} \sum_{j,k} [\mathbb{E} |X_{jk}| \mathbb{1}[|X_{jk}| \geq Dn^{\frac{1}{2}-\phi}]]^2 \leq \frac{C}{n^4}.$$

We may take  $v = v_0 := C_1 n^{-1} \log n$ , with  $C_1 \geq A_0$ . Applying the inequality  $2|ab| \leq a^2 + b^2$ , we get

$$\sup_{u \in \mathcal{U}_0} |\mathbf{R}_{jj}| \leq 3 \sup_{u \in \mathcal{U}_0} |\widetilde{\mathbf{R}}_{jj}|. \tag{5.2}$$

It remains to estimate  $\sup_{u \in \mathcal{U}_0} |\widetilde{\mathbf{R}}_{jj}(u + iv_0)|$ . It is easy to see that

$$\widetilde{\mathbf{R}}(z) = (\widetilde{\mathbf{W}} - z\mathbf{I})^{-1} = \sigma^{-1}(\check{\mathbf{W}} - z\sigma^{-1}\mathbf{I})^{-1} = \sigma^{-1}\check{\mathbf{R}}(\sigma^{-1}z). \tag{5.3}$$

Applying the resolvent identity, we get

$$\check{\mathbf{R}}(z) - \check{\mathbf{R}}(\sigma^{-1}z) = (z - \sigma^{-1}z)\check{\mathbf{R}}(z)\check{\mathbf{R}}(\sigma^{-1}z). \tag{5.4}$$

Combining (5.3) and (5.4), we obtain

$$|\widetilde{\mathbf{R}}_{jj}(z) - \check{\mathbf{R}}_{jj}(z)| \leq (\sigma^{-1} - 1)|\check{\mathbf{R}}_{jj}(\sigma^{-1}z)| + \frac{|z|(\sigma^{-1} - 1)}{v} \sqrt{|\check{\mathbf{R}}_{jj}(z)||\check{\mathbf{R}}_{jj}(\sigma^{-1}z)|}.$$

It is easy to check that  $(\sigma^{-1} - 1) \leq Cn^{-\frac{3}{2}}$  and  $\max(|z\check{\mathbf{R}}_{jj}(z)|, |z\check{\mathbf{R}}_{jj}(\sigma^{-1}z)|) \leq C$  for some constant  $C$ . Similarly to the previous calculations, we get that

$$\sup_{u \in \mathcal{U}_0} |\check{\mathbf{R}}_{jj}| \leq 3 \sup_{u \in \mathcal{U}_0} |\check{\check{\mathbf{R}}}_{jj}|. \tag{5.5}$$

Note, that the matrix  $\check{\check{\mathbf{W}}}$  satisfies the conditions **(C1)**. Applying Lemma 3.1 with  $p = c \log n$ , we obtain

$$\mathbb{P}(|\check{\check{\mathbf{R}}}_{jj}(u + iv_0)| \geq C_0 e^{\frac{6}{c}}) \leq \frac{\mathbb{E} |\check{\check{\mathbf{R}}}_{jj}(u + iv_0)|^p}{(C_0 e^{\frac{6}{c}})^p} \leq \frac{1}{n^6}.$$

We partition the interval  $\mathcal{U}_0$  into  $k_n := n^4$  disjoint sub-intervals of equal length, that is,  $-2u_0 = x_0 \leq x_1 \leq \dots \leq x_{k_n} = 2u_0$ . Then the Newton–Leibniz formula implies

$$\begin{aligned} \sup_{u \in \mathcal{U}_0} |\check{\check{\mathbf{R}}}_{jj}(u + iv_0)| &\leq \max_{1 \leq k \leq k_n} \sup_{x_{k-1} \leq x \leq x_k} |\check{\check{\mathbf{R}}}_{jj}(x + iv_0)| \\ &\leq \max_{1 \leq k \leq k_n} |\check{\check{\mathbf{R}}}_{jj}(x_{k-1} + iv_0)| + \max_{1 \leq k \leq k_n} \int_{x_{k-1}}^{x_k} |\check{\check{\mathbf{R}}}'_{jj}(u + iv_0)| du. \end{aligned}$$

We may write

$$\max_{1 \leq k \leq k_n} \int_{x_{k-1}}^{x_k} |\check{\check{\mathbf{R}}}'_{jj}(u + iv_0)| du \leq \frac{C}{n^{2-\phi}}.$$

Thus we arrive at

$$\mathbb{P}\left(\sup_{u \in \mathcal{L}_0} |\check{\mathbf{R}}_{jj}(u + iv_0)| \geq 2C_0 e^{\frac{6}{c}}\right) \leq \sum_{k=1}^{k_n} \mathbb{P}\left(|\check{\mathbf{R}}_{jj}(x_{k-1} + iv_0)| \geq C_0 e^{\frac{6}{c}}\right) \leq \frac{C}{n^2}. \tag{5.6}$$

