# Limit Theorems for Random Matrices

**Alexander Tikhomirov**

*Dedicated to Friedrich Götze on the occasion of his sixtieth birthday*

**Abstract** This note gives a survey of some results on limit theorems for random matrices that have been obtained during the last 10 years in the joint research of the author and F. Götze. We consider the rate of convergence to the semi-circle law and Marchenko–Pastur law, Stein's method for random matrices, the proof of the circular law, and some limit theorems for powers and products of random matrices.

**Keywords** Limit theorems • Number theory • Probability • Statistics • Random Matrices • Semi-Circle Law • Marchenko-Pastur Law • Circular Law

2010 *Mathematics Subject Classification.* 60F05.

## 1 Introduction

In this note we describe some results obtained jointly with F. Götze in the last 10 years. We consider limit theorems for Wigner matrices (Wigner matrix means symmetric real matrix or Hermitian matrix with independent entries up to symmetry), limit theorems for sample covariance matrices and limit theorems for powers and products of independent Ginibre-Girko matrices (that means matrices with all independent entries without any symmetry). We consider results about

A. Tikhomirov (✉)
Department of Mathematics, Komi Research Center Ural Branch of Russian Academy of Sciences, Syktyvkar State University, Chernova 3a, 197001 Syktyvkar, Russia
e-mail: tichomir@math.uni-bielefeld.de

P. Eichelsbacher et al. (eds.), *Limit Theorems in Probability, Statistics and Number Theory*, Springer Proceedings in Mathematics & Statistics 42,
DOI 10.1007/978-3-642-36068-8_14, 

the rate of convergence to the semi-circular law and Marchenko–Pastur law for the empirical spectral distribution function of Wigner matrices and of sample covariance matrices, respectively. For Girko–Ginibre matrices and their powers and products we discuss the results on convergence to the limit distributions. We consider also random matrices with dependent entries and describe Stein's method for random matrices with some martingale structure of dependence of entries.

## 2 Wigner Matrices

Let $X_{jk}$ ($1 \leq j \leq k \leq n$) be independent random variables (possibly complex) with $\mathbb{E}X_{jk} = 0$ and $\mathbb{E}|X_{jk}|^2 = 1$, defined on the same probability space $\{\Omega, \mathfrak{M}, \mathbb{P}\}$. We define the Hermitian (symmetric in real case) matrix $\mathbf{X}$ with entries $[\mathbf{X}]_{jk} = \frac{1}{\sqrt{n}} X_{jk}$ for $1 \leq j \leq k \leq n$. Consider the eigenvalues of the matrix $\mathbf{X}$ denoted in non-increasing order by $\lambda_1 \geq \cdots \geq \lambda_n$ and define the empirical spectral distribution function of this matrix as

$$\mathcal{F}_n(x) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{I}\{\lambda_j \leq x\},$$

where $\mathbb{I}\{A\}$ denotes indicator of the event $A$. Introduce also the expected spectral distribution function $F_n(x) := \mathbb{E}\mathcal{F}_n(x)$ of matrix $\mathbf{X}$. Wigner [39] considered the symmetric random matrix $\mathbf{X}$ with entries $X_{jk} = \pm 1$ with probability $\frac{1}{2}$ and proved that

$$\Delta_n := \sup_x |F_n(x) - G(x)| \to 0, \quad \text{as} \quad n \to \infty, \tag{1}$$

where $G(x)$ is the distribution function of the semi-circular law with the density $G'(x) = \frac{1}{2\pi}\sqrt{4 - x^2}\,\mathbb{I}\{|x| \leq 2\}$. This problem has been studied by several authors. Wigner's result [39] was extended later to different classes of distributions of random variables $X_{jk}$. In particular, Wigner in [40] proved that (1) holds for symmetric random matrices with sub-Gaussian entries. (A random variable $\xi$ is called subgaussian random variable if there exists a positive constant $\beta > 0$ such that $\mathbb{P}\{|\xi| > x\} \leq \exp\{-\beta x^2\}$ for any $x > 0$.) Later it was shown that the semi-circular law (the statement (1)) holds under the assumption of Lindeberg condition for the distributions of matrix entries, i e.,

$$L_n(\tau) = \frac{1}{n^2} \sum_{j=1}^{n} \sum_{k=j}^{n} \mathbb{E}|X_{jk}|^2 \mathbb{I}\{|X_{jk}| \geq \tau\sqrt{n}\} \to 0 \quad \text{as } n \to \infty, \tag{2}$$

for any $\tau > 0$ (see, e.g., [16]). It was shown also that under the same assumptions

$$\Delta_n^* := \sup_x |\mathcal{F}_n(x) - G(x)| \to 0 \quad \text{in probability as} \quad n \to \infty. \tag{3}$$

We have investigated the rate of convergence in (1) and (3). This problem has been studied by several authors. In particular, Bai [5] proved that $\Delta_n = O(n^{-\frac{1}{4}})$ assuming that $\sup_{1\le j\le k} \mathbb{E}|X_{jk}|^4 \le M_4 < \infty$. Later Bai et al. [9] under the condition that $\sup_{1\le j\le k} \mathbb{E}|X_{jk}|^8 \le M_8 < \infty$ proved that $\Delta_n = O(n^{-\frac{1}{2}})$ and $\mathbb{E}\Delta_n^* = O(n^{-\frac{2}{5}})$. Girko [18] proved that $\Delta_n = O(n^{-\frac{1}{2}})$ assuming $\sup_{1\le j\le k} \mathbb{E}|X_{jk}|^4 \le M_4 < \infty$. A very interesting result was obtained recently by Erdös et al. in [13]. It follows from their results that for random matrices whose entries have distributions with exponential tails, i.e., $\mathbb{P}\{|X_{jk}| > t\} \le A\exp\{-t^{\varkappa}\}$ for some $A, \varkappa > 0$, the following holds

$$\mathbb{P}\Big\{\Delta_n^* \le Cn^{-1}(\log n)^{C\ln\ln n}\Big\} \ge 1 - C\exp\{-(\log n)^{c\ln\ln n}\} \tag{4}$$

with some positive constants $C$ and $c$ depending on $A, \varkappa$ only.

We state the results obtained jointly with F. Götze in several theorems below.

**Theorem 2.1 (Götze and Tikhomirov [20]).** *Let* $\mathbb{E}X_{jk} = 0$ *and* $\mathbb{E}|X_{jk}|^2 = 1$. *Let*

$$\sup_{1\le j\le k} \mathbb{E}|X_{jk}|^4 \le M_4 < \infty. \tag{5}$$

*Then there exist a numerical constant* $C > 0$ *such that*

$$\Delta_n \le CM_4^{\frac{1}{2}}n^{-\frac{1}{2}}. \tag{6}$$

*If in addition*

$$\sup_{1\le j\le k} \mathbb{E}|X_{jk}|^{12} \le M_{12} < \infty,$$

*then*

$$\mathbb{E}\Delta_n^* \le CM_{12}^{\frac{1}{6}}n^{-\frac{1}{2}}. \tag{7}$$

Assuming instead of (5) the condition (8) below, we have obtained the following result.

**Theorem 2.2 (Tikhomirov [37]).** *Let* $X_{jk}$ *be independent random variables with* $\mathbb{E}X_{jk} = 0$ *and* $\mathbb{E}|X_{jk}|^2 = 1$. *Assume that for some* $0 < \delta \le 2$ *the following relation holds*

$$\sup_{1\le j\le k} \mathbb{E}|X_{jk}|^{2+\delta} =: M_{2+\delta} < \infty. \tag{8}$$

*Then there exists a numerical* $C > 0$ *such that*

$$\Delta_n \le C\left(\frac{M_{2+\delta}^{\frac{\delta}{2+\delta}}}{n^{\frac{\delta}{2+\delta}}}\right)^{1-\frac{(1-\delta)_+}{2}},$$

*where* $(1-\delta)_+ = \max\{1-\delta, 0\}$.

