Записки научных
семинаров ПОМИ
Том 364, 2009 г.

S. G. Bobkov[1], **F. Götze**[2], **A. N. Tikhomirov**[3]

ON THE CONCENTRATION OF HIGH DIMENSIONAL MATRICES WITH RANDOMLY SIGNED ENTRIES

Abstract. Results on the concentration and asymptotic behaviour of large dimensional random matrices with random signs are obtained. They extend corresponding results, originating in the 1978 work of V. N. Sudakov, in the scheme of weighted sums with limiting Gaussian mixture families to Wigner distributions.

1. Introduction

Consider a family $\mathbf{X} = \{X_{jk}\}$, $1 \le j \le k \le n$, of random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Put $X_{jk} = X_{kj}$ for $1 \le k < j \le n$, and introduce the symmetric matrices

$$\mathbf{W}(\varepsilon) = \frac{1}{\sqrt{n}} \begin{pmatrix} \varepsilon_{11}X_{11} & \varepsilon_{12}X_{12} & \cdots & \varepsilon_{1n}X_{1n} \\ \varepsilon_{21}X_{21} & \varepsilon_{22}X_{22} & \cdots & \varepsilon_{2n}X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1}X_{n1} & \varepsilon_{n2}X_{n2} & \cdots & \varepsilon_{nn}X_{nn} \end{pmatrix},$$

where $\varepsilon = \{\varepsilon_{jk}\}$ denotes an arbitrary family of signs ± 1, satisfying the symmetry condition $\varepsilon_{jk} = \varepsilon_{kj}$. This constitutes an ensemble of $2^{n(n+1)/2}$ random matrices. Each of them has a random spectrum $\{\lambda_1(\varepsilon), \dots, \lambda_n(\varepsilon)\}$ and an associated spectral distribution function

$$F_{n,\varepsilon}(x, \omega) = \frac{1}{n} \operatorname{card} \{j \le n : \lambda_j(\varepsilon) \le x\}, \quad x \in \mathbb{R}, \quad \omega \in \Omega.$$

Averaging over the random values $X_{ij}(\omega)$, define the expected (non-random) empirical distribution functions

$$F_{n,\varepsilon}(x) = \mathbf{E}\, F_{n,\varepsilon}(x, \omega).$$

[1]Research partially supported by NSF grant.
[3]Research supported by the DFG-Forschergruppe FOR 399/1. Partially supported by INTAS, grant No. 03-51-5018, and by RFBR–DFG, grant No. 04-01-04000.

We shall study the general question about the possible limit laws for $F_{n,\varepsilon}$. In particular, in this randomized model, it is natural to ask whether or not there is for growing n a certain universal limit distribution which approximates $F_{n,\varepsilon}$ for most choices of ε's. Our aim is to show that, under some regularity hypotheses on the distribution of the $\mathbf{X}$'s, such a limit distribution exists and may be characterized as a mixture of the Wigner's semi-circle laws. The whole picture is very similar to that in the case of the weighted sums of dependent random variables, where mixtures of Gaussian measures play the role of typical distributions (a remarkable observation, going back to the work of V. N. Sudakov [23]).

To be more precise, for the truncated random variables $X_{jk}^{(\tau)} = X_{jk} I_{\{|X_{jk}| \leqslant \tau\sqrt{n}\}}$ with truncation level $\tau > 0$ define the quantities

$$(\sigma_{n,j}^{(\tau)})^2 = \frac{1}{n}\sum_{k=1}^{n}(X_{jk}^{(\tau)})^2 \quad \text{and} \quad (\sigma_n^{(\tau)})^2 = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}(X_{jk}^{(\tau)})^2.$$

We impose the following conditions: for any $\tau > 0$, as $n \to \infty$,

(1) $\sup_{j,k} \mathbf{E} X_{jk}^2 = O(1)$;

(2) $L_n(\tau) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} \mathbf{E} X_{jk}^2 I(|X_{jk}| > \tau\sqrt{n}\,) = o(1)$;

(3) $\Delta_1^2 = \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\left((\sigma_{n,j}^{(\tau)})^2 - (\sigma_n^{(\tau)})^2\right)^2 = o(1)$.

The second assumption is the usual Lindeberg-type condition, introduced by L. A. Pastur for random matrices with independent entries [19], often called Wigner ensemble. The third condition may be described as a kind of variance stabilization.

To measure closeness of distributions on the real line, we shall use the Lévy metric: Given distribution functions F and G, the distance $L(F,G)$ is defined as the infimum over all $\delta > 0$ such that for all $x \in \mathbb{R}$,

$$F(x-\delta) - \delta \le G(x) \le F(x+\delta) + \delta.$$

We write $\mathcal{L}(\eta)$ for the distribution of a random variable η. Let μ_n denote the standard Bernoulli measure on the discrete cube $\{-1,1\}^{n(n+1)/2}$. To each point it assigns the mass $2^{-n(n+1)/2}$. Finally, define $\sigma_n^2 = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} X_{jk}^2$ $(\sigma_n \geqslant 0)$.

Theorem 1.1. *Assume $\sigma_n \Rightarrow \sigma$, weakly in distribution, as $n \to \infty$, for some random variable $\sigma \geq 0$. Then, assuming the conditions* (1)–(3) *for any $\delta > 0$, we have*

$$\mu_n\{\varepsilon : L(F_{n,\varepsilon}, \mathcal{L}(\sigma\xi)) > \delta\} \to 0,$$

where ξ is a random variable independent of σ and distributed according to the semi-circle law with variance 1 *(that is, with density $\frac{1}{2\pi}\sqrt{4-x^2}\, I(|x| \leq 2)$).*

The statement may formally be sharpened by replacing δ with a sequence $\delta_n \to 0$. Thus, for large n most of the distribution functions $F_{n,\varepsilon}$ are close to $\mathcal{L}(\sigma\xi)$ in the the metric L. Note that $\mathcal{L}(\sigma\xi)$ may be viewed as a mixture of the semi-circle laws with mixing measure $\mathcal{L}(\sigma)$. If a weak law of large numbers applies to X_{jk}^2, so that σ_n^2 converges in probability to a constant σ^2, then the universal limit distribution $\mathcal{L}(\sigma\xi)$ is the semi-circle law with variance σ^2.

Note as well, when the distribution of $\mathbf{X}$ in $\mathbb{R}^{n(n+1)/2}$ is symmetric with respect to all coordinate axes, there will be no dependence of $\mathbf{W}(\varepsilon)$ of the "parameter" ε, so all $2^{n(n+1)/2}$ distribution functions $F_{n,\varepsilon}$ are identical. Therefore, taking, for example, $\varepsilon_{jk} = 1$, the model reduces to the non-randomized case and yields the Wigner theorem for a certain family of random matrices with dependent entries.

As for the general (non-symmetric) case, there is one interesting concentration aspect, which in essence does not require any hypotheses on the joint distribution of the entries of $\mathbf{X}$. Define the average distribution function with respect to the ε's via

$$F_n(x) = 2^{-n(n+1)/2} \sum_{\varepsilon} F_{n,\varepsilon}(x), \quad x \in \mathbb{R}.$$

It may be described as the expected empirical measure for eigenvalues of $W(\varepsilon)$, where now ε is viewed as a random vector on the discrete cube, independent of $\mathbf{X}$ (with distribution μ_n).

It turns out that already the assumption (1) suffices to guarantee that most of the distribution functions $F_{n,\varepsilon}$ are very close to F_n.

Theorem 1.2. *If $\mathbf{E}X_{jk}^2 \leq \alpha^2$ for all $j,k \leq n$, then for any $\delta \in (0,1)$,*

$$\mu_n\{\varepsilon : L(F_{n,\varepsilon}, F_n) \geq \delta\} \leq C\, e^{-cn^2}, \tag{1.1}$$

where C and c are positive constants depending on δ and α, only.

An important feature of the bound (1.1) is that with respect to the "dimension" the exponent is proportional to n^2. This is an analogue to results on the concentration of distributions of randomized sums for dependent summands. Namely, given a sequence of random variables, say, $\xi_1, \dots, \xi_n$, one considers the sums

$$S_n(\theta) = \theta_1 \xi_1 + \dots + \theta_n \xi_n, \quad \theta = (\theta_1, \dots, \theta_n)$$

with coefficients from the unit sphere $S^{n-1} : \theta_1^2 + \dots \theta_n^2 = 1$. In 1978, V. N. Sudakov discovered [23] that under a mild spectral assumption on the correlation operator of ξ_i's, most of the distributions of $S_n(\theta)$ are close to a certain "typical" distribution F_n which is the average of the $\mathcal{L}(S_n(\theta))$'s with respect to the uniform Lebesgue measure $\mu_n(d\theta)$ on S^{n-1}. For the proof, he applied the Lévy–Schmidt isoperimetric theorem on the sphere and the measure concentration phenomenon related to it. Moreover, F_n itself is close to the law of $\rho_n Z$, where $\rho_n^2 = \frac{1}{n} \sum_{i=1}^n \xi_i^2$ and Z is a standard normal random variable independent of ρ_n. Thus, most of $\mathcal{L}(S_n(\theta))$'s are close to a mixture of symmetric Gaussian measures on the line. A different approach to Sudakov's theorem, involving the case of independent coefficients, was later developed by H. von Weizsäcker [25]. Various extensions and refinements were intensively discussed in the literature (cf., e.g., [1, 5, 18]). One of the results was that, the coefficients may have a special structure. For example, if $\theta_i = \varepsilon_i / \sqrt{n}$ with $\varepsilon_i = \pm 1$, and ξ_i's are orthonormal in $L^2(\Omega, \mathcal{F}, \mathbf{P})$, it was shown in [6] that

$$\nu_n\{\varepsilon : L(\mathcal{L}(S_n), F_n) \geq \delta\} \leq C\, e^{-cn},$$

where now ν_n denotes the normalizing counting measure on the discrete cube $\{-1, 1\}^n$. (In case of the trigonometric system, the problem goes back to the work by R. Salem and A. Zygmund [22].) Thus, the role of the semi-circle law in the framework of randomized matrices, as stated in Theorems 1.1–1.2, is similar to that of the normal distribution in the framework of randomized sums.

The paper is organized as follows.

In Sec. 2, we introduce necessary notations and recall some useful identities and bounds for resolvent matrices.

Section 3 deals with the concentration property of the family of the characteristic functions associated with the expected empirical distributions $F_{n,\varepsilon}$. In Sec. 4, we convert this property in terms of the Lévy distance and derive Theorem 1.2 in a more precised form. A shorter proof of it in case of bounded entries of the matrix $\mathbf{X}$ can be given on the basis of Talagrand's concentration inequality on the cube; this and related lines of arguments was discussed by A. Guionnet and O. Zeitouni, cf. [9]. To save more generality we have chosen a different and more routine way, where a different concentration phenomenon on the cube is applied through the expected Stiltjes transform of the matrix $\mathbf{W}(\varepsilon)$.

In Sec. 5, we show that the average distribution F_n is close to a mixture of the semi-circle laws. For the readers convenience, this section is divided into three subsections. The proof of Theorem 1.1 is completed in Sec. 6. In the Appendix, we prove some auxiliary bounds for resolvent matrices.

2. General identities and bounds for resolvent

Let $\mathcal{L}_n$ denote the collection of all square $n\times n$ matrices $\mathbf{A}$ with complex entries a_{jk}, $1 \le j,k \le n$. Also, denote by $\mathcal{H}_n$ the collection of all $n \times n$ symmetric matrices with real entries. Given a matrix $\mathbf{A}$ in $\mathcal{L}_n$ with entries a_{jk}, $\mathrm{Tr}\,(\mathbf{A}) = \sum_{j=1}^n a_{jj}$ defines its trace. The Hilbert–Schmidt norm of $\mathbf{A}$ is given by $\|\mathbf{A}\|_{HS}^2 = \sum_{j,k=1}^n |a_{jk}|^2$. As usual, $\|A\|$ denotes the spectral norm.

Denote by $\mathbf{I}$ the unit matrix of size $n \times n$. For a complex number $z = u + iv$, the value of the resolvent of $\mathbf{A} \in \mathcal{H}_n$ at z is defined by

$$\mathbf{R}(z) = (\mathbf{A} - z\mathbf{I})^{-1}.$$

It is well-defined, whenever $v > 0$. For short, we also write $\mathbf{R} = \mathbf{R}(z)$.

The next following Lemmas 2.1–2.5 are easily obtained by diagonalization of symmetric matrices.