We choose now  $\lambda := v_0$ . In view of (5.1), (5.2), (5.5) and (5.6), we get that there exist  $C$  and  $C_1$  such that

$$\mathbb{P}\left(\max_{1 \leq j, k \leq n} |u_{jk}|^2 \leq \frac{C_1 \log n}{n}\right) \geq 1 - \frac{C}{n^{2-\phi}},$$

which concludes the proof. □

### 6. Rate of convergence to the semicircle law

In this section, we prove Theorem 1.4 and Theorem 1.6. We estimate the difference between  $F_n$  (resp.  $\mathbb{E} F_n$ ) and  $G_{sc}$  in the Kolmogorov metric via the distance between the corresponding Stieltjes transforms. For this purpose, we formulate the following smoothing inequality proved in [22], Corollary 2.3, which allows to relate distribution functions to their Stieltjes transforms. For all  $x \in [-2, 2]$  let us define  $\gamma(x) := 2 - |x|$ . Given  $\frac{1}{2} > \varepsilon > 0$  we introduce the following intervals  $\mathbb{J}_\varepsilon := \{x \in [-2, 2] : \gamma(x) \geq \varepsilon\}$  and  $\mathbb{J}'_\varepsilon := \mathbb{J}_{\varepsilon/2}$ .

**Lemma 6.1.** *Let  $v_0 > 0$  and  $\frac{1}{2} > \varepsilon > 0$  be positive numbers such that*

$$2(\sqrt{2} + 1)v_0 \leq \varepsilon^{\frac{3}{2}}.$$

*Assume that  $F$  is an arbitrary distribution function with the Stieltjes transform  $f(z)$ . Then for any  $V > 0$  and  $v' := v'(x) := v_0/\sqrt{\gamma(x)}, x \in \mathbb{J}'_\varepsilon$ , there exist absolute positive constants  $C_1, C_2$  and  $C_3$  such that the following inequality holds*

$$\begin{aligned} \Delta(F, G) \leq & C_1 \int_{-\infty}^{\infty} |f(u + iV) - s(u + iV)| du + C_2 v_0 + C_3 \varepsilon^{\frac{3}{2}} \\ & + 2 \sup_{x \in \mathbb{J}'_\varepsilon} \left| \int_{v'}^V (f(x + iv) - s(x + iv)) dv \right|. \end{aligned}$$

**Proof.** See [22], Corollary 2.3, or [26], Proposition 2.1. □

In what follows, we will need the following version of this lemma.

**Corollary 6.2.** *Assuming the conditions of Lemma 6.1, we have*

$$\begin{aligned} \mathbb{E}^{\frac{1}{p}} [\Delta_n^*]^p \leq & C_1 \int_{-\infty}^{\infty} \mathbb{E}^{\frac{1}{p}} |m_n(u + iV) - s(u + iV)|^p du + C_2 v_0 + C_3 \varepsilon^{\frac{3}{2}} \\ & + C_1 \mathbb{E}^{\frac{1}{p}} \sup_{x \in \mathbb{J}'_\varepsilon} \left| \int_{v'}^V (m_n(x + iv) - s(x + iv)) dv \right|^p. \end{aligned} \tag{6.1}$$

**Proof.** The proof is the direct consequence of the previous lemma and we omit it. For details, the interested reader is referred to [26], Corollary 2.1.  $\square$

**Proof of Theorem 1.4.** We start from the part (i). We proceed as in the proof of Theorem 1.1 in [26]. We choose in Corollary 6.2 the following values for the parameters  $v_0, \varepsilon$  and  $V$ . Let us take  $v_0 := A_0 n^{-1} \log n$ ,  $\varepsilon := (2v_0 a)^{\frac{2}{3}}$  and  $V := 4$ . We may partition  $\mathbb{J}'_\varepsilon$  into  $k_n := n^4$  disjoint subintervals of equal length. Let us denote the endpoints of these intervals by  $x_k, k = 0, \dots, k_n$ . We get  $-2 + \varepsilon = x_0 < x_1 < \dots < x_{k_n} = 2 - \varepsilon$ . For simplicity, we denote  $\Lambda_n(u + iv) := m_n(u + iv) - s(u + iv)$  but we will not omit the argument. We start to estimate the second integral in the r.h.s. of (6.1). It is easy to see that

$$\sup_{x \in \mathbb{J}'_\varepsilon} \left| \int_{v'}^V \Lambda_n(x + iv) dv \right| \leq \max_{1 \leq k \leq k_n} \sup_{x_{k-1} \leq x \leq x_k} \left| \int_{v'}^V \Lambda_n(x + iv) dv \right|. \tag{6.2}$$

Applying the Newton–Leibniz formula, we may write

$$\begin{aligned} \sup_{x_{k-1} \leq x \leq x_k} \left| \int_{v'}^V \Lambda_n(x + iv) dv \right| &\leq \left| \int_{v'}^V \Lambda_n(x_{k-1} + iv) dv \right| \\ &+ \int_{x_{k-1}}^{x_k} \int_{v'}^V |\Lambda'_n(x + iv)| dv dx. \end{aligned} \tag{6.3}$$

It follows from Cauchy’s integral formula that for all  $z = x + iv$  with  $v \geq v_0$  we have

$$|\Lambda'_n(x + iv)| \leq \frac{C}{v^2} \leq Cn^2. \tag{6.4}$$

We may conclude from (6.3) and (6.4) that

$$\sup_{x_{k-1} \leq x \leq x_k} \left| \int_{v'}^V \Lambda_n(x + iv) dv \right| \leq \left| \int_{v'}^V \Lambda_n(x_{k-1} + iv) dv \right| + \frac{C}{n}.$$