Under stronger assumptions on the distribution of $X_{jk}$ we get bounds for $\Delta_n$ of order $O(n^{-\frac{1}{2}-\gamma})$ with some positive $\gamma > 0$. In particular, in the paper of Bobkov et al. [13] we consider random variables $X_{jk}$ with distributions satisfying a Poincaré-type inequality. Let us recall that a probability measure $\mu$ on $\mathbb{R}^d$ is said to satisfy a Poincaré-type, $PI(\sigma^2)$, or a spectral gap inequality with constant $\sigma^2$ if for any bounded smooth function $g$ on $\mathbb{R}^d$ with gradient $\nabla g$

$$\mathrm{Var}(g) \le \sigma^2 \int_{\mathbb{R}^d} |\nabla g|^2 d\mu, \tag{9}$$

where $\mathrm{Var}(g) = \int_{\mathbb{R}^d} g^2 d\mu - \left(\int_{\mathbb{R}^d} g d\mu\right)^2$.

**Theorem 2.3 (Bobkov et al. [13]).** *If the distributions of $X_{jk}$'s satisfy the Poincaré-type inequality $PI(\sigma^2)$ on the real line, then*

$$\Delta_n \le Cn^{-2/3},$$

*where the constant $C$ depends on $\sigma$ only. Moreover,*

$$\mathbb{E}\Delta_n^* \le Cn^{-2/3}\log^2(n+1).$$

For any positive constants $\alpha > 0$ and $\varkappa > 0$ define the quantities

$$l_{n,\alpha} := \log n\,(\log\log n)^{\alpha} \quad \text{and} \quad \beta_n := (l_{n,\alpha})^{\frac{1}{\varkappa}+\frac{1}{2}}. \tag{10}$$

The best known result for the rate of convergence in probability to the semi-circular law is the following:

**Theorem 2.4 (Götze and Tikhomirov [28]).** *Let $\mathbb{E}X_{jk} = 0$, $\mathbb{E}X_{jk}^2 = 1$. Assume that there exist constants $A$ and $\varkappa > 0$ such that*

$$\mathbb{P}\{|X_{jk}| \ge t\} \le A\exp\{-t^{\varkappa}\}, \tag{11}$$

*for any $1 \le j \le k \le n$ and any $t \ge 1$. Then, for any positive $\alpha > 0$ there exist positive constants $C$ and $c$ depending on $A$ and $\varkappa$ and $\alpha$ only, such that*

$$\mathbb{P}\Big\{\sup_x |\mathcal{F}_n(x) - G(x)| > n^{-1}\beta_n^2\Big\} \le C\,\exp\{-c\,l_{n,\alpha}\}.$$

*Remark 2.5.* In the result of (4) [13] $\Delta_n^* = O_P(n^{-1}(\log n)^{O(\log\log n)})$. In our result $\Delta_n^* = O_P(n^{-1}(\log n)^{O(1)})$.

*Remark 2.6.* If $\mathbf{X}$ belongs to Gaussian Unitary Ensemble (GUE) [23] or Gaussian Orthogonal Ensemble (GOE) [38] then there exists an absolute constant $C > 0$ such that

$$\Delta_n \le Cn^{-1}. \tag{12}$$

## 3 Sample Covariance Matrices

In this section we consider the so called sample covariance matrices and their generalization. Let $\mathbf{X}$ be rectangular matrices of order $[n \times p]$ with independent entries (possible complex) $X_{jk}$, $j = 1, \ldots, n$; $k = 1 \ldots, p$. We shall assume that $\mathbb{E}X_{jk} = 0$ and $\mathbb{E}|X_{jk}|^2 = 1$. Consider the matrix $\mathbf{W} = \frac{1}{p}\mathbf{X}\mathbf{X}^*$. Such matrices are called sample covariance matrices and they were first considered in 1928 by Wishart [41]. He obtained the joint distribution of entries of the matrix $\mathbf{W}$ as $X_{jk}$ are standard Gaussian random variables. We shall be interested in the asymptotic distribution of the spectrum of the matrix $\mathbf{W}$. Note that the matrix $\mathbf{W}$ is semi-positive definite and its eigenvalues are non-negative. Denote the eigenvalues of the matrix $\mathbf{W}$ in decreasing order by $s_1^2 \geq \ldots \geq s_n^2 \geq 0$. (Note that the numbers $s_1, \ldots, s_n$ are called singular values of matrix $\frac{1}{\sqrt{p}}\mathbf{X}$.) Define the empirical spectral distribution function of the matrix $\mathbf{W}$ by the equality

$$\mathcal{H}_n(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{I}\{s_j^2 \leq x\}. \tag{13}$$

Let $H_y(x)$ be the distribution function with the density

$$H_y'(x) = \frac{\sqrt{(b-x)(x-a)}}{2\pi x y} + (1 - \frac{1}{y})_+ \delta_0(x), \tag{14}$$

where $a = (1-\sqrt{y})^2$, $b = (1+\sqrt{y})^2$, and $\delta_0(x)$ denotes Dirac $\delta$-function, $a_+ = \max\{a, 0\}$ for any real $a$. This distribution is called Marchenko–Pastur distribution with parameter $y$. Assuming that $p = p(n)$ where $\lim_{n\to\infty} \frac{n}{p} = y$, and assuming the moment condition (5), Marchenko and Pastur [29] have shown that there exists

$$\lim_{n\to\infty} \mathbb{E}\mathcal{H}_n(x) = H_y(x), \tag{15}$$

The result of Marchenko–Pastur was improved by many authors. As a final result we have the following Theorem.

**Theorem 3.1.** *Let the random variables $X_{jk}$, $1 \leq j \leq n$, $1 \leq k \leq p$ be independent for any $n \geq 1$ and have zero mean and unit variance. Assume that $p = p(n)$ such that $\lim_{n\to\infty} \frac{n}{p} = y$. Further suppose that the Lindeberg condition holds, i.e.,*

$$L_n(\tau) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{p} \mathbb{E}|X_{jk}|^2 \mathbb{I}\{|X_{jk}| \geq \tau\sqrt{n}\} \to 0,$$

*for any $\tau > 0$. Then*

$$\sup_x |\mathbb{E}\mathcal{H}_n(x) - H_y(x)| \to 0, \text{ as } n \to \infty.$$

*Moreover, $\mathcal{H}_n(x)$ converges to the Marchenko–Pastur distribution in probability.*

The proof of this result may be found in [8]. We have investigated the rate of convergence of the expected and empirical spectral distribution function of sample covariance matrix to the Marchenko–Pastur law. This question was considered also in the papers of Bai [6], and in Bai and co-authors [10]. Bai et al. in [10] and independently Götze and Tikhomirov in [21] established the bound of the rate of convergence in Kolmogorov distance $\Delta_n = \sup_x |\mathbb{E}\mathcal{H}_n(x) - H_y(x)| = O(n^{-\frac{1}{2}})$, assuming that

$$\max_{j,k\geq 1} \mathbb{E}|X_{jk}|^8 \leq C \tag{16}$$

for some positive constant $C > 0$ independent of $n$. Götze and and Tikhomirov [21] proved as well that $\Delta_n^* = \mathbb{E}\sup_x |\mathcal{H}_n(x) - H_y(x)| = O(n^{-\frac{1}{2}})$, assuming

$$\max_{j,k\geq 1} \mathbb{E}|X_{jk}|^{12} \leq C \tag{17}$$

for some positive constant $C > 0$ independent of $n$. Somewhat later these bounds were improved in the paper of Götze and Tikhomirov [26] and in the paper of Tikhomirov [36]. We formulate the following result.