Lemma 2.1. *For any matrix $\mathbf{A}$ in $\mathcal{H}_n$ and $z = u + iv$ with $v > 0$,*

$$|\mathrm{Tr}\,(\mathbf{R})| \le \frac{n}{v}, \quad \|\mathbf{R}\|_{HS} \le \frac{\sqrt{n}}{v}, \quad \|\mathbf{R}^2\|_{HS} \le \frac{\sqrt{n}}{v^2}.$$

Lemma 2.2. *For any $\mathbf{A}$ in $\mathcal{H}_n$ and $z = u + iv$ with $v > 0$, we have $\|\mathbf{R}\| \le \frac{1}{v}$. In particular, for any $j = 1, \dots, n$,*

$$\sum_{k=1}^n |R_{jk}|^2 \le \frac{1}{v^2}.$$

Lemma 2.3. *For any* $\mathbf{A}$ *in* $\mathcal{H}_n$ *and* $v > 0$,

$$\int_{-\infty}^{+\infty} \|\mathbf{R}(u+iv)\|_{HS}^2 \, du = \frac{n\pi}{v}.$$

More precisely, for any $j = 1, \dots, n$,

$$\int_{-\infty}^{+\infty} \sum_{k=1}^{n} |R(u+iv)_{jk}|^2 \, du = \frac{\pi}{v}. \tag{2.1}$$

Lemma 2.4. *Given* $\mathbf{A} \in \mathcal{H}_n$ *and* $\mathbf{M} \in \mathcal{L}_n$, *for any* $v > 0$,

$$\int_{-\infty}^{+\infty} |\mathrm{Tr}\,(\mathbf{M}\mathbf{R}^2(u+iv)| \, du \le \frac{\sqrt{n}\,\pi}{v} \, \|\mathbf{M}\|_{HS}.$$

We will also use the following elementary identity:

Lemma 2.5. *Given matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *in* $\mathcal{H}_n$, *let* $\mathbf{R} = \mathbf{R}_{\mathbf{A}}(z)$ *be the resolvent of* $\mathbf{A}$ *and* $\mathbf{R}' = \mathbf{R}_{\mathbf{A}-\mathbf{B}}(z)$ *be the resolvent of* $\mathbf{A} - \mathbf{B}$. *Then,*

$$\mathbf{R}' = \mathbf{R} + \mathbf{R}\mathbf{B}\mathbf{R}' = \mathbf{R} + \mathbf{R}\mathbf{B}\mathbf{R} + \mathbf{R}\mathbf{B}\mathbf{R}\mathbf{B}\mathbf{R}'.$$

3. Concentration of characteristic functions

To study the concentration problem, we mainly work with the Stieltjes transform of the expected spectral distributions $F_{n,\varepsilon}$, defined by

$$S_n(z,\varepsilon) = \int_{-\infty}^{+\infty} \frac{dF_{n,\varepsilon}(x)}{x-z} = \frac{1}{n}\,\mathrm{Tr}\,(\mathbf{R}(z,\varepsilon)),$$

where $\mathbf{R}(z,\varepsilon) = \mathbf{E}\,(\mathbf{W}(\varepsilon) - z\mathbf{I})^{-1}$ represents the expected resolvent of the random matrix $\mathbf{W}(\varepsilon)$ at the point $z = u + iv$. We always assume that $v > 0$. The Stieltjes transform is related to the characteristic function $f_n(t,\varepsilon)$ of $F_{n,\varepsilon}$ through the simple relation

$$f_n(t,\varepsilon) = \frac{e^{v|t|}}{\pi} \int_{-\infty}^{+\infty} e^{itu}\,\mathrm{Im}\,S_n(z,\varepsilon)\,du, \quad t \in \mathbb{R}.$$

Correspondingly, the μ_n-average of such characteristic functions represents the characteristic function of $F_n(x)$:

$$f_n(t) = \frac{e^{v|t|}}{\pi} \int_{-\infty}^{+\infty} e^{itu} \,\mathrm{Im}\, S_n(z)\, du. \tag{3.1}$$

Note that the imaginary part of the Stieltjes transform as a function of the real variable u is always integrable with respect to Lebesgue measure.

As a first step, we show that for any $z = u + iv$ and $v > 0$, the values of the function $\varepsilon \to S_n(z, \varepsilon)$ are strongly concentrated about its μ_n-mean

$$S_n(z) = \int S_n(z, \varepsilon)\, d\mu_n(\varepsilon) = 2^{-n(n+1)/2} \sum_{\varepsilon} S_n(x, \varepsilon).$$

This can be seen by virtue of the so-called concentration phenomenon on the discrete cube.

Namely, given a complex-valued function $g = g(\varepsilon)$ on $\{-1, 1\}^N$, consider its discrete gradient with modulus given by

$$|\nabla g(\varepsilon)|^2 = \sum_{k=1}^{N} |g(\varepsilon) - g(\varepsilon^k)|^2,$$

where ε^k represents the ± 1-sequence which is obtained from the sequence ε by replacing ε_k with $-\varepsilon_k$ on the kth place. The distribution of the modulus of the gradient is occurs in a number of deviation inequalities. As the simplest example, one may consider the so-called Poincaré-type inequality.

Lemma 3.1. *For any function $g : \{-1, 1\}^N \to \mathbf{C}$ such that $\int g\, d\mu_N = 0$ with respect to the normalized counting measure μ_N, we have*

$$\int |g|^2 \, d\mu_N \leq \frac{1}{4} \int |\nabla g|^2 \, d\mu_N.$$

Moreover, there is an inequality of Gaussian-type in terms of the ℓ^∞-norm of the modulus of the gradient (cf. [7, 15]):

Lemma 3.2. *If* $|\nabla g(\varepsilon)| \leq \alpha$, *for all* $\varepsilon \in \{-1,1\}^N$, *then*

$$\mu_N \left\{\varepsilon : \left|g(\varepsilon) - \int g\, d\mu_N\right| \geq h\right\} \leq 4\, e^{-h^2/4\alpha^2}, \quad h \geq 0.$$

We need to bound the modulus of the gradient $|\nabla f_n(t,\varepsilon)|$ for the specific function $g(\varepsilon) = f_n(t,\varepsilon)$ uniformly over all points ε in $\{-1,1\}^{n(n+1)/2}$. In our setting,

$$|\nabla f_n(t,\varepsilon)|^2 \;=\; \sum_{1 \leq j \leq k \leq n} |f_n(t,\varepsilon) - f_n(t,\varepsilon^{jk})|^2,$$

where, ε^{jk} represents the symmetric ± 1-matrix of size $n \times n$, which is obtained from the matrix ε by replacing the entry ε_{jk} with $-\varepsilon_{jk}$ on the jkth and kjth places.

Lemma 3.3. *If* $\mathbf{E}X_{jk}^2 \leq \alpha^2$, *for all* j,k, *then for all* $\varepsilon \in \{-1,1\}^{n(n+1)/2}$ *and* $t \in \mathbb{R}$,

$$|\nabla f_n(t,\varepsilon)| \leq \frac{C\alpha|t|}{n}\, \max\{1, \alpha|t|\}, \tag{3.2}$$

where C *is an absolute constant.*

By the inequalities in Lemmas 3.1–3.2, we therefore obtain:

Theorem 3.1. *Let* $\mathbf{E}X_{jk}^2 \leq \alpha^2$, *for all* j,k. *Then we get, with respect to the normalized counting measure* μ_n *on* $\{-1,1\}^{n(n+1)/2}$, *for all* $t \in \mathbb{R}$,

$$\int |f_n(t,\varepsilon) - f_n(t)|^2\, d\mu_n(\varepsilon) \leq \frac{C^2 b^2}{n^2}. \tag{3.3}$$

where $b = \alpha|t|\, \max\{1, \alpha|t|\}$. *Moreover, with some universal* $c > 0$, *for all* $h \geq 0$,

$$\mu_n \left\{\varepsilon : |f_n(t,\varepsilon) - f_n(t)| \geq h\right\} \leq 4\, e^{-c\, n^2 h^2/b^2}. \tag{3.4}$$

Proof of Lemma 3.3. Let $t \neq 0$. We use an elementary bound

$$|\nabla f_n(t,\varepsilon)| \leq \sqrt{2}\, \sup_A \left| \sum_{1 \leq j \leq k \leq n} a_{jk} \left(f_n(t,\varepsilon) - f_n(t,\varepsilon^{jk})\right) \right|, \tag{3.5}$$

where the sup is taken over all matrices $\mathbf{A}$ in $\mathcal{L}_n$ with real entries $\{a_{jk}\}$ such that

$$\sum_{1 \leq j \leq k \leq n} a_{jk}^2 = 1. \tag{3.6}$$

Here we assume that $a_{jk} = 0$ for $j > k$. Fix such a matrix $\mathbf{A}$.

According to the relation between characteristic functions and Stieltjes transform, we have for $z = u + iv$, $v > 0$, and a_{jk} real,

$$\sum_{1 \le j \le k \le n} a_{jk} \left(f_n(t, \varepsilon) - f_n(t, \varepsilon^{jk}) \right) \\ = \frac{e^{v|t|}}{\pi} \int_{-\infty}^{+\infty} e^{itu} \operatorname{Im} \sum_{1 \le j \le k \le n} a_{jk} \left(S_n(z, \varepsilon) - S_n(z, \varepsilon^{jk}) \right) du;$$

hence,

$$\Big| \sum_{1 \le j \le k \le n} a_{jk} \left(f_n(t, \varepsilon) - f_n(t, \varepsilon^{jk}) \right) \Big| \\ \le \frac{e^{v|t|}}{\pi} \int_{-\infty}^{+\infty} \Big| \sum_{1 \le j \le k \le n} a_{jk} \left(S_n(z, \varepsilon) - S_n(z, \varepsilon^{jk}) \right) \Big| du. \quad (3.7)$$

Given $1 \le j \le k \le n$, write

$$\mathbf{W}(\varepsilon^{jk}) = \mathbf{W}(\varepsilon) - \frac{2\varepsilon_{jk} X_{jk}}{\sqrt{n}} \, \mathbf{D}^{jk},$$

where we $\mathbf{D}^{jk}$ denotes the symmetric matrix which has zero entries everywhere except at the jkth and kjth entries, where it equals 1. In particular, $\mathbf{D}^{jj}$ has zero entries everywhere except in the jjth entry, where it is 1. By Lemma 2.5 with $\mathbf{B} = \frac{2\varepsilon_{jk} X_{jk}}{\sqrt{n}} \, \mathbf{D}^{jk}$, the resolvents $\mathbf{R}$ of $\mathbf{W}(\varepsilon)$ and $\mathbf{R}^{jk}$ of $\mathbf{W}(\varepsilon^{jk})$ are related by

$$\mathbf{R}^{jk} = \mathbf{R} + \frac{2\varepsilon_{jk} X_{jk}}{\sqrt{n}} \, \mathbf{R}\mathbf{D}^{jk}\mathbf{R} + \frac{4X_{jk}^2}{n} \, \mathbf{R}\mathbf{D}^{jk}\mathbf{R}\mathbf{D}^{jk}\mathbf{R}^{jk}.$$

Taking the trace of the both sides and using commutativity of the trace, we obtain that

$$\operatorname{Tr}(\mathbf{R}^{jk}) = \operatorname{Tr}(\mathbf{R}) + \frac{2\varepsilon_{jk} X_{jk}}{\sqrt{n}} \, \operatorname{Tr}(\mathbf{D}^{jk}\mathbf{R}^2) + \frac{4X_{jk}^2}{n} \, \operatorname{Tr}(\mathbf{D}^{jk}\mathbf{R}\mathbf{D}^{jk}\mathbf{R}^{jk}\mathbf{R}).$$

Taking expectations and dividing by n, we get

$$S_n(z, \varepsilon^{jk}) - S_n(z, \varepsilon) = \frac{2}{n^{3/2}} \, \mathbf{E} \, \varepsilon_{jk} X_{jk} \, \operatorname{Tr}(\mathbf{D}^{jk}\mathbf{R}^2) + \frac{4}{n^2} \, \mathbf{E} \, b_{jk} X_{jk}^2,$$

where $b_{jk} = \mathrm{Tr}\,(\mathbf{D}^{jk}\mathbf{R}\mathbf{D}^{jk}\mathbf{R}^{jk}\mathbf{R})$. Hence,

$$\sum_{1\le j\le k\le n} a_{jk}\left(S_n(z,\varepsilon^{jk}) - S_n(z,\varepsilon)\right) = \frac{2}{n^{3/2}}\,\mathbf{E}\,\mathrm{Tr}\,(\mathbf{M}\mathbf{R}^2) + \frac{4}{n^2}\sum_{1\le j\le k\le n} a_{jk}\mathbf{E}\,b_{jk}X_{jk}^2,$$

and

$$\int_{-\infty}^{+\infty}\Big|\sum_{1\le j\le k\le n} a_{jk}\left(S_n(z,\varepsilon^{jk}) - S_n(z,\varepsilon)\right)\Big|\,du \tag{3.8}$$
$$\le \frac{2}{n^{3/2}}\,\mathbf{E}\int_{-\infty}^{+\infty}|\mathrm{Tr}\,(\mathbf{M}\mathbf{R}^2)|\,du + \frac{4}{n^2}\sum_{1\le j\le k\le n}|a_{jk}|\,\mathbf{E}\Big[X_{jk}^2\int_{-\infty}^{+\infty}|b_{jk}|\,du\Big],$$

where $\mathbf{M}$ is the matrix with entries $M_{jk} = a_{jk}\varepsilon_{jk}X_{jk}$.