Applying this inequality to (6.2) together with expectations we obtain

$$\begin{aligned} \mathbb{E} \sup_{x \in \mathbb{J}'_\varepsilon} \left| \int_{v'}^V \Lambda_n(x + iv) dv \right|^p &\leq \mathbb{E} \max_{1 \leq k \leq k_n} \left| \int_{v'}^V \Lambda_n(x_{k-1} + iv) dv \right|^p + \frac{C^p}{n^p} \\ &\leq \sum_{k=1}^{k_n} \left| \int_{v'}^V \mathbb{E}^{\frac{1}{p}} |\Lambda_n(x_{k-1} + iv)|^p dv \right|^p + \frac{C^p}{n^p}. \end{aligned} \tag{6.5}$$

Since  $x \in \mathbb{J}'_\varepsilon$  it follows from Theorem 1.1 that

$$\mathbb{E} |\Lambda_n(x + iv)|^p \leq \left( \frac{Cp}{nv} \right)^p. \tag{6.6}$$

Choosing  $p = A_1(nv_0)^{\frac{1-2\alpha}{2}} = c \log n$  we finally get from (6.5) and (6.6) that

$$\mathbb{E}^{\frac{1}{p}} \sup_{x \in \mathbb{J}'_\varepsilon} \left| \int_{v'}^V \Lambda_n(x + iv) dv \right|^p \leq \frac{Ck_n^{\frac{1}{p}} \log^2 n}{n} + \frac{C}{n} \leq \frac{C \log^2 n}{n}. \tag{6.7}$$

It remains to estimate the first of the integrals in (6.1). It was proved in [19], Inequality 2.8, that

$$\mathbb{E}^{\frac{1}{p}} |\Lambda_n(u + iV)|^p \leq \frac{Cp|s(z)|^{\frac{p+1}{p}}}{n}, \tag{6.8}$$

which holds for all  $z = u + iV, u \in \mathbb{R}$ . Hence,

$$\int_{-\infty}^{\infty} \mathbb{E}^{\frac{1}{p}} |\Lambda_n(u + iV)|^p du \leq \frac{Cp}{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{du dG_{sc}(x)}{((x-u)^2 + V^2)^{\frac{p+1}{p}}} \leq \frac{C \log^2 n}{n}. \tag{6.9}$$

Combining now (6.1), (6.7) and (6.9), we get

$$\mathbb{E}^{\frac{1}{p}} [\Delta_n^*]^p \leq \frac{C \log^2 n}{n}.$$

Since  $\mathbb{E}^{\frac{1}{p}} [\Delta_n^*]^p$  is non-decreasing function of  $p$ , the last inequality remains valid for all  $1 \leq p \leq c \log n$ . To finish the proof of Theorem 1.4, it remains to apply Markov's inequality

$$\mathbb{P}(\Delta_n^* \geq K) \leq \frac{\mathbb{E}[\Delta_n^*]^p}{K^p} \leq \frac{C^p \log^{2p} n}{K^p n^p}. \tag{6.10}$$

**Proof of Theorem 1.6.** Applying Lemma 6.1 with  $F := \mathbb{E} F_n$ , we get

$$\begin{aligned} \Delta_n \leq & 2 \int_{-\infty}^{\infty} |\mathbb{E} m_n(u + iV) - s(u + iV)| du + C_1 v_0 + C_2 \varepsilon^{\frac{3}{2}} \\ & + 2 \sup_{x \in \mathbb{J}'_\varepsilon} \left| \int_{v'}^V (\mathbb{E} m_n(x + iv) - s(x + iv)) dv \right|. \end{aligned} \tag{6.10}$$

Let us take  $v_0 := A_0 n^{-1}, \varepsilon := (2v_0 a)^{\frac{2}{3}}$  and  $V := 4$ . As before, we denote  $\Lambda_n(u + iv) := m_n(u + iv) - s(u + iv)$ . The bound for the first integral follows from [28], Inequality 3.11

$$|\mathbb{E} \Lambda_n(u + iV)| \leq \frac{C|s(z)|^2}{n}.$$

This bound gives

$$\int_{-\infty}^{\infty} |\mathbb{E} \Lambda_n(u + iV)| du \leq \frac{C}{n}. \tag{6.11}$$

To estimate the second integral we shall use the same arguments as in the proof of Lemma 3.5. We denote by  $Y_{jk}, 1 \leq j \leq k \leq n$  a triangular set of random variables such that  $|Y_{jk}| \leq D$ , for some  $D$  chosen later, and

$$\mathbb{E} X_{jk}^s = \mathbb{E} Y_{jk}^s \quad \text{for } s = 1, \dots, 4.$$

By Lemma 3.4 these random variables exist. Let us denote  $\mathbf{W}^y := \frac{1}{\sqrt{n}}\mathbf{Y}, \mathbf{R}^y := (\mathbf{W}^y - z\mathbf{I})^{-1}$  and  $m_n^y(z) := \frac{1}{n} \text{Tr} \mathbf{R}^y(z)$ . We will show below that for all  $j, j = 1, \dots, n$ , there exists  $\phi > 0$  such that

$$|\mathbb{E} \mathbf{R}_{jj}(z) - \mathbb{E} \mathbf{R}_{jj}^y(z)| \leq \frac{C}{n^{1+\phi\nu}} \tag{6.12}$$

and, hence,

$$|\mathbb{E} m_n(z) - \mathbb{E} m_n^y(z)| \leq \frac{C}{n^{1+\phi\nu}}. \tag{6.13}$$