**Theorem 3.2.** *Let the random variables $X_{jk}$, $1 \leq j \leq n$, $1 \leq k \leq p$ be independent for any fixed $n \geq 1$ and have zero mean and unit variance. Assume that $p = p(n)$, where $\frac{n}{p} = y \leq 1$. Let for some $0 < \delta \leq 2$*

$$M_{2+\delta} := \sup_{j,k,n} \mathbb{E}|X_{jk}|^{2+\delta} < \infty.$$

*Then there exist a positive constant $C = C(\delta)$, depending on $\delta$ only, such that*

$$\Delta_n \leq C M_{2+\delta}^{\frac{\delta}{2+\delta}} n^{-\frac{\delta}{2+\delta}}. \tag{18}$$

The bound (18) for $\delta = 2$ was obtained in [26], the bound for the case $0 < \delta < 2$ in [36]. The question about optimality of the above mentioned bounds is still open. But assuming that the random variables $X_{jk}$ are independent standard complex Gaussian random variables (so-called Laguerre unitary ensemble) the optimal bound of the rate of convergence of the expected spectral distribution of the matrix $\mathbf{W}$ was obtained. It turns out that $\Delta_n = O(n^{-1})$, which was proved by Götze and Tikhomirov [23]. Recall that the distribution of a random variable $X$ has so-called exponential tail means that there exist constants $A > 0$ and $\varkappa > 0$ such that

$$\mathbb{P}\{|X| \geq t\} \leq A\exp\{-t^{\varkappa}\}. \tag{19}$$

Assuming that the entries of the matrix $\mathbf{X}$ have distribution with exponential tails, Götze and Tikhomirov have proved in [27] that

$$\mathbb{P}\Big\{\sup_x |\mathcal{H}_n(x) - H_y(x)| > n^{-1}\beta_n^2\Big\} \le C\ \exp\{-c\, l_{n,\alpha}\}, \tag{20}$$

for any $\alpha > 0$. Here $\beta_n$ and $l_{n,\alpha}$ were defined in (10). The constants $C > 0$ and $c > 0$ depend on $A, \varkappa$ and $\alpha$ only. It would be interesting to extend the results about sample covariance matrices to more general situations. First we consider the singular values of powers of random matrices. And then we consider the asymptotic distribution of singular values of products of independent random matrices.

## 3.1 Powers of Random Matrices

Let $\mathbf{X} = (X_{jk})_{j,k=1}^n$ be a square random matrix of order $n$ with independent entries such that $\mathbb{E}X_{jk} = 0$ and $\mathbb{E}|X_{jk}|^2 = 1$. In this section we shall investigate the asymptotic distribution of the singular values of the matrix $\mathbf{W} = n^{-\frac{m}{2}}\mathbf{X}^m$ or the eigenvalues of the matrix $\mathbf{V} = \mathbf{W}\mathbf{W}^*$. For $m = 1$ it is the case of sample covariance matrix with parameter $y = 1$. Denote by $s_1^2 \ge \ldots \ge s_n^2$ the eigenvalues of the matrix $\mathbf{V}$. (Note that $s_1 \ge \ldots \ge s_n$ are the singular values of the matrix $\mathbf{W}$.) Let

$$\mathcal{H}_n^{(m)}(x) = \frac{1}{n}\sum_{j=1}^{m} \mathbb{I}\{s_j^2 \le x\} \tag{21}$$

denote the empirical spectral distribution function of the matrix $\mathbf{V}$. Let $FC(k,m) = \frac{1}{mk+1}\binom{mk+k}{k}$ denote the $k$th Fuss–Catalan number with parameter $m$, for $k \ge 1$. These numbers are the moments of some distribution which we denote by $H^{(m)}(x)$. It is well known that the Stieltjes transform of this distribution, $s^{(m)}(z) = \int \frac{1}{x-z} dH^{(m)}(x)$, satisfies the equation

$$1 + zs^{(m)}(z) + (-1)^{m+1}z^{m+1}(s^{(m)}(z))^{m+1} = 0,$$

In the joint papers of Alexeev et al. [2] and [1] the following was proved:

**Theorem 3.3.** *Let random variables $X_{jk}$ be independent for any fixed $n \ge 1$ and for any $1 \le j,k \le n$. Assume that $\mathbb{E}X_{jk} = 0$ and $\mathbb{E}|X_{jk}|^2 = 1$ for any $j,k \ge 1$ and*

$$\sup_{j,k\ge 1} \mathbb{E}|X_{jk}|^4 \le C, \tag{22}$$

*for some positive constant $C > 0$. Assume also that for any $\tau > 0$*

$$L_n(\tau) = \frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}|X_{jk}|^4 \mathbb{I}\{|X_{jk}| \ge \tau\sqrt{n}\} \to 0 \quad \text{as } n \to \infty. \tag{23}$$

*Then*

$$\lim_{n\to\infty} \sup_x |\mathbb{E}\mathcal{H}_n^{(m)}(x) - H^{(m)}(x)| = 0. \tag{24}$$

For the proof of this result we use the method of moments, see [2]. The proof of Theorem 3.2 by the method of Stieltjes transform is given in [4].

## 3.2 Product of Random Matrices

Let $m \geq 1$ be a fixed integer. For any $n \geq 1$ consider an $(m+1)$-tuple of integers $(p_0, \ldots, p_m)$ with $p_0 = n$ and $p_\nu = p_\nu(n)$ for $\nu = 1, \ldots, m$, such that

$$\lim_{n\to\infty} \frac{n}{p_\nu(n)} = y_\nu \in (0, 1]. \tag{25}$$

Furthermore, we consider an array of independent complex random variables $X_{jk}^{(\nu)}$, $1 \leq j \leq p_{\nu-1}, 1 \leq k \leq p_\nu$, $\nu = 1, \ldots, m$ defined on a common probability space $\{\Omega_n, \mathbb{F}_n, \mathbb{P}\}$ with $\mathbb{E}X_{jk}^{(\nu)} = 0$ and $\mathbb{E}|X_{jk}^{(\nu)}|^2 = 1$. Let $\mathbf{X}^{(\nu)}$ denote the $p_{\nu-1} \times p_\nu$ matrix with entries $[\mathbf{X}^{(\nu)}]_{jk} = \frac{1}{\sqrt{p_\nu}} X_{jk}^{(\nu)}$, for $1 \leq j \leq p_{\nu-1}, 1 \leq k \leq p_\nu$. The random variables $X_{jk}^{(\nu)}$ may depend on $n$ but for simplicity we shall not make this explicit in our notations. Denote by $s_1 \geq \ldots \geq s_n$ the singular values of the random matrix $\mathbf{W} := \prod_{\nu=1}^m \mathbf{X}^{(\nu)}$ and define the empirical distribution of its squared singular values by

$$\mathcal{H}_n^{(m)}(x) = \frac{1}{n}\sum_{k=1}^n \mathbb{I}\{s_k{}^2 \leq x\}.$$

We shall investigate the approximation of the expected spectral distribution $H_n^{(m)}(x) = \mathbb{E}\mathcal{H}_n^{(m)}(x)$ by the distribution function $H_{\mathbf{y}}(x)$ which is defined by its Stieltjes transform $s_{\mathbf{y}}(z)$ in the following way:

$$1 + zs_{\mathbf{y}}(z) - s_{\mathbf{y}}(z)\prod_{l=1}^m (1 - y_l - zy_l s_{\mathbf{y}}(z)) = 0, \tag{26}$$

where $0 \leq y_l \leq 1$.

*Remark 3.4.* In the case $y_1 = y_2 = \cdots = y_m = 1$ the distribution $H_{\mathbf{y}}$ has moments $M(k,m) = FC(k,m)$. The Stieltjes transform of the distribution $H_{\mathbf{y}}(x)$ satisfies in this case the equation

$$1 + zs(z) + (-1)^{m+1} z^m s(z)^{m+1} = 0.$$

The main result of this subsection.