By Lemma 2.4,

$$\int_{-\infty}^{+\infty}|\mathrm{Tr}\,(\mathbf{M}\mathbf{R}^2)|\,du \le \frac{\sqrt{n}\,\pi}{v}\,\|\mathbf{M}\|_{HS}.$$

Since $\|\mathbf{M}\|_{HS}^2 = \sum_{j,k}|a_{jk}|^2X_{jk}^2$, by (3.6), we have, $\mathbf{E}\,\|\mathbf{M}\|_{HS}^2 \le \alpha^2$, so $\mathbf{E}\,\|\mathbf{M}\|_{HS} \le \alpha$. Thus,

$$\mathbf{E}\int_{-\infty}^{+\infty}|\mathrm{Tr}\,(\mathbf{M}\mathbf{R}^2)|\,du \le \frac{\sqrt{n}\,\pi}{v}\,\alpha. \tag{3.9}$$

Now, let's turn to the last term in (3.7) and recall the definition of b_{jk}. Note that, for all symmetric matrices $\mathbf{G}$ and $\mathbf{H}$ in $\mathcal{L}_n$, we have in general

$$\mathrm{Tr}\,(\mathbf{D}^{jk}\mathbf{G}\mathbf{D}^{jk}\mathbf{H}) = G_{jj}H_{kk} + G_{kk}H_{jj} + 2\,G_{jk}H_{jk}, \quad \text{for} \quad j<k.$$

Also, $\mathrm{Tr}\,(\mathbf{D}^{jj}\mathbf{G}\mathbf{D}^{jj}\mathbf{H}) = G_{jj}H_{jj}$. We apply these identities to $\mathbf{G} = \mathbf{R}$ and $\mathbf{H} = \mathbf{R}^{jk}\mathbf{R}$. By Lemma 2.2, $|R_{jk}| \le \frac{1}{v}$ for all k, so, whether $j<k$ or $j=k$, we have

$$|\mathrm{Tr}\,(\mathbf{D}^{jk}\mathbf{G}\mathbf{D}^{jk}\mathbf{H})| \le \frac{4}{v}\,|H_{jk}|. \tag{3.10}$$

Write $H_{jk} = (\mathbf{R}^{jk}\mathbf{R})_{jk} = \sum_{l=1}^n (\mathbf{R}^{jk})_{jl} R_{lk}$. By the usual Cauchy inequality,

$$|H_{jk}| \le \sqrt{\sum_{l=1}^n |(R^{jk})_{jl}|^2}\ \sqrt{\sum_{l=1}^n |R_{lk}|^2},$$

so integrating over u and using Cauchy–Bunyakovski's inequality, we get

$$\int_{-\infty}^{+\infty} |H_{jk}|\,du \le \sqrt{\int_{-\infty}^{+\infty} \sum_{l=1}^n |(R^{jk})_{jl}|^2\,du}\ \sqrt{\int_{-\infty}^{+\infty} \sum_{l=1}^n |R_{lk}|^2\,du}.$$

By Lemma 2.3, applied to the matrices $\mathbf{R}^{jk}$ and $\mathbf{R}$, and in view of the symmetry of $\mathbf{R}$, the above right-hand side is equal to $\frac{\pi}{v}$. Hence, from (3.9),

$$\int_{-\infty}^{+\infty} |b_{jk}|\,du \le \frac{4\pi}{v^2},$$

and thus the second expectation in (3.7) is bounded by $\frac{4\pi}{v^2}\,\alpha^2$. In addition, by the Cauchy's inequality and (3.6), $\sum_{1\le j\le k\le n} |a_{jk}| \le n$. Therefore, the second term in (3.7) is bounded by $\frac{16\pi}{nv^2}\,\alpha^2$. Together with (3.8), this bound yields via (3.7) that

$$\int_{-\infty}^{+\infty} \Big| \sum_{1\le j\le k\le n} a_{jk}\,\big(S_n(z,\varepsilon^{jk}) - S_n(z,\varepsilon)\big)\Big|\,du$$
$$\le \frac{2\pi\alpha}{nv} + \frac{16\pi\alpha^2}{nv^2} \le \frac{18\pi\alpha}{nv}\,\max\Big\{1, \frac{\alpha}{v}\Big\}.$$

Recalling (3.5) and (3.7), we obtain that for any $v > 0$,

$$|\nabla f_n(t,\varepsilon)| \le 18\sqrt{2}\,e^{v|t|}\,\frac{\alpha}{nv}\,\max\Big\{1, \frac{\alpha}{v}\Big\}.$$

Choosing $v = 1/|t|$ finishes the proof of the lemma with $C = 18\pi e\sqrt{2}$. $\square$

4. Concentration of distributions in Lévy metric

We shall now study the closeness of the randomized distribution functions $F_{n,\varepsilon}$ to the mean distribution functions F_n in terms of the Lévy metric. Note that this metric is related to the usual Kolmogorov's supremum distance by

$$0 \leq L(F,G) \leq \|F-G\| \leq 1.$$

Conversely, if G has a Lipschitz semi-norm bounded by a constant C, then we get an opposite bound $\|F-G\| \leq (1+C)\,L(F,G)$.

Inequality (3.4) of Theorem 3.1 about the concentration property of the characteristic functions may be transformed into a concentration property of the distribution functions by virtue of Zolotarev's general bound, cf. [28]:

Lemma 4.1. *If f and g are the characteristic functions of the distribution functions F and G, then with some absolute constants $C>0$ and $T_0>1$,*

$$L(F,G) \leq C\bigg[\int_0^T \frac{|f(t)-g(t)|}{t}\,dt + \frac{C\log T}{T}\bigg], \quad T \geq T_0. \tag{4.1}$$

We shall use as well the following general elementary result.

Lemma 4.2. *If $\int x^2\,dF(x) \leq \alpha^2$ and $\int x^2\,dG(x) \leq \alpha^2$, then the function $\psi(t) = \frac{f(t)-g(t)}{t}$ has derivative satisfying $|\psi'(t)| \leq \alpha^2$, for any $t>0$.*

Indeed, write

$$\psi'(t) = \int_{-\infty}^{+\infty} \frac{e^{itx}(1-itx)-1}{(tx)^2}\,x^2\,d(F(x)-G(x)).$$

But the function $\xi(s) = e^{is}(1-is)-1$ has the derivative $\xi'(s) = s\,e^{is}$, so $|\xi(s)| \leq \frac{1}{2}\,s^2$ for all $s \in \mathbb{R}$.

Now fix a number δ, such that $0<\delta<1$, fix a number $h>0$, and take $T \geq T_0$ to be specified later on. Consider a partition of $[0,T]$ into m consecutive intervals Δ_i, $i=0,1,\dots,m-1$, of equal length, not exceeding h. These intervals have endpoints $t_i = \frac{T}{m}\,i$, and

$$t_{i+1}-t_i = \frac{T}{m} \leq h, \quad 0 \leq i \leq m-1. \tag{4.2}$$

Introduce the subsets of the discrete cube

$$\Omega_i = \left\{ \varepsilon : \frac{|f_n(t_i,\varepsilon) - f_n(t_i)|}{t_i} < h \right\}, \quad 1 \le i \le m.$$

By Theorem 3.1, $1 - \mu_n(\Omega_i) \le 4\exp\{-cn^2h^2/B^2\}$, where $B = \alpha\max\{1, \alpha T\}$, and c is a positive universal constant. Hence,

$$1 - \mu_n\left\{ \bigcap_{i=1}^m \Omega_i \right\} \le 4m\,\exp\{-cn^2h^2/B^2\}.$$

In (4.2), we may take $m = \left[\frac{T}{h}\right] + 1$ (the integer part), so

$$1 - \mu_n\left\{ \bigcap_{i=1}^m \Omega_i \right\} \le \frac{4\,(T+h)}{h}\,\exp\{-cn^2h^2/B^2\}. \tag{4.3}$$

Recall that $f_n(t,\varepsilon) = \frac{1}{n}\,\mathbf{E}\,\mathrm{Tr}\,(e^{it\mathbf{W}(\varepsilon)})$. Therefore,

$$f_n''(0,\varepsilon) = -\frac{1}{n}\,\mathbf{E}\,\mathrm{Tr}\,(\mathbf{W}(\varepsilon)^2) = -\frac{1}{n^2}\sum_{j=1}^n\sum_{k=1}^n \mathbf{E}\,X_{jk}^2.$$

This identity holds for all ε on the discrete cube, so the same is true for $f_n''(0)$. Thus, $|f_n''(0,\varepsilon)| \le \alpha^2$ and the same is true for $f_n''(0)$. Hence, by Lemma 4.2, the function

$$\psi(t,\varepsilon) = \frac{|f_n(t,\varepsilon) - f_n(t)|}{t}$$

has a Lipschitz semi-norm satisfying $\|\psi(t,\varepsilon)\|_{\mathrm{Lip}} \le \alpha^2$ on the half-axis $t > 0$.

Let $\varepsilon \in \bigcap_{i=1}^m \Omega_i$, so that $|\psi(t_i)| < h$ for all i. Any point $t \in [0,T]$ belongs to some interval Δ_i, so $|t - t_i| \le h$ for some $i \le m$. By the Lipschitz property shown above, we obtain that

$$\psi(t,\varepsilon) \le \psi(t_i,\varepsilon) + \|\psi(t,\varepsilon)\|_{\mathrm{Lip}}|t - t_i| < (1+\sigma^2)\,h.$$

Hence, $\int_0^T \psi(t,\varepsilon)\,dt \le (1+\alpha^2)\,Th$, and by Lemma 4.1,

$$L(F_{n,\varepsilon}, F_n) \le C\left[(1+\alpha^2)\,Th + \frac{\log T}{T}\right].$$

Thus, by (4.3), if

$$\delta \geq C\left[(1+\alpha^2)\,Th + \frac{\log T}{T}\right], \tag{4.4}$$

then

$$\mu_n\{L(F_{n,\varepsilon}, F_n) \geq \delta\} \leq \frac{4\,(T+h)}{h}\,\exp\left\{-\frac{c\,n^2 h^2}{\alpha^2 \max\{1, \alpha^2 T^2\}}\right\}. \tag{4.5}$$

The next step is to minimize the right-hand side of (4.5) for all (T, h), satisfying (4.4), together with the assumption $T \geq T_0$. In fact, as an almost optimal choice, we may take $T = \beta\,\frac{\log\frac{2}{\delta}}{\delta}$ with β large enough. Then,

$$\frac{\log T}{T} \leq \delta\,(\frac{\log\beta}{\beta\log 2} + \frac{2}{\beta}) \leq \frac{\delta}{2C},$$

where C is the constant in Zolotarev's inequality (4.1) and (4.4). Furthermore, if we take

$$h = \frac{\delta}{2C\,(1+\alpha^2)\,T},$$

(4.4) holds. As for the right-hand side of (4.5), then for some constant $C' > 0$,

$$\frac{h}{\max\{1, \alpha T\}} \geq \frac{h}{1+\alpha T} \geq \frac{\delta}{C'(1+\alpha^3)\,T^2} = \frac{\delta^3}{\beta^2\,C'(1+\alpha^3)\,\log^2\frac{2}{\delta}}.$$

In addition,

$$\frac{T+h}{h} = \frac{2\beta^2\,C\,(1+\alpha^2)\,\log^2\frac{2}{\delta} + \delta^3}{\delta^3} \leq \frac{C''(1+\alpha^2)}{\delta^4}.$$

Thus, we may conclude:

Theorem 4.1. *For any $\delta \in (0,1)$,*

$$\mu_n\{\varepsilon : L(F_{n,\varepsilon}, F_n) \geq \delta\} \leq C\,e^{-cn^2}, \tag{4.6}$$

where $C = C(\alpha,\delta)$ and $c = c(\alpha,\delta)$ depend on δ and α, only. Here we may choose for some absolute positive constants C and c,

$$C(\alpha,\delta) = \frac{C(1+\alpha^2)}{\delta^4}, \quad c(\alpha,\delta) = \frac{c\,\delta^6}{\alpha^2(1+\alpha^6)\,\log^4\frac{2}{\delta}}.$$

Thus, we get a more precise version of Theorem 1.2. The bound (4.6) seems to be correct with respect to the "dimension" n^2. However, we do not know the optimal order of the function $c(\sigma,\delta)$ for small δ. Choosing $\delta = \delta_n$ of order $\frac{\log n}{n^{1/3}}$ for large n yields a right-hand side smaller than any power of $\frac{1}{n}$. Hence:

Corollary 4.3. *For "most" of ε, we have*

$$L(F_{n,\varepsilon}, F_n) \leq C\, \frac{\log n}{n^{1/3}}$$

with some constant $C = C(\alpha)$. In particular,

$$\int L(F_{n,\varepsilon}, F_n)\, d\mu_n(\varepsilon) \leq C\, \frac{\log n}{n^{1/3}}. \tag{4.7}$$

Remark. The $\log n$ term in (4.7) may be removed by using a Poincaré-type inequality (3.3). Moreover, for bounded mean zero entries X_{jk}, estimate (4.7) may be sharpened to $\frac{\log n}{n^{1/2}}$, and similarly one may also improve dependence of constants in Theorem 4.1 with respect to small values of δ.

Let us describe an argument in this case, which is more/less standard. The map $T : \mathcal{H}_n \to \mathbb{R}^n$, assigning to any symmetric $n \times n$ matrix $Y = (y_{jk})$ the vector of its eigenvalues $(\lambda_1, \dots, \lambda_n)$, written in the increasing order, is Lipschitz with respect to the Hilbert–Schmidt norm, that is, $||T(Y_1) - T(Y_2)|| \leq \|Y_1 - Y_2\|_{HS}$. Hence, for any function f on the real line with finite Lipschitz norm $\|f\|_{\rm Lip}$, the functional

$$T_f(Y) = \frac{f(\lambda_1) + \dots + f(\lambda_n)}{n}$$

has Lipschitz norm at most $\|f\|_{\rm Lip}/n$. Moreover, if f is convex, then T_f is convex, as well. These two properties may be used to study the variance T_f and other similar quantities in case of random Y by posing natural hypotheses on the distribution of the entries Y_{jk}, cf., e.g., [16, 9]. For our randomized model, where $Y_{jk} = \varepsilon_{jk} X_{jk}$ with fixed ε_{jk} and random X_{jk} such that $|X_{jk}| \leq 1$, we may conclude that the functional

$$Q_f(\varepsilon) = \int f\, dF_{n,\varepsilon}$$

is convex with respect to $\varepsilon \in \mathbb{R}^{n(n+1)/2}$ (as mixture of convex functions after averaging over X_{jk}'s) and has Lipschitz norm at most $2\|f\|_{\rm Lip}/n$ (as mixture of functions with Lipschitz norm at most $\|f\|_{\rm Lip}/n$). Therefore, we are in position to apply to Q_f Talagrand's concentration inequality on the discrete cube $\{-1, 1\}^{n(n+1)/2}$, cf. [24], which gives

$$\mu_n\bigg\{\varepsilon : \bigg|\int f\, dF_{n,\varepsilon} - \int f\, dF_n\bigg| \geq h\bigg\} \leq 2e^{-cn^2h^2/\|f\|_{\rm Lip}^2}, \quad h \geq 0, \tag{4.8}$$

with some positive absolute constant c.