It follows from [28], Theorem 1.3, that

$$|\mathbb{E} m_n^y(z) - s(z)| \leq \frac{C}{nv^{\frac{3}{4}}} + \frac{C}{n^{\frac{3}{2}}v^{\frac{3}{2}}|z^2 - 4|^{\frac{1}{4}}}. \tag{6.14}$$

Inequalities (6.14) and (6.13) together imply that

$$|\mathbb{E} m_n(z) - s(z)| \leq \frac{C}{nv^{\frac{3}{4}}} + \frac{C}{n^{1+\phi\nu}} + \frac{C}{n^{\frac{3}{2}}v^{\frac{3}{2}}|z^2 - 4|^{\frac{1}{4}}}.$$

Calculating the second integral in (6.10), we get

$$\Delta_n \leq \frac{C}{n}.$$

It remains to prove (6.12). For every  $j, j = 1, \dots, n$ , we do  $\frac{n(n+1)}{2}$  replacements and corresponding exchanges of  $\mathbf{R}_{jj}$  by  $\mathbf{R}_{jj}^y$ . Let  $\mathbb{J}, \mathbb{K} \subset \mathbb{T}$ . We denote by  $\mathbf{W}^{(\mathbb{J}, \mathbb{K})}$  a random matrix  $\mathbf{W}$  with entries  $\frac{1}{\sqrt{n}}Y_{\mu\nu}$  in the positions  $(\mu, \nu), \mu \in \mathbb{J}, \nu \in \mathbb{K}$ . Assume that we have already replaced the entries in positions  $(\mu, \nu), \mu \in \mathbb{T}, \nu \in \mathbb{K}$  and want to replace in addition the entry in position  $(a, b), a \in \mathbb{T} \setminus \mathbb{J}, b \in \mathbb{T} \setminus \mathbb{K}$ . Without loss of generality, we may assume that  $\mathbb{J} = \emptyset, \mathbb{K} = \emptyset$  (hence  $\mathbf{W}^{(\mathbb{J}, \mathbb{K})} = \mathbf{W}$ ) and denote by  $\mathbf{V} := \mathbf{W}^{\{(a), \{b\}}$ . Introduce

$$\mathbf{E}^{(a,b)} = \begin{cases} \mathbf{e}_a \mathbf{e}_b^\top + \mathbf{e}_b \mathbf{e}_a^\top, & 1 \leq a < b \leq n, \\ \mathbf{e}_a \mathbf{e}_a^\top, & a = b. \end{cases}$$

and  $\mathbf{U} := \mathbf{W} - \frac{X_{ab}}{\sqrt{n}}\mathbf{E}^{(a,b)}$ , where  $\mathbf{e}_j$  denotes a unit column-vector with all entries zeros except in the  $j$ th position. Using these notations, we may write

$$\mathbf{W} = \mathbf{U} + \frac{1}{\sqrt{n}}X_{ab}\mathbf{E}^{(a,b)}, \quad \mathbf{V} = \mathbf{U} + \frac{1}{\sqrt{n}}Y_{ab}\mathbf{E}^{(a,b)}.$$

Recall that  $\mathbf{R} := (\mathbf{W} - z\mathbf{I})^{-1}$  and denote  $\mathbf{S} := (\mathbf{V} - z\mathbf{I})^{-1}$  and  $\mathbf{T} := (\mathbf{U} - z\mathbf{I})^{-1}$ . Applying (3.16) and (3.17) we get

$$\begin{aligned} \mathbb{E} \mathbf{R}_{jj} - \mathbb{E} \mathbf{S}_{jj} &= \sum_{\mu=5}^m \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E}[X_{ab}^\mu - Y_{ab}^\mu] \mathbb{E}[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj} \\ &\quad + \frac{(-1)^{m+1}}{n^{\frac{m+1}{2}}} \mathbb{E} X_{ab}^{m+1} [(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jj} \\ &\quad - \frac{(-1)^{m+1}}{n^{\frac{m+1}{2}}} \mathbb{E} Y_{ab}^{m+1} [(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{S}]_{jj}. \end{aligned} \tag{6.15}$$

Without loss of generality we may assume that  $|X_{jk}| \leq Dn^{\frac{1}{2}-\phi}$  for some  $\phi > 0$  depending on  $\alpha$ . Choosing  $m$  and applying Lemma 3.2 we estimate the third term in (6.15) as follows

$$\frac{1}{n^{\frac{m+1}{2}}} \mathbb{E} |X_{ab}|^{m+1} |[(\mathbf{TE}^{(a,b)})^{m+1} \mathbf{R}]_{jj}| \leq \frac{C}{n^{3+\phi\nu}}.$$

The same bound obviously holds for the fourth term in (6.15). Let us consider now the first term and investigate the part corresponding to  $X_{ab}$  (the same estimates are valid for the part corresponding to  $Y_{ab}$ )

$$\sum_{\mu=5}^m \frac{(-1)^\mu}{n^{\frac{\mu}{2}}} \mathbb{E}[X_{ab}^\mu] \mathbb{E}[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj}.$$