**Theorem 3.5.** *Assume that condition (25) holds. Let* $\mathbb{E}X_{jk}^{(\nu)} = 0$, $\mathbb{E}|X_{jk}^{(\nu)}|^2 = 1$. *Suppose that the Lindeberg condition holds, i.e.,*

$$L_n(\tau) := \max_{\nu=1,\dots,m} \frac{1}{p_{\nu-1}p_\nu} \sum_{j=1}^{p_{\nu-1}} \sum_{k=1}^{p_\nu} \mathbb{E}|X_{jk}^{(\nu)}|^2 I_{\{|X_{jk}^{(\nu)}| \ge \tau\sqrt{n}\}} \to 0 \quad \text{as } n \to \infty, \tag{27}$$

*for any* $\tau > 0$. *Then*

$$\lim_{n\to\infty} \sup_x |H_n^{(m)}(x) - H_{\mathbf{y}}(x)| = 0.$$

*Remark 1.* For $m = 1$ we get the well-known result of Marchenko-Pastur for sample covariance matrices [29].

*Remark 2.* We see that the limit distribution for the distribution of singular values of product of independent square random matrices is the same as for powers of random matrices with independent entries, see [2].

The statement of Theorem 3.5 was published in [1] and a proof of this result is given in [3].

## 4 Circular Law and Its Generalization

### *4.1 Circular Law*

Let $X_{jk}, 1 \le j,k < \infty$, be complex random variables with $\mathbb{E}X_{jk} = 0$ and $\mathbb{E}|X_{jk}|^2 = 1$. For a fixed $n \ge 1$, denote by $\lambda_1, \dots, \lambda_n$ the eigenvalues of the $n \times n$ matrix

$$\mathbf{X} = (X_n(j,k))_{j,k=1}^n, \quad X_n(j,k) = \frac{1}{\sqrt{n}} X_{jk} \text{ for } 1 \le j,k \le n, \tag{28}$$

and define its empirical spectral distribution function by

$$G_n(x,y) = \frac{1}{n} \sum_{j=1}^n \mathbb{I}\{\text{Re}\{\lambda_j\} \le x,\ \text{Im}\{\lambda_j\} \le y\}. \tag{29}$$

We investigate the convergence of the expected spectral distribution function $\mathbb{E}G_n(x,y)$ to the distribution function $G(x,y)$ of the uniform distribution in the unit disc in $\mathbb{R}^2$.

The main results which was obtained in [19] is the following.

**Theorem 4.1.** *Let* $X_{jk}$ *be independent random variables with*

$$\mathbb{E}X_{jk} = 0, \qquad \mathbb{E}|X_{jk}|^2 = 1 \quad \text{and} \quad \mathbb{E}|X_{jk}|^2\varphi(X_{jk}) \le \varkappa,$$

*where* $\varphi(x) = (\ln(1+|x|))^{19+\eta}$ *for some* $\eta > 0$. *Then* $\mathbb{E}G_n(x,y)$ *converges weakly to the distribution function* $G(x,y)$ *as* $n \to \infty$.

We shall prove the same result for the following class of sparse matrices. Let $\varepsilon_{jk}$, $j,k = 1,\dots,n$ denote Bernoulli random variables which are independent in aggregate and independent of $(X_{jk})_{j,k=1}^n$ with success probability $p_n := \mathbb{P}\{\varepsilon_{jk} = 1\}$. Consider the matrix $\mathbf{X}^{(\varepsilon)} = \frac{1}{\sqrt{np_n}}(\varepsilon_{jk}X_{jk})_{j,k=1}^n$. Let $\lambda_1^{(\varepsilon)},\dots,\lambda_n^{(\varepsilon)}$ denote the (complex) eigenvalues of the matrix $\mathbf{X}^{(\varepsilon)}$ and denote by $G_n^{(\varepsilon)}(x,y)$ the empirical spectral distribution function of the matrix $\mathbf{X}^{(\varepsilon)}$, i. e.

$$G_n^{(\varepsilon)}(x,y) := \frac{1}{n}\sum_{j=1}^n \mathbb{I}\{\mathrm{Re}\{\lambda_j^{(\varepsilon)}\} \le x,\ \mathrm{Im}\{\lambda_j^{(\varepsilon)}\} \le y\}. \tag{30}$$

**Theorem 4.2.** *Let* $X_{jk}$ *be independent random variables with*

$$\mathbb{E}X_{jk} = 0, \qquad \mathbb{E}|X_{jk}|^2 = 1 \quad \textit{and} \quad \mathbb{E}|X_{jk}|^2\varphi(X_{jk}) \le \varkappa,$$

*where* $\varphi(x) = (\ln(1+|x|))^{19+\eta}$ *for some* $\eta > 0$. *Assume that* $p_n^{-1} = \mathcal{O}(n^{1-\theta})$ *for some* $1 \ge \theta > 0$. *Then* $\mathbb{E}G_n^{(\varepsilon)}(x,y)$ *converges weakly to the distribution function* $G(x,y)$ *as* $n \to \infty$.

*Remark 4.3.* The crucial problem of the proofs of Theorems 4.1 and 4.2 is to find bounds for the smallest singular values $s_n(z)$ respectively $s_n^{(\varepsilon)}(z)$ of the shifted matrices $\mathbf{X}-z\mathbf{I}$ respectively $\mathbf{X}^{(\varepsilon)}-z\mathbf{I}$. These bounds are based on the results obtained by Rudelson and Vershynin in [32]. In the version of paper [25] we have used the corresponding results of Rudelson [31] proving the circular law in the case of i.i.d. sub-Gaussian random variables. In fact, the results in [25] actually imply the circular law for i.i.d. random variables with $\mathbb{E}|X_{jk}|^4 \le \varkappa_4 < \infty$ in view of the fact (explicitly stated by Rudelson in [31]) that in his results the sub-Gaussian condition is needed for the proof of $\mathbb{P}\{\|\mathbf{X}\| > K\} \le C\exp\{-cn\}$ only. This result was written by Pan and Zhou in [30].

The strong circular law assuming moment condition of order larger than 2 and comparable sparsity assumptions was proved by Tao and Vu in [33] based on their results in [34] in connection with the multivariate Littlewood Offord problem. In [35] Tao and Vu proved the circular law without sparsity assuming a moment condition of order 2 only.

The investigation of the convergence of the spectral distribution functions of real or complex (non-symmetric and non-Hermitian) random matrices with independent entries has a long history. Ginibre in 1965, [14], studied the real, complex and quaternion matrices with i. i. d. Gaussian entries. He derived the joint density for the distribution of eigenvalues of matrix and determined the density of the expected spectral distribution function of random matrix with Gaussian entries with independent real and imaginary parts and deduced the circle law. Using the Ginibre results, Edelman in 1997, [12], proved the circular law for matrices with i. i. d. Gaussian real entries. Girko in 1984, [15], investigated the circular law for general matrices with independent entries assuming that the distributions of the entries have densities. As pointed out by Bai [7], Girko's proof had serious gaps. Bai in [7] gave a proof of the circular law for random matrices with independent entries assuming that the entries have bounded densities and finite sixth moments. It would be interesting to consider the following generalization of the circular law.

## *4.2 Asymptotic Spectrum of the Product of Independent Random Matrices*

Let $m \geq 1$ be a fixed integer. For any $n \geq 1$ consider mutually independent identically distributed (i.i.d.) omplex random variables $X_{jk}^{(\nu)}$, $1 \leq j,k \leq n$, $\nu = 1,\dots,m$, with $\mathbb{E}X_{jk}^{(\nu)} = 0$ and $\mathbb{E}|X_{jk}^{(\nu)}|^2 = 1$ defined on a common probability space $(\Omega_n, \mathbb{F}_n, \mathbb{P})$. Let $\mathbf{X}^{(\nu)}$ denote the $n \times n$ matrix with entries $[\mathbf{X}^{(\nu)}]_{jk} = \frac{1}{\sqrt{n}} X_{jk}^{(\nu)}$, for $1 \leq j,k \leq n$. Denote by $\lambda_1,\dots,\lambda_n$ the eigenvalues of the random matrix $\mathbf{W} := \prod_{\nu=1}^m \mathbf{X}^{(\nu)}$ and define its empirical spectral distribution function by

$$\mathcal{F}_n(x,y) := \mathcal{F}_n^{(m)}(x,y) = \frac{1}{n}\sum_{k=1}^n \mathbb{I}\{\operatorname{Re}\lambda_k \leq x, \operatorname{Im}\lambda_k \leq y\},$$

where $\mathbb{I}\{B\}$ denotes the indicator of an event $B$. We shall investigate the convergence of the expected spectral distribution $F_n(x,y) = \mathbb{E}\mathcal{F}_n(x,y)$ to the distribution function $F(x,y)$ corresponding to the $m$-th power of the uniform distribution on the unit disc in the plane $\mathbb{R}^2$ with Lebesgue-density

$$f(x,y) = \frac{1}{\pi m (x^2+y^2)^{\frac{m-1}{m}}} \mathbb{I}\{x^2+y^2 \leq 1\}.$$

We consider the Kolmogorov distance between the distributions $F_n(x,y)$ and $F(x,y)$,

$$\Delta_n := \sup_{x,y} |F_n(x,y) - F(x,y)|.$$

We have proved the following.