This bound is already very similar to estimate (4.6) in Theorem 4.1. In order to relate it to the Lévy distance between $F_{n,\varepsilon}$ and F_n, first we extend (4.8) to a larger family of admissible functions. Namely, if $f = f_1 - f_2$ for some convex f_1 and f_2 with $\|f_i\|_{\rm Lip} \leq \sigma$, it follows from (4.8) that

$$\mu_n\bigg\{\varepsilon : \bigg|\int f\,dF_{n,\varepsilon} - \int f\,dF_n\bigg| \geq h\bigg\} \leq 4e^{-cn^2h^2/\sigma^2}, \quad h \geq 0, \tag{4.9}$$

with some (other) absolute $c > 0$. Now, fix $a \in \mathbb{R}$, $h > 0$, and $\sigma > 0$, and apply the latter to $f_1(x) = \sigma(x-a)^+$ and $f_2(x) = \sigma(x-(a+h))^+$. Then the function f is vanishing on $(-\infty, a]$, is equal to σh on $[a+h, +\infty)$, and is linear on the interval $[a, a+h]$. Therefore,

$$\int f\,dF_{n,\varepsilon} - \int f\,dF_n \geq \sigma h\,(F_{n,\varepsilon}(a) - F_n(a+h)),$$
$$\int f\,dF_n - \int f\,dF_{n,\varepsilon} \geq \sigma h\,(F_n(a) - F_{n,\varepsilon}(a+h)).$$

Choosing $\sigma = \frac{1}{h}$, we obtain from (4.9) that the set $\Omega(a,h)$ of ε's, for which

$$F_{n,\varepsilon}(a) \leq F_n(a+h) + h \quad \text{and} \quad F_n(a) \leq F_{n,\varepsilon}(a+h) + h, \tag{4.10}$$

has μ_n-measure at least $1 - 4e^{-cn^2h^4}$. Now, given a natural number N, introduce $\Omega(h) = \cap_{i=-Nh}^{i=Nh}\Omega(ih,h)$, so that by the previous step,

$$\mu_n(\Omega(h)) \geq 1 - 4(2N+1)\,e^{-cn^2h^4}. \tag{4.11}$$

If ε is in $\Omega(h)$, then (4.10) is fulfilled for all $2N+1$ points of the form $a = ih$, $i = -N, \dots, N$. In case $a \in (-Nh, Nh)$, choose $i = -N+1, \dots, N$ such that $(i-1)h < a \leq ih$. Then, by (4.10),

$$F_{n,\varepsilon}(a) \leq F_{n,\varepsilon}(ih) \leq F_n(ih+h) + h \leq F_n(a+2h) + h,$$

and similarly $F_n(a) \leq F_{n,\varepsilon}(a+2h) + h$. In case $a < -Nh$ and N, is large enough so that $F(-(N-1)h) \leq h$, we also have

$$F_{n,\varepsilon}(a) \leq F_{n,\varepsilon}(-(N-1)h) \leq F_n(-Nh) + h \leq F_n(a) + 2h,$$

and $F_n(a) \le h \le F_{n,\varepsilon}(a+h)+h$. Finally, assume $a > Nh$ and $1-F_n((N-1)h) \le h$. Then, by (4.10),

$$F_n(a) \le F_n((N-1)h) + h \le F_{n,\varepsilon}(Nh) + 2h \le F_{n,\varepsilon}(a+h) + 2h,$$

and $F_{n,\varepsilon}(a) \le 1 \le F_n(a+h) + h$.

Thus, in all cases for all $a \in \mathbb{R}$, we obtain that $F_{n,\varepsilon}(a) \le F_n(a+2h)+2h$ and $F_n(a) \le F_{n,\varepsilon}(a+2h)+2h$, which yields $L(F_{n,\varepsilon}, F_n) \le 2h$. Hence, according to (4.11),

$$\mu_n\{\varepsilon : L(F_{n,\varepsilon}, F_n) \ge 2h\} \le 4(2N+1)\, e^{-cn^2h^4}. \tag{4.12}$$

It remains to estimate the least possible N. Using the basic assumption $|X_{jk}| \le 1$, we have

$$\int_{-\infty}^{+\infty} x^2\, dF_n(x) = \frac{1}{n^2} \sum_{j,k} \mathbf{E} X_{jk}^2 \le 1,$$

so that by Chebyshev's inequality, $1 - F_n((N-1)h) \le \frac{1}{(N-1)^2h^2} \le h$, as long as $N - 1 \ge \frac{1}{h^{3/2}}$. Then also $F_n(-(N-1)h) \le h$, so we may take $N = [\frac{1}{h^{3/2}}] + 2$. Replacing $\delta = 2h$ in (4.12), we arrive at the following sharpening of Theorem 4.1.

Theorem 4.2. *If $|X_{jk}| \le 1$, for any $\delta \in (0,1)$,*

$$\mu_n\{\varepsilon : L(F_{n,\varepsilon}, F_n) \ge \delta\} \le \frac{C}{\delta^{3/2}}\, e^{-cn^2\delta^4},$$

where C and c are absolute positive constants.

It is now easy to deduce with some absolute constant C:

Corollary 4.4. *If $|X_{jk}| \le 1$, then $\int L(F_{n,\varepsilon}, F_n)\, d\mu_n(\varepsilon) \le C\, \frac{\log n}{n^{1/2}}$.*

5. Asymptotic behaviour of weighted random spectral distributions

In this section, we consider the behaviour of the spectral distribution functions that are averaged with respect to ε's,

$$\widetilde{F}_n(x) = F_n(x,\omega) \equiv 2^{-n(n+1)/2} \sum_{\varepsilon} F_{n,\varepsilon}(x,\omega), \tag{5.1}$$

under the conditions that, for any $\tau > 0$, as $n \to \infty$,

(1) $\widetilde{L}_n(\tau) = \frac{1}{n^2}\sum_{j=1}^n \sum_{k=1}^n X_{jk}^2 I(|X_{jk}| > \tau\sqrt{n}\,) = o(1)$;

(2) $\widetilde{\Delta}_1^2 := \frac{1}{n}\sum_{j=1}^n \Big((\sigma_{n,j}^{(\tau)})^2 - (\sigma_n^{(\tau)})^2\Big)^2 = o(1)$.

The study of the behavior of the spectral distribution function $F_n(x,\omega)$ for large n is essentially equivalent to the study of the spectra of a random matrix whose entries are independent and symmetrically distributed random variables, taking at most two opposite values. Note this is a non-i.i.d. model. First, we need to prove a (random) bound for the Lévy distance between the distribution function $F_n(x,\omega)$ and $\mathcal{L}(\sigma_n\xi)$, where ξ is a random variable having a standard semi-circle law (with variance (1).

Recall the notations introduced in Sec. 1:

$$(\sigma_{nj}^{(\tau)})^2 = \frac{1}{n}\sum_{k=1}^n (X_{jk}^{(\tau)})^2, \quad (\sigma_n^{(\tau)})^2 = \frac{1}{n}\sum_{j=1}^n (\sigma_{nj}^{(\tau)})^2.$$

As before, $X_{jk}^{(\tau)}$ stands for X_{jk}, truncated at level τ. Consider the random symmetric matrices, subject to the truncation procedure,

$$\mathbf{W}^{(\tau)} = \frac{1}{\sqrt{n}}\Big(\varepsilon_{jk}X_{jk}^{(\tau)}\Big)_{j,k=1}^n, \quad \mathbf{R}^{(\tau)} = (\mathbf{W}^{(\tau)} - z\mathbf{I})^{-1},$$

and introduce the Stieltjes transform $S_n^{(\tau)}(z) = \frac{1}{n}\operatorname{Tr}(\mathbf{R}^{(\tau)})$. Let

$$\widetilde{B} = \frac{\sqrt{\widetilde{L}_n(\tau)}}{(\sigma_n^{(\tau)})^2} + 2\tau + \frac{1}{n} + \frac{\sigma_n^{(\tau)}}{\sqrt{n}} + \widetilde{\Delta}_1 + \frac{\widetilde{\Delta}_1^2}{(\sigma_n^{(\tau)})^2}.$$

Theorem 5.1. *Given $\tau > 0$ and $\omega \in \Omega$, assume that*

$$\widetilde{B} \leqslant \sigma_n^{(\tau)}. \tag{5.2}$$

Then, there exist absolute constants C and $T_0 > 1$ such that for any $T > T_0$,

$$L(\widetilde{F}_n, \mathcal{L}(\sigma_n^{(\tau)}\xi)) \leqslant C\widetilde{B}(\sigma_n^{(\tau)})^{-1}\exp\{2T\sigma_n^{(\tau)}\}\log T + \frac{C\log T}{T} + T\frac{\widetilde{L}_n(\tau)}{2\sigma_n^{(\tau)}}.$$

Proof. Let $\widetilde{f}_n$ and f denote the characteristic functions of $\widetilde{F}_n$ and the semi-circle law, respectively. By Zolotarev's Lemma 4.1, we have for all $T \geq T_0$,

$$L(\widetilde{F}_n, \mathcal{L}(\sigma_n^{(\tau)}\xi)) \leq C \int_0^T \frac{|\widetilde{f}_n(t) - f(t\sigma_n^{(\tau)})|}{t}\, dt + \frac{C \log T}{T}.$$

Note that $|\sigma_n - \sigma_n^{(\tau)}| \leq \frac{\widetilde{L}_n(\tau)}{2\sigma_n^{(\tau)}}$. This easily implies the bound

$$|f(t\sigma_n^{(\tau)}) - f(t\sigma_n)| \;\leq\; |t|\, |\sigma_n - \sigma_n^{(\tau)}| \leq |t|\, \frac{\widetilde{L}_n(\tau)}{2\sigma_n^{(\tau)}}. \tag{5.3}$$

Then, applying equality (3.1) and inequality (5.3) together, we get for any $v > 0$,

$$L(\widetilde{F}_n, \mathcal{L}(\sigma_n\xi)) \leqslant C \exp\{vT\} \log T \int_{-\infty}^{\infty} |S_n(u+iv)$$

$$-(\sigma_n^{(\tau)})^{-1} S((u+iv)({\sigma_n}^{(\tau)})^{-1})|du + \frac{\log T}{T} + T\frac{\widetilde{L}_n(\tau)}{2\sigma_n^{(\tau)}}, \tag{5.4}$$

where $S_n(z)$ and $S(z)$ denote the Stieltjes transforms of the distribution function $\widetilde{F}_n(x)$ and the semi-circle law, respectively.

To bound the right hand side of this inequality, we shall investigate the Stieltjes transform

$$S_n(z) = \frac{1}{n}\, \mathbf{E}_\varepsilon \operatorname{Tr}(\mathbf{R}(z)).$$

Theorem 5.2. *Under the condition of Theorem* 5.1, *there exists an absolute positive constant C such that*

$$\int_{-\infty}^{\infty} \left| S_n(u+iv) - (\sigma_n^{(\tau)})^{-1} S((u+iv)(\sigma_n^{(\tau)})^{-1}) \right| du \leqslant C\frac{\widetilde{B}}{v}.$$

Theorem 5.2 and inequality 5.4 together will complete the proof of Theorem 5.1. $\square$

Proof of Theorem 5.2. First we bound the difference between Stieltjes transforms of the spectrum of the matrix $\mathbf{W}$ and the one of the spectrum of the matrix $\mathbf{W}^{(\tau)}$ with truncated entries.

Lemma 5.1. *For any $v > 0$, we have*

$$\int_{-\infty}^{\infty} \left| S_n(u+iv) - S_n^{(\tau)}(u+iv) \right| du \leqslant \frac{\sqrt{2(\sigma_n^{(\tau)})^2 + v^2}}{v^{\frac{3}{2}} \sqrt{\sigma_n^{(\tau)}}} \sqrt{L_n(\tau)}.$$

The proof of this lemma is postponed to the Appendix.

Now, assuming $|X_{jk}| \leqslant \tau\sqrt{n}$, we prove:

Theorem 5.3. *Under the condition of Theorem* 5.1, *there exists an absolute positive constant C such that for all $v \geqslant 2\sigma_n^{(\tau)}$,*

$$\int_{-\infty}^{+\infty} \left| S_n^{(\tau)}(u+iv) - (\sigma_n^{(\tau)})^{-1} S((u+iv)(\sigma_n^{(\tau)})^{-1}) \right| du \leqslant \frac{C\widetilde{B}}{v}.$$

The proof of this theorem is given in Sec. 5.1. Lemma 5.1 and Theorem 5.3 together imply the result of Theorem 5.2. □

5.1. Proof of Theorem 5.3. Uniform bound

In the sequel, we assume that

$$|X_{jk}| \leqslant \tau\sqrt{n}.$$

We will omit the index τ in the notation. In this section, we show that the Stieltjes transform of the spectral distribution of the matrix $\mathbf{W}$ satisfies a certain approximate equation that characterizes the semi-circle law. Furthermore, we give a bound for the error of this approximation which is uniform in u. Then, in Sec. 5.3, using the obtained representation, we derive an integral bound for the difference between the Stieltjes transforms of the semi-circle law and the spectral distribution of the matrix.

5.1.1. The main equation

We recall the following notations. Let $\mathbf{X} = (X_{jk})_{j,k=1}^{n}$ denote a symmetric matrix of order n. Let ε_{jk}, $1 \leqslant j \leqslant k \leqslant n$, denote the Bernoulli i.i.d. random variables. Consider a symmetric random matrix $\mathbf{W}$ of order n with entries

$$W_{jk} = \frac{1}{\sqrt{n}} \varepsilon_{jk} X_{jk} \quad \text{for} \quad 1 \leqslant j \leqslant k \leqslant n.$$

Introduce the resolvent matrix $\mathbf{R} = (\mathbf{W} - z\mathbf{I}_n)^{-1}$, where $\mathbf{I}_n$ is the identity matrix of order n and $z = u + vi$ with $v > 0$.

Let $\mathbf{j}_\nu$ denote the multi-index $\mathbf{j}_\nu = (j_1, \dots, j_\nu)$ with distinct numbers $j_1, \dots, j_\nu$ from $\{1.\dots, n\}$. Introduce the matrix $\mathbf{W}^{(\mathbf{j}_\nu)}$ obtained from $\mathbf{W}$ by deleting the j_1th,…,j_νth both rows and columns. Consider also the resolvent matrix

$$\mathbf{R}^{(\mathbf{j}_\nu)} = (\mathbf{W}^{(\mathbf{j}_\nu)} - z\mathbf{I}_{n-\nu})^{-1}.$$

If $\nu = 0$, the set of indices $\{j_1, \dots, j_\nu\}$ is empty, and we may also write $\mathbf{W} = \mathbf{W}^{(\mathbf{j}_0)}$ and $\mathbf{R} = \mathbf{R}^{(\mathbf{j}_0)}$. Let $\mathbf{a}_j^{(\mathbf{j}_\nu)}$ denote the jth column without jth elements of the matrix $\mathbf{W}^{(\mathbf{j}_\nu)}$.