It is straightforward to check that  $[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj}$  is the sum of  $2^\nu$  terms of the following type

$$\mathbf{T}_{j i_1} \mathbf{T}_{i_1 i_2} \cdots \mathbf{T}_{i_m i_m} \mathbf{T}_{i_m j},$$

where  $i_l = a$  or  $i_l = b$  for  $l = 1, \dots, m$ . Assume that  $a$  and  $b$  are not equal to  $j$ . Then, the first and the last terms in the last product are off diagonal entries of the resolvent  $\mathbf{T}$ . Applying Hölder’s inequality, Lemma A.8 and Lemma 3.2, we get

$$\sum_{\mu=5}^m \frac{1}{n^{\frac{\mu}{2}}} \mathbb{E} |X_{ab}|^\mu \mathbb{E} |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj}| \leq \frac{C}{n^{3+\phi\nu}}.$$

Assume now that  $a$  or (and)  $b$  are equal to  $j$ . The number of configurations of this type is of order  $n$ . Similarly to the previous inequality, we thus obtain the bound

$$\sum_{\mu=5}^m \frac{1}{n^{\frac{\mu}{2}}} \mathbb{E} |X_{ab}|^\mu \mathbb{E} |[(\mathbf{TE}^{(a,b)})^\mu \mathbf{T}]_{jj}| \leq \frac{C}{n^{2+\phi}}.$$

Repeating the same steps for all  $\frac{n(n+1)}{2}$  pairs of  $(a, b)$  we arrive at (6.12) and (6.13) respectively. □

## Appendix: Auxiliary lemmas

### A.1. Inequalities for resolvent matrices

In this section, we collect some inequalities for the resolvent of the matrix  $\mathbf{W}$ .

**Lemma A.1.** *For any  $z = u + iv \in \mathbb{C}^+$  we have for any  $s \geq 1$*

$$|\mathbf{R}_{jj}^{(\mathbb{J})}(u + iv/s)| \leq s |\mathbf{R}_{jj}^{(\mathbb{J})}(u + iv)|$$

and

$$\frac{1}{|u + iv/s_0 + m_n^{(\mathbb{J})}(u + iv/s_0)|} \leq \frac{s_0}{|u + iv + m_n^{(\mathbb{J})}(u + iv)|}. \tag{A.1}$$

**Proof.** See [18], Lemma C.1. □

**Lemma A.2.** *Let  $g(v) := g(u + iv)$  be the Stieltjes transform of some distribution function  $G(x)$ . Then for any  $s \geq 1$*

$$\operatorname{Im} g(v/s) \leq s \operatorname{Im} g(v) \quad \text{and} \quad \operatorname{Im} g(v) \leq s \operatorname{Im} g(v/s).$$

**Proof.** See [18], Lemma C.2. □

**Lemma A.3.** *For any  $z = u + iv \in \mathbb{C}^+$  we have*

$$\frac{1}{n} \sum_{l,k \in \mathbb{T}_{\mathbb{J}}} |\mathbf{R}_{kl}^{(\mathbb{J})}|^2 \leq \frac{1}{v} \operatorname{Im} m_n^{(\mathbb{J})}(z). \tag{A.2}$$

For any  $l \in \mathbb{T}_{\mathbb{J}}$

$$\sum_{k \in \mathbb{T}_{\mathbb{J}}} |\mathbf{R}_{kl}^{(\mathbb{J})}|^2 \leq \frac{1}{v} \operatorname{Im} \mathbf{R}_{ll}^{(\mathbb{J})}. \tag{A.3}$$

**Proof.** See [18], Lemma C.4. □

### A.2. Moment inequalities for linear and quadratic forms in the sub-Gaussian case

In this subsection, we estimate the moments of  $\varepsilon_{\nu j}^{(\mathbb{J})}$  for  $\nu = 1, \dots, 4$  (recall the definition (4.2)) in the sub-Gaussian case. It is well known that the random variables  $\xi$  is sub-Gaussian if and only if  $\mathbb{E} |\xi|^p = O(p^{\frac{p}{2}})$  as  $p \rightarrow \infty$ . We define the sub-Gaussian norm of  $\xi$  as

$$\|\xi\|_{\psi_2} := \sup_{p \geq 1} p^{-\frac{1}{2}} \mathbb{E}^{\frac{1}{p}} |\xi|^p.$$

We may conclude that  $\mathbb{E}|\varepsilon_{1j}^{(\mathbb{J})}|^p \leq Cp^{\frac{p}{2}}n^{-\frac{p}{2}}$ . The following lemma is the Hanson–Wright inequality for quadratic forms, see [30]. The following improved version is due to [34].

**Lemma A.4 (Hanson–Wright inequality).** *Let  $X = (X_1, \dots, X_n) \in \mathbb{R}^n$  be a random vector with independent components  $X_i$  which satisfy  $\mathbb{E} X_i = 0$  and  $\|X\|_{\psi_2} \leq K$ . Let  $\mathbf{A} = [a_{jk}]_{j,k=1}^n$  be an  $n \times n$  matrix. Then, for every  $t \geq 0$*

$$\mathbb{P}\left(\left|\sum_{j,k=1}^n a_{jk} X_j X_k - \sum_{j=1}^n a_{jj} \mathbb{E} X_{jj}^2\right| \geq t\right) \leq 2 \exp\left[-c \min\left(\frac{t^2}{K^4 \|A\|_2^2}, \frac{t}{K^2 \|A\|}\right)\right].$$

**Proof.** See [34], Theorem 1.1. □

A direct consequence of this lemma is the following result.