**Theorem 4.4.** *Let* $\mathbb{E}X_{jk}^{(\nu)} = 0$, $\mathbb{E}|X_{jk}^{(\nu)}|^2 = 1$. *Then, for any fixed* $m \geq 1$,

$$\lim_{n\to\infty} \Delta_n = 0.$$

The result holds in the non-i.i.d. case as well.

**Theorem 4.5.** *Let* $\mathbb{E}X_{jk}^{(\nu)} = 0$, $\mathbb{E}|X_{jk}^{(\nu)}|^2 = 1$, $\nu = 1, \ldots, m$, $j, k = 1, \ldots, n$. *Assume that the random variables* $X_{jk}^{(\nu)}$ *have uniformly integrable second moments, i. e.*

$$\max_{\nu,j,k,n} \mathbb{E}|X_{jk}^{(\nu)}|^2 \mathbb{I}\{|X_{jk}^{(\nu)}| > M\} \to 0 \quad as \quad M \to \infty. \tag{31}$$

*Then for any fixed* $m \geq 1$,

$$\lim_{n\to\infty} \Delta_n = 0.$$

**Definition 4.6.** Let $\mu_n(\cdot)$ denote the empirical spectral measure of an $n \times n$ random matrix $\mathbf{X}$ (uniform distribution on the eigenvalues of matrix $\mathbf{X}$) and let $\mu(\cdot)$ denote the uniform distribution on the unit disc in the complex plane $\mathbb{C}$. We say that the circular law holds for the random matrices $\mathbf{X}$ if $\mathbb{E}\mu_n(\cdot)$ converges weakly to the measure $\mu(\cdot)$ in the complex plane $\mathbb{C}$.

*Remark 4.7.* For $m = 1$ we recover the well-known circular law for random matrices [19, 35].

## 5 Bounds on Levy and Kolmogorov Distance in Terms of Stieltjes Transform

One of the first bounds on the Kolmogorov distance between distribution functions via their Stieltjes transforms was obtained by Girko in [17]. Bai in [5] proved a new inequality bounding the Kolmogorov distance of distribution functions by their Stieltjes transforms. The proofs of Theorems 2.1–3.3 are based on a smoothing inequality for the Kolmogorov distance between distribution functions in terms of their Stielties transform. Recall that the Stieltjes transform $S_F(z)$ of a distribution function $F(x)$ is defined by the equality

$$S_F(z) := \int_{\mathbb{R}} \frac{1}{z - x} dF(x),$$

for all $z = u + iv$ with $u \in \mathbb{R}$ and $v > 0$. For any distribution functions $F$ and $G$ define the Levy distance as

$$L(F, G) := \inf\{\delta > 0 : \ F(x - \delta) - \delta \leq G(x) \leq F(x + \delta) + \delta, \quad \text{for all } x \in \mathbb{R}\}. \tag{32}$$

In [13] the following result was proved.

**Theorem 5.1.** *Let* $F$ *and* $G$ *be distribution functions. Given* $v > 0$, *let an interval* $[\alpha, \beta] \subset \mathbb{R}$ *be chosen to satisfy* $G(\alpha) < v$ *and* $1 - G(\beta) < v$. *Then*

$$L(F,G) \le \sup_{x\in[\alpha-2v,\beta+2v]} \Big| \int_{-\infty}^{x} (S_F(x+iv) - S_G(x+iv))dx \Big| + 4v + 50\,\mathrm{Im}\, S_G(x+iv). \tag{33}$$

The following corollaries are very important for applications.

**Corollary 5.2 (Bobkov et al. [13]).** *Let $F$ and $G$ be arbitrary distribution functions. With some universal constant $c > 0$, for any $v_1 > v_0 > 0$,*

$$cL(F,G) \le v_0 + v_0 \sup_{x\in\mathbb{R}} \mathrm{Im}\, S_G(x+iv_0) + \int_{\mathbb{R}} |S_G(u+iv_1) - S_F(u+iv_1)|du + \sup_{x\in[\alpha-2v,\beta+2v]} \int_{v_0}^{v_1} |S_G(x+iv) - S_F(x+iv)|dv, \tag{34}$$

*where $\alpha < \beta$ are chosen to satisfy $G(\alpha) < v_0$, and $1 - G(\beta) < v_0$.*

**Corollary 5.3.** *If $G$ is the distribution function of the standard semi-circular law, and $F$ is any distribution function, we have for all $v_1 > v_0 > 0$, up to some universal constant $c > 0$,*

$$c\|F-G\| := c\sup_{x\in\mathbb{R}} |F(x) - G(x)| \le v_0 + \int_{\mathbb{R}} |S_G(u+iv_1) - S_F(u+iv_1)|du + \sup_{x\in[\alpha-2v,\beta+2v]} \int_{v_0}^{v_1} |S_G(x+iv) - S_F(x+iv)|dv \tag{35}$$

This result improved a similar inequality (2.4) in [20]. The main idea of such type of inequalities belongs to F. Götze. We consider the first integral in the right hand side of (35) ("horizontal") far from the real line. A distance from a real line in the second integral ("vertical") has an order $O(n^{-1}\log n^b)$ in one point only. To obtain a bound of order $O(n^{-1}\log n^b)$ for $\Delta_n$ we need some modification of the last result. Let $\gamma = \sqrt{4-x^2}$.

**Theorem 5.4.** *Let $v > 0$ and $a$ and $\varepsilon > 0$ be positive numbers such that*

$$\alpha = \frac{1}{\pi}\int_{|u|\le a} \frac{1}{u^2+1}du = \frac{3}{4}, \tag{36}$$

*and*

$$2va \le \varepsilon\sqrt{\gamma}. \tag{37}$$

*If $G$ denotes the distribution function of the standard semi-circular law, and $F$ is any distribution function, there exists some absolute constant $c > 0$, such that*

$$c\|F-G\| \le \int_{-\infty}^{\infty} |S_F(u+iV) - S_G(u+iV)|du + v + \varepsilon^{\frac{3}{2}}$$
$$+ \sup_{x\in[\alpha-2v,\beta+2v]} \int_{v'}^{V} |S_F(x+iu) - S_G(x+iu)|du, \qquad (38)$$

*where* $\alpha$, $\beta$ *are defined in Theorem 5.2 and* $v' = v/\sqrt{\gamma}$.

In the inequality (38) the right hand side is "sensitive" to the closeness of the point $x$ to the end points of the support of semi-circular distribution function.

## 6 Stein's Method for Random Matrices

One of the more interesting direction of our joint work with F. Götze was a development of Stein's method for random matrices. This idea belongs exclusively to F. Götze. The obtained results were published in several papers [22, 24, 25], and we give a short review of them in this section. The goal of this review is to illustrate the possibilities of Stein's method for the investigation of the convergence of the empirical spectral distribution function of random matrices. We consider two ensembles of random matrices: real symmetric matrices and sample covariance matrices of real observations. We give a simple characterization of both semicircle and Marchenko-Pastur distributions via linear differential equations. Using conjugate differential operators, we give a simple criterion for convergence to these distributions. We state also the general sufficient conditions for the convergence of the expected spectral distribution functions of random matrices.