According to Lemma 7.1 below, we have the following equality for the diagonal entries of the resolvent matrices: For all $j \in \{1, \dots, n\} \setminus \{j_1, \dots, j_\nu\}$, $\nu = 0, 1, \dots, n-1$.

$$R_{jj}^{(\mathbf{j}_\nu)} = \frac{1}{\frac{\varepsilon_{jj} X_{jj}}{\sqrt{n}} - z - \frac{1}{n}\sum_{l \neq k}^{(\mathbf{j}_{\nu+1})} \varepsilon_{jk}\varepsilon_{jl} X_{jk} X_{jl} R_{kl}^{(\mathbf{j}_{\nu+1})} - \frac{1}{n}\sum_{l}^{(\mathbf{j}_{\nu+1})} X_{jl}^2 R_{ll}^{(\mathbf{j}_{\nu+1})}}, \quad (5.5)$$

where $\mathbf{j}_{\nu+1} = (j_1, \dots, j_\nu, j)$ and where $\sum^{(\mathbf{j}_{\nu+1})}$ indicates summation over all indices from $\{1, \dots, n\} \setminus \{j_1, \dots, j_{\nu+1}\}$. Introduce the following notations:

$$\gamma_{j,1}^{(\mathbf{j}_\nu)} = \frac{1}{\sqrt{n}} \varepsilon_{jj} X_{jj}, \quad \gamma_{j,2}^{(\mathbf{j}_\nu)} = -\frac{1}{n}\sum_{l \neq k}^{(\mathbf{j}_{\nu+1})} \varepsilon_{jk}\varepsilon_{jl} X_{jl} X_{jk} R_{lk}^{(\mathbf{j}_{\nu+1})},$$

$$\gamma_{j,3}^{(\mathbf{j}_\nu)} = -\frac{1}{n}\sum_{l}^{(\mathbf{j}_{\nu+1})} X_{jl}^2 \left(R_{ll}^{(\mathbf{j}_{\nu+1})} - \frac{1}{n} \operatorname{Tr} \mathbf{R}^{(\mathbf{j}_{\nu+1})} \right),$$

$$\gamma_{j,4}^{(\mathbf{j}_\nu)} = -\left(\frac{1}{n}\sum_{l}^{(\mathbf{j}_{\nu+1})} X_{jl}^2 \right) \left(\frac{1}{n} \operatorname{Tr} \mathbf{R}^{(\mathbf{j}_{\nu+1})} - \frac{1}{n} \operatorname{Tr} \mathbf{R}^{(\mathbf{j}_\nu)} \right)$$

$$\gamma_{j,5}^{(\mathbf{j}_\nu)} = -\left(\frac{1}{n}\sum_{l=1}^{n} X_{jl}^2 \right) \left(\frac{1}{n} \operatorname{Tr} \mathbf{R}^{(\mathbf{j}_\nu)} - \frac{1}{n} \mathbf{E}_\varepsilon \operatorname{Tr} \mathbf{R}^{(\mathbf{j}_\nu)} \right),$$

$$\gamma_{j,6}^{(\mathbf{j}_\nu)} = \left(\frac{1}{n}\sum_{p=1}^{\nu+1} X_{jj_p}^2 \right) \frac{1}{n} \operatorname{Tr} \mathbf{R}^{(\mathbf{j}_\nu)},$$

$$\gamma_{j,7}^{(\mathbf{j}_\nu)} = -\left(\frac{1}{n}\sum_{l=1}^{n} X_{jl}^2 - \sigma_n^2 \right) \frac{1}{n} \mathbf{E}_\varepsilon \operatorname{Tr} \mathbf{R}^{(\mathbf{j}_\nu)}.$$

Using these notations, we may rewrite equality (5.5) in the form

$$R_{jj}^{(\mathbf{j}_\nu)} = -\frac{1}{z+\sigma_n^2 S_n^{(\mathbf{j}_\nu,\tau)}(z)} + \frac{1}{z+\sigma_n^2 S_n^{(\mathbf{j}_\nu)}(z)}\Gamma_j R_{jj}^{(\mathbf{j}_\nu)}, \tag{5.6}$$

where

$$S_n^{(\mathbf{j}_\nu)}(z) = \frac{1}{n}\mathbf{E}_\varepsilon \operatorname{Tr}\mathbf{R}^{(\mathbf{j}_\nu)}, \quad \Gamma_j^{(\mathbf{j}_\nu)} = \sum_{s=1}^{7}\gamma_{j,s}^{(\mathbf{j}_\nu)}.$$

Taking the mean value of Eq. (5.6) with respect to both j and ε, we get

$$S_n^{(\mathbf{j}_\nu)}(z) = -\frac{1}{z+\sigma_n^2 S_n^{(\mathbf{j}_\nu)}(z)} + \frac{1}{z+\sigma_n^2 S_n^{(\mathbf{j}_\nu)}(z)}\frac{1}{n}\sum_j{}^{(\mathbf{j}_\nu)}\mathbf{E}_\varepsilon\Gamma_j^{(\mathbf{j}_\nu)}R_{jj}^{(\mathbf{j}_\nu)}. \tag{5.7}$$

In particular, for $\nu = 0$, we have

$$S_n(z) = -\frac{1}{z+\sigma_n^2 S_n(z)} + \delta_n(z), \tag{5.8}$$

where

$$\delta_n(z) = \frac{1}{n(z+\sigma_n^2 S_n(z))}\sum_{j=1}^{n}\mathbf{E}_\varepsilon\Gamma_j R_{jj}.$$

Assuming that $|X_{jk}| \leqslant \tau\sqrt{n}$, we prove the following

Theorem 5.4. *For $v \geqslant 2\sigma_n$, we have $|\delta_n(z)| \leqslant \frac{\widetilde{B}_1}{\sigma_n^2}$, where*

$$\widetilde{B}_1 = 2\tau + \frac{\sigma_n}{\sqrt{n}} + \frac{\sqrt{\Delta_1^2}}{\sigma_n} + \frac{\Delta_1^2}{\sigma_n^3} + \frac{1}{\sqrt{n}}.$$

Corollary 5.2. *Assume that $\widetilde{B}_1 \leqslant \frac{1}{2}v$. Then, for $v \geqslant \max\{1, 2\sigma_n\}$,*

$$S_n(z) = \sigma_n^{-1}S((z+\delta_n(z))\sigma_n^{-1}) + \delta_n(z)$$

where $S(z) = -\frac{z}{2} + \frac{\sqrt{z^2-4}}{2}$.

Proof. Note that for $v \geqslant \max\{1, 2\sigma_n\}$, the assumption implies $\operatorname{Im}\{z + \sigma_n^2\delta_n(z)\} > 0$. Solving equation (5.8) with respect to $S_n(z)$, we arrive at the desired representation. □

5.2. Proof of Theorem 5.4

We start with straightforward bounds for the $\gamma_{j,s}$ with $s \neq 3$. For $s = 3$, the bound will be based on a certain recurrence procedure. The bounds uniformly in $u = \Re z$ and allow us to estimate the difference between Stieltjes transforms.

Proof of 5.4. Put $A_s = \frac{1}{n}\sum_{j=1}^n \mathbf{E}_\varepsilon\,|\gamma_{j,s}|$ and note that

$$|\delta_n(z)| \leqslant \frac{1}{v^2}\frac{1}{n}\sum_{j=1}^n \mathbf{E}_\varepsilon\,|\Gamma_j| \leqslant \frac{1}{v^2}\sum_{s=1}^7 A_s.$$

Lemmas 5.3–5.9 below together imply the result. $\square$

5.2.1. Auxiliary bounds

In this subsection, we will consider the matrix $\mathbf{W}^{(\tau)}$. For simplicity of notations, we assume that, for some $\tau > 0$ and for any $1 \leqslant j, k \leqslant n$,

$$|X_{jk}| \leqslant \tau\sqrt{n}.$$

We shall estimate the $\gamma_{j,s}$, $s = 1, \ldots, 7$, error terms of the approximation of the Stieltjes transform of the distribution function $F_n(x)$ by the Stieltjes transform of the semi-circle law using several auxiliary lemmas. We start with obvious estimates.

Lemma 5.3. *We have*

$$A_1^{(\mathbf{j}_\nu)} \equiv \frac{1}{n}{\sum_j}^{(\mathbf{j}_\nu)}\mathbf{E}_\varepsilon|\gamma_{j,1}^{(\mathbf{j}_\nu)}| \leqslant \tau.$$

Proof. It is straightforward to check that

$$A_1^{(\mathbf{j}_\nu)} \leqslant \frac{1}{n^{\frac{3}{2}}}\sum_{j=1}^n |X_{jj}| \leqslant \tau.$$

Lemma 5.4.

$$A_2^{(\mathbf{j}_\nu)} \equiv \frac{1}{n}{\sum_j}^{(\mathbf{j}_\nu)}\mathbf{E}_\varepsilon|\gamma_{j,2}^{(\mathbf{j}_\nu)}| \leqslant \frac{\sigma_n\tau}{v}.$$

Proof. Applying Cauchy's inequality and the definition of $\gamma_{j,2}^{(\mathbf{j}_\nu)}$, we obtain that

$$\begin{aligned}
\frac{1}{n}\sum_j{}^{(\mathbf{j}_\nu)}\mathbf{E}_\varepsilon|\gamma_{j,2}^{(\mathbf{j}_\nu)}| &\leqslant \frac{1}{n}\sum_j{}^{(\mathbf{j}_\nu)}\mathbf{E}_\varepsilon^{\frac{1}{2}}|\gamma_{j,2}^{(\mathbf{j}_\nu)}|^2 \\
&\leqslant \frac{1}{n^2}\sum_j{}^{(\mathbf{j}_\nu)}\Bigg(\sum_{l\neq k}{}^{(\mathbf{j}_{\nu+1})}X_{jl}^2X_{jk}^2\mathbf{E}_\varepsilon|R_{kl}^{(\mathbf{j}_{\nu+1})}|^2\Bigg)^{\frac{1}{2}} \\
&\leqslant \frac{\tau}{n}\sum_j{}^{(\mathbf{j}_\nu)}\Bigg(\frac{1}{n}\sum_{l\neq k}{}^{(\mathbf{j}_{\nu+1})}X_{jl}^2\mathbf{E}_\varepsilon|R_{kl}^{(\mathbf{j}_{\nu+1})}|^2\Bigg)^{\frac{1}{2}} \\
&\leqslant \frac{\tau}{nv}\sum_j{}^{(\mathbf{j}_\nu)}\Bigg(\frac{1}{n}\sum_{l}{}^{(\mathbf{j}_{\nu+1})}X_{jl}^2\Bigg)^{\frac{1}{2}} \leqslant \frac{\tau\sigma_n}{v}.
\end{aligned}$$

Here we have used that for all $l \in \{1,\dots,n\}\setminus\{j_1,\dots,j_{\nu+1}\}$ and for all $\nu = 0,1,\dots,n$,

$$\sum_k{}^{(\mathbf{j}_{\nu+1})}|R_{lk}|^2 \leqslant \frac{1}{v^2}.$$

Lemma 5.5.

$$A_4^{(\mathbf{j}_\nu)} \equiv \frac{1}{n}\sum_j{}^{(\mathbf{j}_\nu)}\mathbf{E}_\varepsilon|\gamma_{j,4}^{(\mathbf{j}_\nu)}| \leqslant \frac{\sigma_n^2}{nv}.$$

Proof. Applying Lemma 7.2, we obtain

$$A_4^{(\mathbf{j}_\nu)} \leqslant \frac{1}{n^2v}\sum_j{}^{(\mathbf{j}_\nu)}\Bigg(\frac{1}{n}\sum_k{}^{(\mathbf{j}_{\nu+1})}X_{jk}^2\Bigg) \leqslant \frac{\sigma_n^2}{nv}.$$

Lemma 5.6.

$$A_5^{(\mathbf{j}_\nu)} \equiv \frac{1}{n}\sum_j{}^{(\mathbf{j}_\nu)}\mathbf{E}_\varepsilon|\gamma_{j,5}^{(\mathbf{j}_\nu)}| \leqslant \frac{\sigma_n}{\sqrt{n}v}.$$

Proof. Applying Lemma 7.3, we get

$$A_5^{(\mathbf{j}_\nu)} \leqslant \frac{1}{n}\sum_j{}^{(\mathbf{j}_\nu)}\frac{1}{n}\sum_k{}^{(\mathbf{j}_{\nu+1})}X_{jk}^2\mathbf{E}_\varepsilon^{\frac{1}{2}}\left|\frac{1}{n}(\operatorname{Tr}\mathbf{R}^{(\mathbf{j}_\nu)} - \mathbf{E}_\varepsilon\operatorname{Tr}\mathbf{R}^{(\mathbf{j}_\nu)})\right|^2 \leqslant \frac{2\sigma_n^2}{\sqrt{n}v}.$$

Lemma 5.7.

$$A_6^{(\mathbf{j}_\nu)} \equiv \frac{1}{n}\sum_j^{(\mathbf{j}_\nu)} \mathbf{E}_\varepsilon|\gamma_{j,6}^{(\mathbf{j}_\nu)}| \leqslant \frac{(\nu+1)\sigma_n^2}{nv} + \frac{\sqrt{\Delta_1^2}}{v}.$$

Proof. The definition of $\gamma_{j,6}^{(\mathbf{j}_\nu)}$ and the inequality $|\frac{1}{n}\mathrm{Tr}\,\mathbf{R}^{(\mathbf{j}_\nu)}| \leqslant \frac{1}{v}$ together imply

$$A_6^{(\mathbf{j}_\nu)} \leqslant \frac{1}{nv}\sum_{p=1}^{\nu}\frac{1}{n}\sum_{j=1}^{n} X_{jj_p}^2 \leqslant \frac{(\nu+1)\sigma_n^2}{nv} + \frac{1}{nv}\sum_{p=1}^{\nu}\left|\frac{1}{n}\sum_{j=1}^{n} X_{jj_p}^2 - \sigma_n^2\right|.$$

Finally, we state a simple bound for $A_7^{(\mathbf{j}_\nu)} \equiv \frac{1}{n}\sum_j^{(\mathbf{j}_\nu)} \mathbf{E}|\gamma_{j,7}^{(\mathbf{j}_\nu)}|$.