**Lemma A.5.** *Assume that  $X_1, \dots, X_n$  are i.i.d. sub-Gaussian random variables which satisfy  $\mathbb{E} X_i = 0$  and  $\|X\|_{\psi_2} \leq K$ . Then for all  $p \geq 2$  there exists a positive constant  $C$  depending on  $K$  such that*

$$\mathbb{E}|\varepsilon_{2j}^{(\mathbb{J})} + \varepsilon_{3j}^{(\mathbb{J})}|^p \leq \frac{C^p p^{\frac{p}{2}}}{(nv)^{\frac{p}{2}}} \mathbb{E} \operatorname{Im}^{\frac{p}{2}} m_n^{(\mathbb{J},j)}(z) + \frac{C^p p^p}{(nv)^p}.$$

**Proof.** Applying  $\mathbb{E}|\xi|^p = p \int_0^\infty t^{p-1} \mathbb{P}(|\xi| \geq t) dt$  and Lemma A.4, we get

$$\mathbb{E}|\varepsilon_{2j}^{(\mathbb{J})} + \varepsilon_{3j}^{(\mathbb{J})}|^p \leq \frac{2p}{n^p} \int_0^\infty t^{p-1} 2 \exp\left[-c \min\left(\frac{t^2}{K^4 \|\mathbf{R}^{(\mathbb{J},j)}\|_2^2}, \frac{t}{K^2 \|\mathbf{R}^{(\mathbb{J},j)}\|}\right)\right] dt.$$

We may split the last integral in two integrals over the regions  $[0, \|\mathbf{R}^{(\mathbb{J},j)}\|_2^2 \|\mathbf{R}^{(\mathbb{J},j)}\|^{-1}]$  and  $[\|\mathbf{R}^{(\mathbb{J},j)}\|_2^2 \|\mathbf{R}^{(\mathbb{J},j)}\|^{-1}, \infty]$  obtaining

$$\begin{aligned} \mathbb{E}|\varepsilon_{2j}^{(\mathbb{J})} + \varepsilon_{3j}^{(\mathbb{J})}|^p &\leq \frac{2p \|\mathbf{R}^{(\mathbb{J},j)}\|_2^p}{n^p} \int_0^\infty t^p e^{-c \frac{t^2}{K^4}} dt + \frac{2p \|\mathbf{R}^{(\mathbb{J},j)}\|^p}{n^p} \int_0^\infty t^p e^{-c \frac{t}{K^2}} dt \\ &\leq \frac{C^p p^{\frac{p}{2}}}{(nv)^{\frac{p}{2}}} \mathbb{E} \operatorname{Im}^{\frac{p}{2}} m_n^{(\mathbb{J},j)}(z) + \frac{C^p p^p}{(nv)^p}, \end{aligned}$$

where in the second inequality we applied Lemma A.3 and  $\|\mathbf{R}^{(\mathbb{J},j)}\|^p \leq v^{-p}$ . □

It remains to estimate  $\mathbb{E}|\varepsilon_{4j}^{(\mathbb{J})}|^p$ . The bound for this term is distribution-independent and given in the following lemma.

**Lemma A.6.** *For  $p \geq 2$ , we have*

$$\mathbb{E}|\varepsilon_{4j}^{(\mathbb{J})}|^p \leq \frac{1}{(nv)^p}.$$

**Proof.** See [18], Lemma A.9. □

In order to bound  $\mathbb{E} \operatorname{Im}^p \mathbf{R}_{jj}$  we also need the following lemma.

**Lemma A.7.** *Assume that  $X_1, \dots, X_n$  are i.i.d. sub-Gaussian random variables which satisfy  $\mathbb{E} X_i = 0$  and  $\|X\|_{\psi_2} \leq K$ . Then for all  $p \geq 2$  there exists a positive constant  $C$  depending on  $K$  such that*

$$\mathbb{E} |\operatorname{Im} \varepsilon_{2j}^{(\mathbb{J})} + \operatorname{Im} \varepsilon_{3j}^{(\mathbb{J})}|^p \leq \frac{C^p p^{\frac{p}{2}}}{(nv)^{\frac{p}{2}}} \mathbb{E} \operatorname{Im}^{\frac{p}{2}} m_n^{(\mathbb{J},j)}(z) + \frac{C^p p^p}{(nv)^p}.$$

**Proof.** The proof is similar to the proof of the previous Lemma A.5. We omit the details. □

### A.3. Moment inequalities for off-diagonal entries of the resolvent

**Lemma A.8.** *Assume that the conditions (C1) hold. Then for all  $1 \leq j < k \leq n$  and  $1 \leq q \leq 2 + \frac{\delta}{2}$  there exists a positive constant  $C$  such that*

$$\mathbb{E} |\mathbf{R}_{jk}|^q \leq \frac{C}{(nv)^{\frac{q}{2}}}.$$

**Proof.** Consider the following equality

$$\mathbf{R}_{jk} = -\frac{1}{\sqrt{n}} \left( \sum_{l \in \mathbb{T}_j} X_{jl} \mathbf{R}_{lk}^{(j)} \right) \mathbf{R}_{jj}.$$