### *6.1 Real Symmetric Matrices*

Let $X_{jk}$, $1 \le j \le k < \infty$, be a triangular array of random variables with $\mathbb{E}X_{jk} = 0$ and $\mathbb{E}X_{jk}^2 = \sigma_{jk}^2$, and let $X_{kj} = X_{jk}$, for $1 \le j < k < \infty$. For a fixed $n \ge 1$, denote by $\lambda_1 \le \ldots \le \lambda_n$ the eigenvalues of a symmetric $n \times n$ matrix

$$\mathbf{W}_n = (W_n(jk))_{j,k=1}^n, \quad W_n(jk) = \frac{1}{\sqrt{n}} X_{jk}, \text{ for } 1 \le j \le k \le n, \qquad (39)$$

and define its empirical spectral distribution function by

$$\mathcal{F}_n(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{I}\{\lambda_j \le x\}. \qquad (40)$$

We investigate the convergence of the expected spectral distribution function, $F_n(x) := \mathbb{E}\mathcal{F}_n(x)$, to the distribution function of Wigner's semicircle law.

Let $g(x)$ and $G(x)$ denote the density and the distribution function of the standard semicircle law, i.e.

$$g(x) = \frac{1}{2\pi}\sqrt{4-x^2}\,\mathbb{I}\{|x| \le 2\}. \quad G(x) = \int_{-\infty}^{x} g(u)du. \tag{41}$$

## 6.2 Stein's Equation for the Semicircle Law

Introduce a class of functions

$$\mathbb{C}^1_{\{-2,2\}} = \{f : \mathbb{R} \to \mathbb{R} : f \in \mathbb{C}^1(\mathbb{R} \setminus \{-2,2\}); \\ \overline{\lim}_{|y|\to\infty}|yf(y)| < \infty;\ \limsup_{y\to\pm 2}|4-y^2||f'(y)| < C\}.$$

By $\mathbb{C}(\mathbb{R})$ we denote the class of continuous functions on $\mathbb{R}$, by $\mathbb{C}^1(B)$, $B \subset \mathbb{R}$, we denote the class of all functions $f : \mathbb{R} \to \mathbb{R}$ differentiable on B with bounded derivative on all compact subsets of $B$. We state the following

**Lemma 6.1.** *Assume that a bounded function $\varphi(x)$ without discontinuity of second order satisfies the following conditions*

$$\varphi(x) \textit{ is continuous at the points } x = \pm 2 \tag{42}$$

*and*

$$\int_{-2}^{2} \varphi(u)\sqrt{4-u^2}du = 0. \tag{43}$$

*Then there exists a function $f \in \mathbb{C}^1_{\{-2,2\}}$ such that, for any $x \neq \pm 2$,*

$$(4-x^2)f'(x) - 3xf(x) = \varphi(x). \tag{44}$$

*If $\varphi(\pm 2) = 0$ then there exists a continuous solution of (44).*

As a simple implication of this Lemma we get

**Proposition 6.2.** *The random variable $\xi$ has distribution function $G(x)$ if and only if the following equality holds, for any function $f \in \mathbb{C}^1_{\{-2,2\}}$,*

$$\mathbb{E}\left((4-\xi^2)f'(\xi) - 3\xi f(\xi)\right) = 0. \tag{45}$$

## 6.3 Stein Criterion for Random Matrices

Let $\mathbf{W}$ denote a symmetric random matrix with eigenvalues $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$. If $\mathbf{W} = \mathbf{U}^{-1}\boldsymbol{\Lambda}\mathbf{U}$, where $\mathbf{U}$ is an orthogonal matrix and $\boldsymbol{\Lambda}$ is a diagonal matrix, one defines $f(\mathbf{W}) = \mathbf{U}^{-1} f(\boldsymbol{\Lambda})\mathbf{U}$, where $f(\boldsymbol{\Lambda}) = \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n))$.

We can now formulate the convergence to the semicircle law for the spectral distribution function of random matrices.

**Theorem 6.3.** *Let $\mathbf{W}_n$ denote a sequence of random matrices of order $n \times n$ such that, for any function $f \in \mathbb{C}^1_{\{-2,2\}}$*

$$\frac{1}{n}\mathbb{E}\mathrm{Tr}(4\mathbf{I}_n - \mathbf{W}_n^2) f'(\mathbf{W}_n) - \frac{3}{n}\mathbb{E}\mathrm{Tr}\mathbf{W}_n f(\mathbf{W}_n) \to 0, \quad \text{as } n \to \infty. \tag{46}$$

*Then*

$$\Delta_n := \sup_x |\mathbb{E}F_n(x) - G(x)| \to 0, \ \text{as } n \to \infty. \tag{47}$$

## 6.4 Resolvent Criterion for the Spectral Distribution Function of a Random Matrix

We introduce the resolvent matrix for a symmetric matrix $\mathbf{W}$ and any non-real $z$,

$$\mathbf{R}(z) = (\mathbf{W} - z\mathbf{I})^{-1}, \tag{48}$$

where $\mathbf{I}$ denotes the identity matrix of order $n \times n$.

**Proposition 6.4.** *Assume that, for any $v \ne 0$,*

$$\mathcal{R}_n(\mathbf{W})(z) := \frac{1}{n}\mathbb{E}\mathrm{Tr}(4\mathbf{I} - \mathbf{W}^2)\mathbf{R}^2(z) + \frac{3}{n}\mathbb{E}\mathrm{Tr}\mathbf{W}\mathbf{R}(z) \to 0, \ \text{as } n \to \infty \tag{49}$$

*uniformly on compact sets in $\mathbb{C}\backslash\mathbb{R}$. Then*

$$\Delta_n \to 0, \ \text{as } n \to \infty. \tag{50}$$

## 6.5 General Conditions of the Convergence of the Expected Distribution Function of Random Matrices to the Semi Circular Law

We shall assume that $\mathbb{E}X_{jl} = 0$ and $\sigma_{jl}^2 := \mathbb{E}X_{jl}^2$, for $1 \le j \le l \le n$. Introduce $\sigma$-algebras $\mathcal{F}^{jl} = \sigma\{X_{km} : 1 \le k \le m \le n, \ \{k, m\} \ne \{j, l\}\}$, $1 \le j \le l \le n$, and

$\mathcal{F}^j = \sigma\{X_{km} : 1 \le k \le m \le n,\ k \ne j$ and $m \ne j\}$, $1 \le j \le n$. We introduce as well Lindeberg's ratio for random matrices, that is for any $\tau > 0$,

$$L_n(\tau) := \frac{1}{n^2} \sum_{j,l=1}^{n} \mathbb{E} X_{jl}^2 I_{\{|X_{jl}| > \tau\sqrt{n}\}}. \tag{51}$$

**Theorem 6.5.** *Assume that the random variables* $X_{jl}, 1 \le j \le l \le n,\ n \ge 1$ *satisfy the following conditions*

$$\mathbb{E}\{X_{jl}|\mathcal{F}^{jl}\} = 0, \tag{52}$$

$$\varepsilon_n^{(1)} := \frac{1}{n^2} \sum_{1 \le j \le l \le n} \mathbb{E}|\mathbb{E}\{X_{jl}^2|\mathcal{F}^j\} - \sigma_{jl}^2| \to 0 \quad as \quad n \to \infty, \tag{53}$$

*there exists* $\sigma^2 > 0,$ *such that*

$$\varepsilon_n^{(2)} := \frac{1}{n^2} \sum_{1 \le j \le l \le n} |\sigma_{jl}^2 - \sigma^2| \to 0 \quad as \quad n \to \infty, \tag{54}$$

*and*

*for any fixed* $\tau > 0$,

$$L_n(\tau) \to 0 \quad as \quad n \to \infty. \tag{55}$$

*Then*

$$\Delta_n := \sup_x |\mathbb{E} F_n(x) - G(x\sigma^{-1})| \to 0 \quad as \quad n \to \infty. \tag{56}$$