Lemma 5.8.

$$A_7^{(\mathbf{j}_\nu)} \leqslant \frac{1}{nv}\sum_{j=1}^{n}\left|\frac{1}{n}\sum_{k=1}^{n} X_{jk}^2 - \sigma_n^2\right| \leqslant \frac{\sqrt{\Delta_1^2}}{v}.$$

The proof follows from the definition of $\gamma_{j,7}^{(\mathbf{j}_\nu)}$.

5.2.2. The bound on A_3

We prove the following:

Lemma 5.9. *The inequality*

$$A_3 = \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}|\gamma_{j,3}| \leqslant 4\widetilde{B}_1,$$

holds for $v \geqslant 2\sigma_n$ with

$$\widetilde{B}_1 = 2\tau + \left(\frac{\sigma_n}{\sqrt{n}} + 1\right)\frac{1}{\sqrt{n}} + \frac{1}{\sigma_n^3}\Delta_1^2 + \frac{1}{\sigma_n}\sqrt{\Delta_1^2}.$$

Proof. Introduce the following quantity

$$\beta_j^{(\nu)} = \max_{\mathbf{j}_\nu:\ j\notin\{j_1,\dots,j_\nu\}} \mathbf{E}_\varepsilon|\gamma_{j,3}^{(\mathbf{j}_\nu)}|.$$

For $\nu=0$, we define

$$\beta_j^{(0)}=\mathbf{E}_\varepsilon|\gamma_{j,3}|.$$

Using equality (5.6), we get

$$\mathbf{E}_\varepsilon|\gamma_{j,3}^{(\mathbf{j}_\nu)}|\leqslant\frac{1}{n}\sum_k^{(\mathbf{j}_{\nu+1})}X_{jk}^2\frac{1}{n}\sum_l^{(\mathbf{j}_{\nu+1})}\left|R_{kk}^{\mathbf{j}_{\nu+1}}-R_{ll}^{\mathbf{j}_{\nu+1}}\right|$$

$$\leqslant\frac{1}{v^2}\frac{1}{n}\sum_k^{(\mathbf{j}_{\nu+1})}X_{jk}^2\frac{1}{n}\sum_l^{(\mathbf{j}_{\nu+1})}\left(\mathbf{E}_\varepsilon\left|\Gamma_l^{(\mathbf{j}_{\nu+1})}\right|+\mathbf{E}_\varepsilon\left|\Gamma_k^{(\mathbf{j}_{\nu+1})}\right|\right).$$

Note that

$$\frac{1}{n}\sum_l^{(\mathbf{j}_\nu)}\mathbf{E}_\varepsilon\left|\Gamma_l^{(\mathbf{j}_{\nu+1})}\right|$$

$$\leqslant\max_{\mathbf{j}_\nu:\,j\notin\{j_1,\dots,j_\nu\}}\left\{\sum_{s=1,\ s\neq 3}^{7}\frac{1}{n}\sum_l^{(\mathbf{j}_\nu)}\mathbf{E}|\gamma_{j,s}^{(\mathbf{j}_{\nu+1})}|\right\}+\frac{1}{n}\sum_l^{(\mathbf{j}_\nu)}\beta_l^{(\nu+1)}.$$

According to Lemmas 5.3–5.8, we obtain

$$\max_{\mathbf{j}_\nu:\,j\notin\{j_1,\dots,j_\nu\}}\left\{\sum_{s\neq 3}\frac{1}{n}\sum_l^{(\mathbf{j}_\nu)}\mathbf{E}|\gamma_{j,s}^{(\mathbf{j}_{\nu+1})}|\right\}\leqslant B_\nu,$$

where

$$B_\nu=\left(1+\frac{\sigma_n}{v}\right)\tau+\frac{\sigma_n}{\sqrt{n}v}\left(\frac{(\nu+2)\sigma_n}{\sqrt{n}}+1\right)+\frac{\sqrt{\Delta_1^2}}{v}.$$

Then

$$\frac{1}{n}\sum_j^{(\mathbf{j}_\nu)}\beta_j^{(\nu)}\leqslant\frac{\sigma_n^2}{v^2}B_\nu+\frac{1}{v^2}\frac{1}{n^2}\sum_j^{(\mathbf{j}_\nu)}\sum_k^{(\mathbf{j}_{\nu+1})}X_{jk}^2\frac{1}{n}\sum_l^{(\mathbf{j}_{\nu+1})}\beta_l^{(\nu+1)}$$

$$+\frac{1}{v^2}\frac{1}{n^2}\sum_j^{(\mathbf{j}_\nu)}\sum_k^{(\mathbf{j}_{\nu+1})}X_{jk}^2\beta_k^{(\nu+1)},\tag{5.9}$$

In (5.9), we note the estimates

$$\frac{1}{n}\sum_j^{(\mathbf{j}_\nu)}\beta_j^{(\nu)}\leqslant\frac{\sigma_n^2}{v^2}B_\nu+\frac{1}{v^2}\frac{1}{n^2}\sum_j^{(\mathbf{j}_\nu)}\sum_k^{(\mathbf{j}_{\nu+1})}X_{jk}^2\frac{1}{n}\sum_l^{(\mathbf{j}_{\nu+1})}\beta_l^{(\nu+1)}$$

$$+\frac{1}{v^2}\frac{1}{n}\sum_k^{(\mathbf{j}_\nu)}\left(\frac{1}{n}\sum_{j\neq k}^{(\mathbf{j}_\nu)}X_{jk}^2\right)\beta_k^{(\nu+1)}.$$

Thus we may rewrite the previous inequality as

$$\frac{1}{n}\sum_j^{(\mathbf{j}_\nu)}\beta_j^{(\nu)}\leqslant\frac{\sigma_n^2}{v^2}B_\nu+\frac{1}{v^2}\frac{1}{n^2}\sum_j^{(\mathbf{j}_\nu)}\sum_k^{(\mathbf{j}_{\nu+1})}X_{jk}^2\left\{\frac{1}{n}\sum_l^{(\mathbf{j}_\nu)}\beta_l^{(\nu+1)}\right\}$$

$$+\frac{1}{v^2}\frac{1}{n}\sum_k^{(\mathbf{j}_\nu)}\left(\frac{1}{n}\sum_{j\neq k}^{(\mathbf{j}_\nu)}X_{jk}^2\right)\beta_k^{(\nu+1)}. \tag{5.10}$$

Note that $\beta_k^{(\nu+1)}$ in the second term on the right-hand side of (5.10), does not depend on j and the quantity $\frac{1}{n}\sum_l^{(\mathbf{j}_\nu)}\beta_l^{(\nu+1)}$ in the first term does not depend on j and k, respectively. We also note that

$$\beta_j^{(\nu)}\leqslant\frac{\sigma_n^2}{v}+\frac{1}{v}\left|\frac{1}{n}\sum_{l=1}^{n}X_{kl}^2-\sigma_n^2\right|. \tag{5.11}$$

Inequalities (5.10) and (5.11), and the above remark imply that

$$\frac{1}{n}\sum_j^{(\mathbf{j}_\nu)}\beta_j^{(\nu)}\leqslant\frac{\sigma_n^2}{v^2}B_\nu+\frac{2\sigma_n^2}{v^2}\frac{1}{n}\sum_l^{(\mathbf{j}_\nu)}\beta_l^{(\nu+1)}+\frac{2\sigma_n^2}{v^3}\sqrt{\Delta_1^2}+\frac{2}{v^3}\Delta_1^2.$$

Put

$$\widetilde{B}_\nu=B_\nu+\frac{2\sqrt{\Delta_1^2}}{v}+\frac{2}{v\sigma_n^2}\Delta_1^2.$$

Using this notation, we have

$$\frac{1}{n}\sum_j^{(\mathbf{j}_\nu)}\beta_j^{(\nu)}\leqslant\frac{\sigma_n^2}{v^2}\widetilde{B}_\nu+\frac{2\sigma_n^2}{v^2}\frac{1}{n}\sum_l^{(\mathbf{j}_\nu)}\beta_l^{(\nu+1)}.$$

Since the last inequality does not depend on $\mathbf{j}_\nu$, we may write

$$\frac{1}{n}\sum_{j=1}^{n}\beta_j^{(\nu)}\leqslant\frac{\sigma_n^2}{v^2}\widetilde{B}_\nu+\frac{2\sigma_n^2}{v^2}\frac{1}{n}\sum_{l=1}^{n}\beta_l^{(\nu+1)}.$$

For $\nu = 0, 1, \dots, n$, $n \geqslant 1$, introduce the quantity $D_{n\nu} = \frac{1}{n}\sum_{j=1}^{n} \beta_j^{(\nu)}$. Then inequality (5.9) may be rewritten as

$$D_{n,\nu} \leqslant \frac{2\sigma_n^2}{v^2}\widetilde{B}_\nu + \frac{2\sigma_n^2}{v^2} D_{n,\nu+1}.$$

Note that $D_{n,n} = 0$. We may take $v^2 \geqslant 4\sigma_n^2$. Then we get $D_{n,\nu} \leqslant \frac{1}{2}\widetilde{B}_\nu + \frac{1}{2}D_{n,\nu+1}$, which, for $v \geqslant 2\sigma_n$, implies that $D_{n,0} \leqslant 4\widetilde{B}_1$, where

$$\widetilde{B}_1 = 2\tau + \left(\frac{\sigma_n}{\sqrt{n}} + 1\right)\frac{1}{\sqrt{n}} + \frac{1}{\sigma_n^3}\Delta_1^2 + \frac{1}{\sigma_n}\sqrt{\Delta_1^2}.$$

Finally, we note that

$$\frac{1}{n}\sum_{j=1}^{n} \mathbf{E}|\gamma_{j,3}| \leqslant \frac{1}{n}\sum_{j=1}^{n} \beta_j^{(0)}.$$

The last inequalities together imply that $\frac{1}{n}\sum_{j=1}^{n} \mathbf{E}|\gamma_{j,3}| \leqslant 4\widetilde{B}_1$. □

5.3. Proof of Theorem 5.4. An integral bound

In this section, we prove Theorem 5.4. In particular, we obtain some integral bounds for the difference between the Stieltjes transforms of the spectral distribution function and of the semi-circle law. As a simple corollary, we get bounds for the distance between the corresponding characteristic functions.

First we consider the matrix $\mathbf{W}^{(\tau)}$, but for simplicity of notation, we omit the symbol τ, assuming that $|X_{jk}| \leqslant \tau\sqrt{n}$ for all $j,\ k$. Note that according to Corollary 5.2, we may write the solution of Eq. (5.8) in the form

$$S_n(z) = \frac{1}{\sigma_n} S\left(\frac{z + \sigma_n^2 \delta_n(z)}{\sigma_n}\right) + \delta_n(z). \tag{5.12}$$

Using $|S'(z)| \leqslant \frac{1}{v}$, this implies that, for $v \geqslant 2\sigma_n$,

$$\left|S_n(z) - \frac{1}{\sigma_n} S\Big(\frac{z}{\sigma_n}\Big)\right| \leqslant (1 + \sigma_n^{-1})\,|\delta_n(z)|$$

and

$$\int_{-\infty}^{\infty} \left|S_n(z) - \sigma_n^{-1} S\left(z\sigma_n{}^{-1}\right)\right| du \leqslant (1 + \sigma_n^{-1}) \int_{-\infty}^{\infty} |\delta_n(u + iv)| du.$$

We return to Eq. (5.7) and to the definition $\delta_n(z)$ as

$$\delta_n(z) = \frac{1}{n(z+\sigma_n^2 S_n(z))} \sum_{j=1}^n \mathbf{E}_\varepsilon \Gamma_j R_{jj},$$

where $\Gamma_j = \sum_{p=1}^7 \gamma_{j,p}$. Integrating, we get

$$\int_{-\infty}^{\infty} |\delta_n(z)|\, du \leqslant \frac{1}{n} \sum_{j=1}^n \mathbf{E}_\varepsilon \int_{-\infty}^{\infty} \left| \Gamma_j \frac{R_{jj}}{z+\sigma_n^2 S_n(z)} \right| du$$

$$\leqslant \frac{1}{n} \sum_{j=1}^n \sum_{s=1}^7 \int_{-\infty}^{\infty} \frac{1}{|z+\sigma_n^2 S_n(z)|} \mathbf{E}_\varepsilon \, |\gamma_{j,s}| \, |R_{jj}(u+iv)| du.$$

Introduce the quantities

$$\widetilde{A}_{j,s}^{(\mathbf{j}_\nu)} = \frac{1}{n} \sum_j{}^{(\mathbf{j}_\nu)} \int_{-\infty}^{\infty} \frac{1}{|z+\sigma_n^2 S_n(z)|} \mathbf{E}_\varepsilon \, |\gamma_{j,s}| \, |R_{jj}(u+iv)| du.$$

Note that

$$\int_{-\infty}^{\infty} \frac{|R_{jj}(u+iv)| du}{|z+\sigma_n^2 S_n(z)|} \leqslant$$

$$\frac{1}{2} \left(\int_{-\infty}^{\infty} |R_{jj}(u+iv)|^2 du + \int_{-\infty}^{\infty} \frac{du}{|z+\sigma_n^2 S_n(z)|^2} \right).$$

Using the representation of the diagonal entries of the matrix $\mathbf{R}$ via eigenvalues $\lambda_1, \dots, \lambda_n$ and eigenvectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ of the matrix $\mathbf{W}$, we obtain

$$\int_{-\infty}^{\infty} |R_{jj}(z)|^2 du \leqslant \sum_{k=1}^n u_{jk}^2 \int_{-\infty}^{\infty} \frac{1}{(\lambda_j - u)^2 + v^2} du \leqslant \frac{\pi}{v}.$$

Equation (5.12) implies the following inequality

$$\int_{-\infty}^{\infty} \frac{1}{|z+\sigma_n^2 S_n(z)|^2} du$$

$$\leqslant 2 \int_{-\infty}^{\infty} |S_n(z)|^2 du + 2 \int_{-\infty}^{\infty} \frac{\widetilde{B}^2}{v^2 |z+\sigma_n^2 S_n(z)|^2} du.$$

It is straightforward to check that

$$\int_{-\infty}^{\infty} |S_n(z)|^2 du \leqslant \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{(u-x)^2+v^2} du dF_n(x) \leqslant \frac{\pi}{v}.$$

According to condition 5.2 of Theorem 5.1, we have $|\frac{\widetilde{B}^2}{v^2}| \leqslant \frac{1}{4}$. It then follows that

$$\int_{-\infty}^{\infty} \frac{1}{|z+\sigma_n^2 S_n(z)|^2} du \leqslant \frac{C\pi}{v} \tag{5.13}$$

for $z = u+iv$ with $v \geqslant 2\sigma_n$. Using these bounds, similarly to Lemmas 5.3–5.8, we show that for $s \neq 3$

$$\sum_{s\neq 3} \widetilde{A}_s \leqslant \frac{\widetilde{B}}{v}. \tag{5.14}$$

Applying the same argument as in the proof of Lemma 5.9, we show that, for $v \geqslant 2\sigma_n$, $\widetilde{A}_3 \leqslant \frac{C\widetilde{B}}{v}$. From these inequalities it follows that

$$\int_{-\infty}^{\infty} |S_n(z) - \sigma_n^{-1} S(z\sigma_n^{-1}|\, du \leqslant \frac{C\widetilde{B}}{v}.$$

The last inequality completes the proof of Theorem 5.4.