Applying Hölder’s inequality, we get

$$\mathbb{E} |\mathbf{R}_{jk}|^q \leq n^{-\frac{q}{2}} \mathbb{E}^{\frac{1}{2}} \left| \sum_{l \in \mathbb{T}_j} X_{jl} \mathbf{R}_{lk}^{(j)} \right|^{2q} \mathbb{E}^{\frac{1}{2}} |\mathbf{R}_{jj}|^{2q}.$$

Conditioning, applying Rosenthal’s inequality and Lemma 3.2 we obtain

$$\mathbb{E} |\mathbf{R}_{jk}|^q \leq C^q n^{-\frac{q}{2}} \left( q^{\frac{q}{2}} \mathbb{E}^{\frac{1}{2}} \left( \sum_{l \in \mathbb{T}_j} |\mathbf{R}_{lk}^{(j)}|^2 \right)^q + q^q \mu_{2q}^{\frac{1}{2}} \mathbb{E}^{\frac{1}{2}} \left( \sum_{l \in \mathbb{T}_j} |\mathbf{R}_{lk}^{(j)}|^{2q} \right) \right).$$

This inequality and Lemma A.3 together imply that

$$\mathbb{E} |\mathbf{R}_{jk}|^q \leq C^q q^{\frac{q}{2}} (nv)^{-\frac{q}{2}} + q^q \mu_{2q}^{\frac{1}{2}} (nv)^{-\frac{q}{2}}.$$

Hence, for  $1 \leq q \leq 2 + \frac{\delta}{2}$  we get

$$\mathbb{E} |\mathbf{R}_{jk}|^q \leq C (nv)^{-\frac{q}{2}}. □$$

## Acknowledgements

We would like to thank the Associate Editor and the Reviewers for helpful comments and suggestions.

All authors were supported by CRC 701 “Spectral Structures and Topological Methods in Mathematics”. A. Naumov, A. Tikhomirov and D. Timushev were supported by RFBR N 14-01-00500. A. Tikhomirov and D. Timushev were also supported by Programs of UD RAS, project N 15-16-1-3. A. Naumov was also supported by Hong Kong RGC GRF 403513, RFBR N 16-31-00005 and President’s of Russian Federation Grant for young scientists N 4596.2016.1.

## References

- [1] Anderson, G.W., Guionnet, A. and Zeitouni, O. (2010). *An Introduction to Random Matrices*. Cambridge Studies in Advanced Mathematics **118**. Cambridge: Cambridge Univ. Press. [MR2760897](#)
- [2] Arnold, L. (1967). On the asymptotic distribution of the eigenvalues of random matrices. *J. Math. Anal. Appl.* **20** 262–268. [MR0217833](#)
- [3] Bai, Z., Hu, J., Pan, G. and Zhou, W. (2011). A note on rate of convergence in probability to semicircular law. *Electron. J. Probab.* **16** 2439–2451. [MR2861680](#)
- [4] Bai, Z. and Silverstein, J.W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. *Springer Series in Statistics*. New York: Springer. [MR2567175](#)
- [5] Bai, Z.D. (1993). Convergence rate of expected spectral distributions of large random matrices. I. Wigner matrices. *Ann. Probab.* **21** 625–648. [MR1217559](#)
- [6] Bai, Z.D., Miao, B. and Tsay, J. (2002). Convergence rates of the spectral distributions of large Wigner matrices. *Int. Math. J.* **1** 65–90. [MR1825933](#)
- [7] Bobkov, S.G., Götze, F. and Tikhomirov, A.N. (2010). On concentration of empirical measures and convergence to the semi-circle law. *J. Theoret. Probab.* **23** 792–823. [MR2679957](#)
- [8] Cacciapuoti, C., Maltsev, A. and Schlein, B. (2015). Bounds for the Stieltjes transform and the density of states of Wigner matrices. *Probab. Theory Related Fields* **163** 1–59. [MR3405612](#)
- [9] Erdős, L., Knowles, A., Yau, H.-T. and Yin, J. (2012). Spectral statistics of Erdős–Rényi Graphs II: Eigenvalue spacing and the extreme eigenvalues. *Comm. Math. Phys.* **314** 587–640. [MR2964770](#)
- [10] Erdős, L., Knowles, A., Yau, H.-T. and Yin, J. (2013). The local semicircle law for a general class of random matrices. *Electron. J. Probab.* **18** 58. [MR3068390](#)
- [11] Erdős, L., Knowles, A., Yau, H.-T. and Yin, J. (2013). Spectral statistics of Erdős–Rényi graphs I: Local semicircle law. *Ann. Probab.* **41** 2279–2375. [MR3098073](#)
- [12] Erdős, L., Schlein, B. and Yau, H.-T. (2009). Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.* **287** 641–655. [MR2481753](#)
- [13] Erdős, L., Schlein, B. and Yau, H.-T. (2009). Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.* **37** 815–852. [MR2537522](#)
- [14] Erdős, L., Schlein, B. and Yau, H.-T. (2010). Wegner estimate and level repulsion for Wigner random matrices. *Int. Math. Res. Not. IMRN* **3** 436–479. [MR2587574](#)
- [15] Girko, V. (1985). Spectral theory of random matrices. *Uspekhi Mat. Nauk* **40** 67–106. [MR0783605](#)
- [16] Girko, V. (1998). Convergence rate of the expected spectral functions of symmetric random matrices is equal to  $O(n^{-1/2})$ . *Random Oper. Stoch. Equ.* **6** 359–408. [MR1657835](#)
- [17] Girko, V.L. (2002). Extended proof of the statement: Convergence rate of the expected spectral functions of symmetric random matrices  $\Xi_n$  is equal to  $O(n^{-1/2})$  and the method of critical steepest descent. *Random Oper. Stoch. Equ.* **10** 253–300. [MR1923427](#)