**Corollary 6.6.** *Let* $X_{lj}^{(n)}$, $1 \le l \le j \le n$ *be distributed uniformly in the ball of the radius* $\sqrt{N}$ *in* $\mathbb{R}^N$ *with* $N = \frac{n(n+1)}{2}$ *, for any* $n \ge 1$*. Then*

$$\Delta_n \to 0, \quad as \quad n \to \infty. \tag{57}$$

## 6.6 Sample Covariance Matrices

Let $X_{jk}, 1 \le j, k < \infty$, be random variables with $\mathbb{E} X_{jk} = 0$ and $\mathbb{E} X_{jk}^2 = \sigma_{jk}^2$. For fixed $n \ge 1$ and $m \ge 1$, we introduce a matrix $n \times m$

$$\mathbf{X} = \left(X_{lj}\right)_{1 \le l \le n,\ 1 \le j \le m}. \tag{58}$$

Denote by $\lambda_1 \le \ldots \le \lambda_n$ the eigenvalues of the symmetric $n \times n$ matrix

$$\mathbf{W}_n = \frac{1}{p}\mathbf{X}\mathbf{X}^T, \tag{59}$$

and define its empirical spectral distribution function by

$$F_n(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{I}\{\lambda_j \le x\}. \tag{60}$$

We investigate the convergence of the expected spectral distribution function $\mathbb{E}F_n(x)$ to the distribution function of the Marchenko-Pastur law.

Let $g_\alpha(x)$ and $G_\alpha(x)$ denote the density and the distribution function of the Marchenko-Pastur law with parameter $\alpha \in (0, \infty)$, that is

$$g_\alpha(x) = \frac{1}{x\pi}\sqrt{(x-a)(b-x)}\; I_{\{x\in[a,b]\}}, \quad G_\alpha(x) = \int_{-\infty}^{x} g_\alpha(u)du, \tag{61}$$

where $a = (1-\sqrt{\alpha})^2$, $b = (1+\sqrt{\alpha})^2$.

## 6.7 Stein's Equation for the Marchenko-Pastur Law

Introduce a class of functions

$$\mathbb{C}^1_{\{a,b\}} = \{f : \mathbb{R} \to \mathbb{R} : f \in \mathbb{C}^1(\mathbb{R}\setminus\{a,b\});$$
$$\overline{\lim}_{|y|\to\infty}|yf(y)| < \infty;\ \limsup_{y\to\frac{a-b}{2}\pm\frac{a+b}{2}} |(\frac{(a-b)^2}{4} - (y - \frac{a+b}{2})^2||f'(y)| < C\}.$$

At first we state the following

**Lemma 6.7.** *Let $\alpha \ne 1$. Assume that a bounded function $\varphi(x)$ without discontinuity of second order satisfies the following conditions*

$$\varphi(x) \textit{ is continuous in the points } x = a, x = b \tag{62}$$

*and*

$$\int_a^b \varphi(u)g_\alpha(u)du = 0. \tag{63}$$

*Then there exists a function $f \in \mathbb{C}^1_{\{a,b\}}$ such that, for any $x \ne a$ or $x \ne b$,*

$$(x-a)(b-x)xf'(x) - 3x(x - \frac{a+b}{2})f(x) = \varphi(x). \tag{64}$$

*If $\varphi(a) = 0$ ($\varphi(b) = 0$) then there exists a continuous solution of the equation (64).*

**Proposition 6.8.** *The random variable $\xi$ has distribution function $G_\alpha(x)$ if and only if the following equality holds, for any function $f \in \mathbb{C}^1_{\{a,b\}}$,*

$$\mathbb{E}\left((\xi-a)(b-\xi)\xi f'(\xi) - 3\xi(\xi - \frac{a+b}{2})f(\xi)\right) = 0. \tag{65}$$

## 6.8 Stein's Criterion for Sample Covariance Matrices

Let $\mathbf{W}$ denote a sample covariance matrix with eigenvalues $0 \le \lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$. If $\mathbf{W} = \mathbf{U}^{-1}\mathbf{\Lambda}\mathbf{U}$, where $\mathbf{U}$ is an orthogonal and $\mathbf{\Lambda}$ a diagonal matrix, one defines $f(\mathbf{W}) = \mathbf{U}^{-1}f(\mathbf{\Lambda})\mathbf{U}$, where $f(\mathbf{\Lambda}) = \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n))$.

We can now formulate the convergence to the Marchenko-Pastur law for the spectral distribution function of random matrices.

**Theorem 6.9.** *Let $\mathbf{W}_n$ denote a sequence of sample covariance matrices of order $n \times n$ such that, for any function $f \in \mathbb{C}^1_{\{a,b\}}$*

$$\frac{1}{n}\mathbb{E}\mathrm{Tr}(\mathbf{W}_n - a\mathbf{I}_n)(b\mathbf{I}_n - \mathbf{W}_n)\mathbf{W}_n f'(\mathbf{W}_n) - \frac{3}{n}\mathbb{E}\mathrm{Tr}\mathbf{W}_n(\mathbf{W}_n - \frac{a+b}{2}\mathbf{I}_n)f(\mathbf{W}_n) \to 0, \quad \text{as } n \to \infty. \tag{66}$$

*Then*

$$\Delta_n := \sup_x |\mathbb{E}F_n(x) - G_\alpha(x)| \to 0, \text{ as } n \to \infty. \tag{67}$$

## 6.9 Resolvent Criterion for Sample Covariance Matrices

Denote by $\mathbf{R}(z)$ the resolvent matrix for the sample covariance matrix $\mathbf{W}$.

**Proposition 6.10.** *Assume that, for any $v \ne 0$,*

$$\mathcal{R}_n(\mathbf{W})(z) := \frac{1}{n}\mathbb{E}\mathrm{Tr}\mathbf{W}(\mathbf{W} - a\mathbf{I})(b\mathbf{I} - \mathbf{W})\mathbf{R}^2(z) + \frac{3}{n}\mathbb{E}\mathrm{Tr}(\mathbf{W} - \frac{a+b}{2}\mathbf{I})\mathbf{W}\mathbf{R}(z) \to 0, \text{ as } n \to \infty \tag{68}$$

*uniformly on compacts sets in* $\mathbb{C}\setminus\mathbb{R}$. *Then*

$$\Delta_n \to 0, \text{ as } n \to \infty. \tag{69}$$

## *6.10 Convergence to the Marchenko-Pastur Distribution*

We shall assume that $\mathbb{E}X_{jl} = 0$ and $\sigma_{jl}^2 := \mathbb{E}X_{jl}^2$, for $1 \le j \le n$ and $1 \le l \le m$. Introduce $\sigma$-algebras $\mathcal{F}^{jl} = \sigma\{X_{kq} : 1 \le k \le n, 1 \le q \le m, \{k,q\} \ne \{j,l\}\}$, $1 \le j \le n, 1 \le l \le m$, and $\mathcal{F}^{l} = \sigma\{X_{js} : 1 \le j \le n, 1 \le s \le m, s \ne l\}$, $1 \le l \le m$. We introduce as well Lindeberg's ratio for random matrices, that is for any $\tau > 0$,

$$L_n(\tau) = \frac{1}{nm}\sum_{j=1}^{n}\sum_{l=1}^{m}\mathbb{E}X_{jl}^2\mathbb{I}\{|X_{jl}| > \tau\sqrt{n}\}, \tag{70}$$

as well as the notation $X_{jl}^{(\tau)} := X_{jl}\mathbb{I}\{|X_{jl}| \le \tau\sqrt{n}\} - \mathbb{E}\mathbb{X}_{jl}\mathbb{I}\{|X_{jl}| \le \tau\sqrt{n}\}$, $\xi_{jl}^{(\tau)} := \mathbb{E}\Big\{X_{jl}^{(\tau)}\Big|\mathcal{F}^{(jl)}\Big\}$. Introduce also the vectors $\mathbf{X}_l^{(\tau)} = (X_{1,l}^{(\tau)}, \dots, X_{n,l}^{(\tau)})^T$ and $\,_{l}^{(\tau)} = (\xi_{1,l}^{(\tau)}, \dots, \xi_{n,l}^{(\tau)})^T$.