6. Proof of Theorem 1.1

By Theorem 1.2, it remains to show that $F_n(x) \to \mathbf{E}_{\mathbf{X}}\widetilde{G}(x)$, as $n \to \infty$. Here $\widetilde{G}(x) = G(x\sigma^{-1})$. Recall that $F_n(x) = \mathbf{E}\ \widetilde{F}_n(x)$. The latter expectation may be splitted into the three integrals such that, whenever $0 < m < M$,

$$\begin{aligned}&\mathbf{E}\ \widetilde{F}_n(x)\\ &\quad= \mathbf{E}\ \widetilde{F}_n(x) I_{\{m\leqslant\sigma_n\leqslant M\}} + \mathbf{E}\ \widetilde{F}_n(x) I_{\{\sigma_n<m\}} + \mathbf{E}\ \widetilde{F}_n(x) I_{\{M<\sigma_n\}}.\end{aligned} \tag{6.1}$$

Without loss of generality we will assume $\mathbf{P}\{\sigma^2 = 0\} = 0$. Let

$$\psi_n(\mathbf{X}) = L(\widetilde{F}_n, \widetilde{G}_n),$$

where $\widetilde{G}_n(x) = G(x\sigma_n^{-1})$, $x \in \mathbb{R}$. By the definition of the Lévy distance, we have

$$\begin{aligned}\mathbf{E}\ \widetilde{F}_n(x) &\geqslant \mathbf{E}\ \widetilde{F}_n(x) I_{\{m\leqslant\sigma_n\leqslant M\}}\\ &\geqslant \mathbf{E}\ \widetilde{G}_n(x-\psi_n(\mathbf{X}))I_{\{m\leqslant\sigma_n\leqslant M\}} - \mathbf{E}\ \psi_n(\mathbf{X})I_{\{m\leqslant\sigma_n\leqslant M\}}\end{aligned}$$

Note that $\sup_x |(\widetilde{G}_n'(x))| \leqslant C\sigma_n^{-1}$ with some absolute positive constant C. This implies

$$\begin{aligned}\mathbf{E}\ \widetilde{F}_n(x) &\geqslant \mathbf{E}\ \widetilde{G}_n(x)I_{\{m\leqslant\sigma_n\leqslant M\}} - \mathbf{E}\ (1+C\sigma_n^{-1})\psi_n(\mathbf{X})I_{\{m\leqslant\sigma_n\leqslant M\}}\\ &\geqslant \mathbf{E}\ \widetilde{G}_n(x)I_{\{m\leqslant\sigma_n\leqslant M\}} - (1+Cm^{-1})\mathbf{E}\ \psi_n(\mathbf{X})I_{\{m\leqslant\sigma_n\leqslant M\}}. \qquad (6.2)\end{aligned}$$

Furthermore,

$$\begin{aligned}\limsup_{n\to\infty} \mathbf{E}\ \psi_n(\mathbf{X})I_{\{m\leqslant\sigma_n\leqslant M\}}\\ \leqslant \limsup_{n\to\infty} \mathbf{E}\ \psi_n(\mathbf{X})I_{\{m\leqslant\sigma_n\leqslant M\}}I_{\{\widetilde{B}\leqslant\sigma_n\}} +\\ \limsup_{n\to\infty} \mathbf{E}\ \psi_n(\mathbf{X})I_{\{m\leqslant\sigma_n\leqslant M\}}I_{\{\widetilde{B}>\sigma_n\}}.\end{aligned}$$

According to Theorem 5.1, for any $T > T_0$,

$$\begin{aligned}\limsup_{n\to\infty} \mathbf{E}\ \psi_n(\mathbf{X})I_{\{m\leqslant\sigma_n\leqslant M\}}I_{\{\widetilde{B}\leqslant\sigma_n\}}\\ \leqslant Cm^{-1}\limsup_{n\to\infty} \mathbf{E}\widetilde{B}\exp\{2TM\}\log T + \frac{C\log T}{T} + \frac{T}{m}\limsup_{n\to\infty}\mathbf{E}\widetilde{L}_n(\tau).\end{aligned}$$

Since $\limsup_{n\to\infty} \mathbf{E}\widetilde{B} \leqslant \tau$, we obtain for any $\tau > 0$ and for any $T > T_0$,

$$\limsup_{n\to\infty} \mathbf{E}\ \psi_n(\mathbf{X})I_{\{m\leqslant\sigma_n\leqslant M\}}I_{\{\widetilde{B}\leqslant\sigma_n\}} \leqslant C(T,M,m)\ \tau + \frac{C\log T}{T},$$

where $C(T,M) = C\log T\exp\{2TM\}$. The left-hand side of the last inequality does not depend on τ and T. In the limit with $\tau \to 0$ and $T \to \infty$ we get

$$\limsup_{n\to\infty} \mathbf{E}\ \psi_n(\mathbf{X})I_{\{m\leqslant\sigma_n\leqslant M\}}I_{\{\widetilde{B}\leqslant\sigma_n\}} = 0. \qquad (6.3)$$

Relations (6.2) and (6.3) yield

$$\liminf_{n\to\infty} \mathbf{E}\ \widetilde{F}_n(x) \geqslant \liminf_{n\to\infty} \mathbf{E}\ \widetilde{G}_n(x)I_{\{m\leqslant\sigma_n\leqslant M\}}.$$

Since σ_n converges weakly in distribution to σ as $n \to \infty$ we obtain

$$\liminf_{n\to\infty} \mathbf{E}\ \widetilde{F}_n(x) \geqslant \mathbf{E}\ \widetilde{G}(x) I_{\{m \leqslant \sigma \leqslant M\}},$$

In the limit with $m \to 0$ and $M \to \infty$ we get

$$\liminf_{n\to\infty} \mathbf{E}\ \widetilde{F}_n(x) \geqslant \mathbf{E}\ \widetilde{G}(x). \tag{6.4}$$

Representation (6.1) yields the following inequality

$$F_n(x) \leqslant \mathbf{E}\ \widetilde{F}_n(x) I_{\{m \leqslant \sigma_n \leqslant M\}} + \mathbf{E}\ I_{\{\sigma_n < m\}} + \mathbf{E}\ I_{\{M < \sigma_n\}}.$$

From here, using the same argument as above, we get

$$\limsup_{n\to\infty} \mathbf{E}\ \widetilde{F}_n(x) \leqslant \mathbf{E}\ \widetilde{G}(x). \tag{6.5}$$

Relations (6.4) and (6.5) together complete the proof.

7. Appendix

7.1. Auxiliary lemmas

In order to make the paper self-contained, we collect here some auxiliary lemmas similar to those used in [2].

Lemma 7.1. *Let $\mathbf{A} = (a_{kj})$ denote a nondegenerate matrix of order n with inverse $\mathbf{A}^{-1} = \left(a^{jk}\right)$ and $\mathbf{A}_k$ its nondegenerate major sub-matrix of order $n-1$. Let α_k denote the vector obtained from the kth row of $\mathbf{A}$ by removing the kth entry and β_k the vector obtained from the kth column by removing the kth entry. Then,*

$$a^{kk} = \frac{1}{a_{kk} - \alpha_k' \mathbf{A}_k^{-1} \beta_k}.$$

Proof. Consider the obvious equality

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix},$$

which implies

$$\det \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \det(\mathbf{A}) \det\left(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right).$$

Since
$$a^{kk} = \frac{\det(\mathbf{A}_k)}{\det(\mathbf{A})},$$
the above equality with $\mathbf{A} = \mathbf{A}_k$, $\mathbf{D} = a_{kk}$, $\mathbf{C} = \alpha_k$ and $\mathbf{B} = \beta_k$ yields the result. □

As a trivial corollary of the Sturmian separation theorem (cf.,e.g., [4, Chap. 7, Theorem 4]), we have following:

Lemma 7.2. *Let $z = u + iv$, and $\mathbf{A}$ be an $n \times n$ symmetric matrix. Then*
$$\left|\mathrm{Tr}\ (\mathbf{A} - z\mathbf{I}_n)^{-1} - \mathrm{Tr}\ (\mathbf{A}_k - z\mathbf{I}_{n-1})^{-1}\right| \leqslant v^{-1}.$$

Proof. Consider a nonsingular block matrix
$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11}\mathbf{S}_{12} \\ \mathbf{S}_{21}\mathbf{S}_{22} \end{bmatrix}.$$
Applying the Schur complements formula ([13, Chap. 08, p. 21]) with $\mathbf{S}_{11} = \mathbf{A}_k - z\mathbf{I}_{n-1}$, $\mathbf{S}_{21} = \alpha_k$, $\mathbf{S}_{12} = \alpha_k'$, and $\mathbf{S}_{22} = a_{kk} - z$, a direct calculation yields
$$\mathrm{Tr}\,(\mathbf{A} - z\mathbf{I}_n)^{-1} - \mathrm{Tr}\,(\mathbf{A}_k - z\mathbf{I}_{n-1})^{-1} = \frac{1 + \alpha_k'(\mathbf{A}_k - z\mathbf{I}_{n-1})^{-2}\alpha_k}{a_{kk} - \alpha_k'(\mathbf{A}_k - z\mathbf{I}_{n-1})^{-1}\alpha_k}.$$
Let T be an orthogonal transformation which transforms $\mathbf{A}$ into diagonal form. Denote by $\mu_1 \leqslant \cdots \leqslant \mu_{n-1}$ the eigenvalues of $\mathbf{A}_k$ and let $(y_1, \ldots, y_{n-1}) = \alpha_k' T'$. Then
$$|1 + \alpha_k'(\mathbf{A}_k - z\mathbf{I}_{n-1})^{-2}\alpha_k| = \left|1 + \sum_{l=1}^{n-1} y_l^2(\mu_l - z)^{-2}\right|$$
$$\leqslant 1 + \sum_{l=1}^{n-1} y_l^2\left((\mu_l - u)^2 + v^2\right)^{-1} \leqslant 1 + \alpha_k'\left((A_k - u\mathbf{I}_{n-1})^2 + v^2\mathbf{I}_{n-1}\right)^{-1}\alpha_k.$$
Since for any commuting matrices $\mathbf{A}$, $\mathbf{B}$, such that $\mathbf{A}^2 + \mathbf{B}^2$ is nondegenerate,
$$(\mathbf{A} + i\mathbf{B})^{-1} = (\mathbf{A} - i\mathbf{B})(\mathbf{A}^2 + \mathbf{B}^2)^{-1},$$
we can directly verify that
$$\mathrm{Im}\left(a_{kk} - z - \alpha_k'\,(A - z\mathbf{I}_{n-1})^{-1}\,\alpha_k\right)$$
$$= -v\left(1 + \alpha_k'\left((A - u\mathbf{I}_{n-1})^2 + v^2\mathbf{I}_{n-1}\right)^{-1}\alpha_k\right).$$
The last two relations together imply the result. □

7.2. Estimation of the variance of the Stieltjes transforms

In this section, we give a general bound for the variance of $\mathrm{Tr}\,(\mathbf{R}(z))$, $z = u + iv$, without restrictions on the moments of the matrix entries under the assumption of independence of entries. Let

$$V_n^2 = \frac{1}{n}\,\mathbf{E}\,|\mathrm{Tr}\,(\mathbf{R}(z)) - \mathbf{E}\mathrm{Tr}\,(\mathbf{R}(z))|^2.$$

The bound for the last quantity gives:

Lemma 7.3. *For any $v > 0$,*

$$V_n^2 \leqslant \frac{4}{nv^2}. \tag{7.2}$$

Proof. We apply the martingale decomposition for the difference $\mathrm{Tr}\,(\mathbf{R}(z)) - \mathbf{E}\,\mathrm{Tr}\,(\mathbf{R}(z))$ developed in [10, p. 9]. Let $\mathbf{E}_k$ denote the conditional expectation given the σ-algebra $\mathcal{F}_k = \sigma\{\varepsilon_{ij} : k+1 \leqslant i \leqslant j \leqslant n\}$, where ε_{ij}, $i, j = 1, \dots, n$, denote i.i.d. Bernoulli random variables. Introduce the $(n-1)\times(n-1)$ matrix $\mathbf{W}^{(k)}$ obtained from $\mathbf{W}$ by deleting the kth row and column. Set $\mathbf{R}^{(k)}(z) = (\mathbf{W}^{(k)} - z\mathbf{I}_{n-1})^{-1}$. Let

$$\alpha_k = \mathbf{E}_{k-1}\,\mathrm{Tr}\,(\mathbf{R}(z)) - \mathbf{E}_k\,\mathrm{Tr}\,(\mathbf{R}(z)) = \mathbf{E}_{k-1}\varkappa_k - \mathbf{E}_k\varkappa_k, \tag{7.3}$$

where

$$\varkappa_k = \mathrm{Tr}\,(\mathbf{R}(z)) - \mathrm{Tr}\,(\mathbf{R}^{(k)}(z)).$$

Equation (7.3) follows since $\mathbf{E}\{\mathrm{Tr}\,(\mathbf{R}^{(k)}(z))\big|\mathcal{F}_k\} = \mathbf{E}\{\mathrm{Tr}\,(\mathbf{R}^{(k)}(z))\big|\mathcal{F}_{k-1}\}$.