- [18] Götze, F., Naumov, A. and Tikhomirov, A. (2015). Local semicircle law under moment conditions. Part I: The Stieltjes transform. Preprint. Available at [arXiv:1510.07350](https://arxiv.org/abs/1510.07350).
- [19] Götze, F., Naumov, A. and Tikhomirov, A. (2015). Local semicircle law under moment conditions. Part II: Localization and delocalization. Preprint. Available at [arXiv:1511.00862](https://arxiv.org/abs/1511.00862).
- [20] Götze, F., Naumov, A. and Tikhomirov, A. (2015). Limit theorems for two classes of random matrices with dependent entries. *Theory Probab. Appl.* **59** 23–39.
- [21] Götze, F. and Tikhomirov, A. (2002). Rate of convergence to the semi-circular law for the Gaussian unitary ensemble. *Theory Probab. Appl.* **47** 381–387. [MR2003206](https://doi.org/10.1007/s11222-002-9020-2)
- [22] Götze, F. and Tikhomirov, A. (2003). Rate of convergence to the semi-circular law. *Probab. Theory Related Fields* **127** 228–276. [MR2013983](https://doi.org/10.1007/s00332-003-0001-1)
- [23] Götze, F. and Tikhomirov, A. (2005). The rate of convergence for spectra of GUE and LUE matrix ensembles. *Cent. Eur. J. Math.* **3** 666–704. [MR2171668](https://doi.org/10.1007/s10035-005-0001-1)
- [24] Götze, F. and Tikhomirov, A. (2006). Limit theorems for spectra of random matrices with martingale structure. *Theory Probab. Appl.* **51** 171–192. [MR2324173](https://doi.org/10.1007/s11222-006-9020-2)
- [25] Götze, F. and Tikhomirov, A. (2010). On the rate of convergence to the semi-circular law. Preprint. Available at [arXiv:1109.0611](https://arxiv.org/abs/1109.0611).
- [26] Götze, F. and Tikhomirov, A. (2014). Rate of convergence of the empirical spectral distribution function to the semi-circular law. Preprint. Available at [arXiv:1407.2780](https://arxiv.org/abs/1407.2780).
- [27] Götze, F. and Tikhomirov, A. (2015). Optimal bounds for convergence of expected spectral distributions to the semi-circular law for the  $4 + \varepsilon$  moment ensemble. Preprint. Available at [arXiv:1511.03692](https://arxiv.org/abs/1511.03692).
- [28] Götze, F. and Tikhomirov, A. (2016). Optimal bounds for convergence of expected spectral distributions to the semi-circular law. *Probab. Theory Related Fields* **165** 163–233. [MR3500270](https://doi.org/10.1007/s00332-016-0001-1)
- [29] Gustavsson, J. (2005). Gaussian fluctuations of eigenvalues in the GUE. *Ann. Inst. Henri Poincaré B, Probab. Stat.* **41** 151–178. [MR2124079](https://doi.org/10.1016/j.aip.2005.02.001)
- [30] Hanson, D.L. and Wright, F.T. (1971). A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Stat.* **42** 1079–1083. [MR0279864](https://doi.org/10.1080/01621459.1971.10473864)
- [31] Lee, J.O. and Yin, J. (2014). A necessary and sufficient condition for edge universality of Wigner matrices. *Duke Math. J.* **163** 117–173. [MR3161313](https://doi.org/10.1017/S0013071513000131)
- [32] Naumov, A.A. (2014). Limit theorems for two classes of random matrices with Gaussian elements. *J. Math. Sci.* **204** 140–147.
- [33] Pastur, L. (1973). Spectra of random selfadjoint operators. *Uspekhi Mat. Nauk* **28** 3–64. [MR0406251](https://doi.org/10.1070/10663692197300000000000000000000)
- [34] Rudelson, M. and Vershynin, R. (2013). Hanson–Wright inequality and sub-Gaussian concentration. *Electron. Commun. Probab.* **18** no. 82, 9. [MR3125258](https://doi.org/10.1214/12-ECP928)
- [35] Tao, T. (2012). *Topics in Random Matrix Theory. Graduate Studies in Mathematics* **132**. Providence, RI: Amer. Math. Soc.. [MR2906465](https://doi.org/10.1090/S0270-9139-2012-095465-0)
- [36] Tao, T. and Vu, V. (2013). Random matrices: Sharp concentration of eigenvalues. *Random Matrices Theory Appl.* **2** 1350007, 31. [MR3109424](https://doi.org/10.1142/S179894813500007)
- [37] Timushev, D.A., Tikhomirov, A.N. and Kholopov, A.A. (2007). On the accuracy of the approximation of the GOE spectrum by the semi-circular law. *Theory Probab. Appl.* **52** 180–185. [MR2354577](https://doi.org/10.1007/s11222-007-9020-2)
- [38] Wigner, E.P. (1955). Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math. (2)* **62** 548–564. [MR0077805](https://doi.org/10.2307/2372055)

*Received April 2016 and revised February 2017*