Applying Lemma 7.2 with $\mathbf{A} = \mathbf{W}$ and $\mathbf{A}^{(k)} = \mathbf{W}^{(k)}$ for symmetric matrices, we get

$$|\varkappa_k| \leqslant \frac{1}{v}.$$

This immediately implies that

$$|\gamma_d| \leqslant \frac{2}{v}.$$

Since the martingale differences γ_d represent uncorrelated random variables for $d \le n$, and $\mathrm{Tr}\,(\mathbf{R}(z)) - \mathbf{E}\,\mathrm{Tr}\,(\mathbf{R}(z)) = \sum_{d=1}^{n}\gamma_d$, we obtain the inequality

$$V_n^2 = \frac{1}{n^2}\,\mathbf{E}\,|\mathrm{Tr}\,(\mathbf{R}(z)) - \mathbf{E}\,\mathrm{Tr}\,(\mathbf{R}(z))|^2 \leqslant \frac{4}{nv^2},$$

which completes the proof. □

7.3.Truncation

Here we consider the difference between the Stieltjes transforms of the spectral distributions of the matrix $\mathbf{W}$ and the truncated matrix. In what follows we shall use notation

$$|\mathbf{A}|^2 = \mathbf{A}\mathbf{A}^*,$$

where $\mathbf{A}^*$ denotes complex conjugate matrix $\mathbf{A}$. We start again with some obvious bounds.

Lemma 7.4. *For all λ real and $v > 0$,*

$$\sup_u \left\{ \frac{u^2}{(u-\lambda)^2 + v^2} \right\} \leqslant \frac{\lambda^2 + v^2}{v^2}.$$

Proof. The function $f(u) = \frac{u^2}{(u-\lambda)^2+v^2}$ the first derivative

$$f'(u) = 2u \, \frac{-\lambda u + \lambda^2 + v^2}{((u-\lambda)^2 + v^2)^2},$$

with critical points at $u_0 = 0$ and $u_1 = \frac{\lambda^2+v^2}{\lambda}$. It is easy to see that

$$\sup_u \left\{ \frac{u^2}{(u-\lambda)^2 + v^2} \right\} = f(u_1) = \frac{u^2 + v^2}{v^2}.$$

Thus the Lemma is proved. □

Lemma 7.5.

$$\int_{-\infty}^{\infty} \sqrt{\frac{1}{n} \mathbf{E}_\varepsilon \operatorname{Tr} (|\mathbf{R}|^4)} \, du \leqslant \frac{\pi}{v^{\frac{3}{2}}} \frac{\sqrt{2\sigma_n^2 + v^2}}{\sqrt{\sigma_n}}.$$

Proof. Consider the following equalities

$$\begin{aligned} \int_{-\infty}^{\infty} \sqrt{\frac{1}{n} \mathbf{E}_\varepsilon \operatorname{Tr} |\mathbf{R}|^4} &= \int_{-\infty}^{\infty} \sqrt{\frac{1}{n} \sum_{j=1}^{n} \mathbf{E}_\varepsilon \frac{1}{((u-\lambda_j)^2 + v^2)^2}} du \\ &= \int_{-\infty}^{\infty} \sqrt{\frac{1}{n} \sum_{j=1}^{n} \mathbf{E}_\varepsilon \frac{u^2 + \sigma_n^2}{((u-\lambda_j)^2 + v^2)^2}} \frac{du}{\sqrt{u^2 + \sigma_n^2}}. \end{aligned}$$

Applying Hölder inequality, we get

$$\Big|\int\limits_{-\infty}^{\infty}\sqrt{\frac{1}{n}\mathbf{E}_{\varepsilon}\operatorname{Tr}|\mathbf{R}|^4}du\Big|$$
$$\leqslant\Big(\int\limits_{-\infty}^{\infty}\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}_{\varepsilon}\frac{u^2+\sigma_n^2}{((u-\lambda_j)^2+v^2)^2}du\Big)^{\frac12}\Big(\int\limits_{-\infty}^{\infty}\frac{du}{u^2+\sigma_n^2}\Big)^{\frac12}. \quad (7.5)$$

By Lemma 7.4, we have

$$\int\limits_{-\infty}^{\infty}\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}_{\varepsilon}\frac{u^2+\sigma_n^2}{((u-\lambda_j)^2+v^2)^2}du$$
$$\leqslant\frac{\pi}{nv}\sum_{j=1}^{n}\mathbf{E}_{\varepsilon}\frac{\lambda_j^2+v^2+\sigma_n^2}{v^2}=\frac{\pi}{v}\frac{2\sigma_n^2+v^2}{v^2}. \quad (7.6)$$

Inequalities (7.5) and (7.6) together imply the Lemma. □

We prove the following

Lemma 7.6. *For any $v>0$, the following inequality holds,*

$$\int\limits_{-\infty}^{\infty}\left|S_n(u+iv)-S_n^{(\tau)}\right|du\leqslant\frac{\sqrt{2\sigma_n^2+v^2}}{v^{\frac32}\sqrt{\sigma_n}}\sqrt{L_n(\tau)}.$$

Proof. Applying the resolvent equality

$$(\mathbf{A}+\mathbf{B}-z\mathbf{I})^{-1}=(\mathbf{A}-z\mathbf{I})^{-1}+(\mathbf{A}-z\mathbf{I})^{-1}\mathbf{B}(\mathbf{A}-z\mathbf{I})^{-1},$$

we get

$$\left|\frac{1}{n}\mathbf{E}_{\varepsilon}\operatorname{Tr}\mathbf{R}-\frac{1}{n}\mathbf{E}_{\varepsilon}\operatorname{Tr}\mathbf{R}^{(\tau)}\right|\leqslant\frac{1}{n}\mathbf{E}_{\varepsilon}|\operatorname{Tr}\mathbf{R}^{(\tau)}\mathbf{W}^{(\tau)}\mathbf{R}|. \quad (7.7)$$

Using Cauchy's inequality, we get

$$\left|\frac{1}{n}\mathbf{E}_{\varepsilon}\operatorname{Tr}\mathbf{R}-\frac{1}{n}\mathbf{E}_{\varepsilon}\operatorname{Tr}\mathbf{R}^{(\tau)}\right|$$
$$\leqslant\frac{1}{n}\Big(\sum_{j,k=1}^{n}\mathbf{E}_{\varepsilon}\Big|W_{jk}^{(\tau)}\Big|^2\Big)^{\frac12}\Big(\sum_{j,k=1}^{n}\mathbf{E}_{\varepsilon}\Big|(\mathbf{R}\mathbf{R}^{(\tau)})_{jk}\Big|^2\Big)^{\frac12}.$$

We rewrite the last inequality as follows

$$\left|\frac{1}{n}\mathbf{E}_\varepsilon \operatorname{Tr}\mathbf{R} - \frac{1}{n}\mathbf{E}_\varepsilon \operatorname{Tr}\mathbf{R}^{(\tau)}\right| \leqslant \sqrt{L_n(\tau)}\sqrt{\mathbf{E}_\varepsilon \operatorname{Tr}\mathbf{R}\overline{\mathbf{R}}\mathbf{R}^{(\tau)}\overline{\mathbf{R}^{(\tau)}}}.$$

From the last bound it follows that

$$\left|\frac{1}{n}\mathbf{E}_\varepsilon \operatorname{Tr}\mathbf{R} - \frac{1}{n}\mathbf{E}_\varepsilon \operatorname{Tr}\mathbf{R}^{(\tau)}\right| \leqslant \sqrt{L_n(\tau)}\left(\sqrt{\mathbf{E}_\varepsilon \operatorname{Tr}|\mathbf{R}|^4} + \sqrt{\mathbf{E}_\varepsilon \operatorname{Tr}|\mathbf{R}^{(\tau)}|^4}\right).$$

Inequality (7.8) and Lemma 7.5 together imply

$$\int_{-\infty}^{\infty}\left|\frac{1}{n}\mathbf{E}_\varepsilon \operatorname{Tr}\mathbf{R} - \frac{1}{n}\mathbf{E}_\varepsilon \operatorname{Tr}\mathbf{R}^{(\tau)}\right| \leqslant \frac{\pi}{v^{\frac{3}{2}}}\sqrt{\frac{\sigma_n^2 + 2v^2}{\sigma_n}}.$$

Lemma 7.5 and the last inequality together imply the result. □

References

1. M. Anttila, K. Ball, I. Perissinaki, *The central limit problem for convex bodies.* — Trans. Amer. Math. Soc. **355** (2003), No. 12, 4723–4735.
2. Z. D. Bai, *Convergence rate of expected spectral distributions of large random matrices. Part* I. *Wigner matrices.* — Ann. Probab. **21** (1993), 625–648.
3. Z. D. Bai, *Methodologies in spectral analysis of large dimensional random matrices: a review.* — Statistica Sinica **9** (1999), 611–661.
4. R. Bellman, *Introduction to matrix analysis.* McGraw-Hill Book Company, New York, 1970.
5. S. G. Bobkov, *On concentration of distributions of random weighted sums.* — Ann. Probab. **31** (2003), No. 1, 195–215.
6. S. G. Bobkov, *Concentration of distributions of the weighted sums with Bernoullian coefficients.* — Geom. Aspects Funct. Anal., Lect. Notes Math. **1807** (2003), 27–36.
7. S. G. Bobkov, F. Götze, *Exponential integrability and transportation cost related to logarithmic Sobolev inequalities.* — J. Funct. Anal. **163** (1999), No. 1, 1–28.
8. B. V. Bronk, *Topics in the theory of random matrices.* Thesis, Princeton University, 1964.
9. A. Guionnet, O. Zeitouni, *Concentration of the spectral measure for large matrices.* — Electron. Comm. Probab. **5** (2000), 119–136.
10. V. L. Girko, *Asymptotics of the distribution of the spectrum of random matrices.* — Russian Math. Surveys **44** (1989), 3–36.
11. F. Götze, A. N. Tikhomirov, *Rate of convergence to the semi-circular law.* — Probab. Theory Relat. Fields **127** (2003), 228–276.
12. I. S. Gradstein, I. M, Ryzhik, *Table of Integrals, Series, and Products.* Academic Press, Inc., New York, 1994.

13. R. Horn, Ch. Johnson, *Matrix analysis.* Cambridge University Press, New York, 1991.
14. A. Khorunzhy, B. A. Khoruzhenko, L. A. Pastur, *On asymptotic properties of large random matrices with independent entries.* — J. Math. Phys. **37** (1996), 5033–5060.
15. M. Ledoux, *Concentration of measure and logarithmic Sobolev inequalities.* — Séminaire de Probabilités, XXXIII, Lect. Notes Math. **1799** (1999), 120–216.
16. M. Ledoux, *The concentration of measure phenomenon.* Math. Surveys and monographs **89**, AMS (2001).
17. M. L. Mehta, *Random Matrices*, 2nd ed., Academic Press, San Diego, 1991.
18. A. Naor, D. Romik, *Projecting the surface measure of the sphere of* ℓ_p^n. — Ann. Inst. H. Poincaré Probab. Statist. **39** (2003), No. 2, 241–261.
19. L. A.Pastur, *Spectra of random self-adjoint operators.* — Russian Math. Surveys **28** (1973), No. 1 (169), 3–64.
20. L. A. Pastur, *Eigenvalue distribution of random matrices.* — Ann. l'Inst. H. Poincaré **64** (1996), 325–337.
21. N. Rosenzweig, *Statistical mechanics of equally likely quantum systems in Statistical physics.* Brandeis Summer Institute, Vol. 3, Benjamin, New York, 1962.
22. R. M. Salem, A Zygmund, . *Trigonometric series whose terms have random sign.* — Acta Math. **91** (1954), 245–301.
23. V. N. Sudakov, *Typical distributions of linear functionals in finite-dimensional spaces of higher dimension.* — Soviet Math. Dokl. **19** (1978), No. 6, 1578–1582.
24. M. Talagrand, *An isoperimetric theorem on the cube and the Kintchine–Kahane inequalities.* — Proc. Amer. Math. Soc. **104** (1988), No. 3, 905–909.
25. H. von Weizsäcker, *Sudakov's typical marginals, random linear functionals, and a conditional central limit theorem.* — Probab. Theory Rel. Fields **107** (1997), 313–324.
26. E. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions.* — Ann. Math. **62** (1955), 548–564.
27. E. Wigner, *On the distribution of the roots of certain symmetric matrices.* — Ann. Math. **67** (1958), 325–327.
28. V. M. Zolotarev, *An estimate for the difference between distributions in Lévy metric.* — Proceedings of Steklov's Math. Institute **112** (1971), 224–231.

School of Mathematics,
University of Minnesota,
Minneapolis, USA

Поступило 17 января 2008 г.

Department of Mathematics,
University of Bielefeld,
Bielefeld, Germany
E-mail: goetze@mathematik.uni-bielefeld.de

S.-Petersburg University,
Department of Mathematics
and Mechanics, Russia
E-mail: tichomir@math.uni-bielefeld.de