**Theorem 6.11.** *Let* $m = m(n)$ *depend on n, such that*

$$\frac{m(n)}{n} \to \alpha \in (0,1), \text{ as } n \to \infty. \tag{71}$$

*Assume that the random variables* $X_{jl}, 1 \le j \le n, 1 \le l \le m$,
$n, m \ge 1$ *satisfy the following conditions*

$$\mathbb{E}\{X_{jl}|\mathcal{F}^{jl}\} = 0, \tag{72}$$

$$\varepsilon_n^{(1)} := \frac{1}{nm}\sum_{j=1}^{n}\sum_{l=1}^{m}\mathbb{E}|\mathbb{E}\{X_{jl}^2|\mathcal{F}^l\} - \sigma_{jl}^2| \to 0 \quad \text{as} \quad n \to \infty, \tag{73}$$

*there exists* $\sigma^2 > 0$, *such that*

$$\varepsilon_n^{(2)} := \frac{1}{nm}\sum_{j=1}^{n}\sum_{l=1}^{m}|\sigma_{jl}^2 - \sigma^2| \to 0, \tag{74}$$

$$\varepsilon_n^{(3)} := \frac{1}{nm^2}\sum_{l=1}^{m}\sum_{j,k=1}^{n}\mathbb{E}\left|\mathbb{E}\{((X_{jl}^{(\tau)})^2 - \mathbb{E}(X_{jl}^{(\tau)})^2)((X_{kl}^{(\tau)})^2 - \mathbb{E}(X_{kl}^{(\tau)})^2)\Big|\mathcal{F}^l\}\right| \to 0, \tag{75}$$

$$\varepsilon_n^{(4)} := \frac{1}{nm^2}\sum_{l=1}^{m}\sum_{1\le j\ne k\le n}\mathbb{E}\Big|(\mathbb{E}\{((\xi_{jl}^{(\tau)})^2-\mathbb{E}(\xi_{jl}^{(\tau)})^2)$$
$$\times\ ((\xi_{kl}^{(\tau)})^2-\mathbb{E}(\xi_{kl}^{(\tau)})^2)\Big|\mathcal{F}^l\}\Big| \to 0, \quad as\, n\to\infty, \tag{76}$$

*and*

$$L_n(\tau)\to 0, \quad \text{for any fixed } \tau>0, \quad as \quad n\to\infty. \tag{77}$$

*Then*

$$\Delta_n := \sup_x |\mathbb{E}F_n(x) - G_\alpha(x\sigma^{-1})| \to 0 \quad as \quad n\to\infty. \tag{78}$$

*Remark 6.12.* Note that condition (74) implies that

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\mathrm{Tr}\,\mathbf{W}_n = \sigma^2 < \infty. \tag{79}$$

**Corollary 6.13.** *Assume (71). Let, for any $n,m\ge 1$, $X_{jl}$, $1\le j\le n, 1\le l\le m$, be independent and $\mathbb{E}X_{jl}=0$, $\mathbb{E}X_{jl}^2=\sigma^2$. Suppose that, for any fixed $\tau>0$,*

$$L_n(\tau)\to 0, \quad as \quad n\to\infty. \tag{80}$$

*Then the expected spectral distribution function of the sample covariance matrix W converges to the Marchenko-Pastur distribution,*

$$\Delta_n := \sup_x |\mathbb{E}F_n(x) - G_\alpha(x\sigma^{-1})| \to 0, \quad as \quad n\to\infty. \tag{81}$$

**Acknowledgements** Alexander Tikhomirov partially was supported by SFB 701, by grants of RFBR N 11-01-00310-a, N 11-01-122104-ofi-m-2011, and by Program of Basic Research of Urals Division of RAS.

## References

1. N. Alexeev, F. Götze, A.N. Tikhomirov, On the singular spectrum of powers and products of random matrices. Dokl. Math. **82**(1), 505–507 (2010)
2. N. Alexeev, F. Götze, A.N. Tikhomirov, Asymptotic distribution of singular values of powers of random matrices. Lithuanian math. J. **50**(2), 121–132 (2010)
3. N. Alexeev, F. Götze, A.N. Tikhomirov, On the asymptotic distribution of singular values of products of large rectangular random matrices. J. Math. Sci. **408**, 9–43 (2012 (Zapiski nauchnyh seminarov POMI (in Russia)). arXiv:1012.2586
4. N. Alexeev, F. Götze, A.N. Tikhomirov, On the asymptotic distribution of the singular values of powers of random matrices. Preprint, arXiv:1012.2743
5. Z.D. Bai, Convergence rate of expected spectral distributions of large random matrices. I. Wigner matrices. Ann. Probab. **21**(2), 625–648 (1993)
6. Z.D. Bai, Convergence rate of expected spectral distributions of large random matrices. II. Sample covariance matrices. Ann. Probab. **21**(2), 649–672 (1993)

7. Z.D. Bai, Circular law. Ann. Probab. **25**, 494–529 (1997)
8. Z.D. Bai, Methodologies in spectral analysis of large-dimensional random matrices, a review. With comments by G. J. Rodgers and Jack W. Silverstein; and a rejoinder by the author. Statist. Sinica **9**(3), 611–677 (1999)
9. Z.D. Bai, B. Miao, J. Tsay, Convergence rates of the spectral distributions of large Wigner matrices. Int. Math. J. **1**(1), 65–90 (2002)
10. Z.D. Bai, B. Miao, J.-F. Yao, Convergence rates of spectral distributions of large sample covariance matrices. SIAM J. Matrix Anal. Appl. **25**(1), 105–127 (2003)
11. S.G. Bobkov, F. Götze, A.N. Tikhomirov, On concentration of empirical measures and convergence to the semi-circle law. J. Theoret. Probab. **23**(3), 792–823 (2010)
12. A. Edelman, The probability that a random real Gaussian matrix has $k$ real eigenvalues, related distributions, and circular law. J. Mult. Anal. **60**, 203–232 (1997)
13. L. Erdös, H.-T. Yau, J. Yin, Rigidity of eigenvalues of generalized Wigner matrices. Adv. Math. **229**(3), 1435–1515 (2012) arXiv:1007.4652.
14. J. Ginibre, Statistical ensembles of complex, quaternion, and real matrices. J. Math. Phys. **6**, 440–449 (1965)
15. V.L. Girko, The circular law. (Russian) Teor. Veroyatnost. i Primenen. **29**(4), 669–679 (1984)
16. V.L. Girko, Spectral theory of random matrices. (Russian) Uspekhi Mat. Nauk **40** 1(241), 67–106 (1985)
17. V.L. Girko, Asymptotics of the distribution of the spectrum of random matrices. (Russian) Uspekhi Mat. Nauk **44** 4(268), 7–34, 256 (1989); translation in Russ. Math. Surv. **44**(4), 3–36 (1989)
18. V.L. Girko, Extended proof of the statement: convergence rate of expected spectral functions of the sample covariance matrix $\hat{R}_{m_n}(n)$ is equal to $O(n^{-1/2})$ under the condition $\frac{m_n}{n} \leq c < 1$ and the method of critical steepest descent. Random Oper. Stoch. Equat. **10**(4), 351–405 (2002)
19. F. Götze, A. Tikhomirov, The circular law for random matrices. Ann. Probab. **38**(4), 1444–1491 (2010)
20. F. Götze, A.N. Tikhomirov, Rate of convergence to the semi-circular law. Probab. Theor. Relat. Fields **127**, 228–276 (2003)
21. F. Götze, A.N. Tikhomirov, Rate of convergence in probability to the Marchenko-Pastur law. Bernoulli **10**(3), 503–548 (2004)
22. F. Götze, A.N. Tikhomirov, *Limit Theorems for Spectra of Random Matrices with Martingale Structure. Stein's Method and Applications.* Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., vol. 5 (Singapore University Press, Singapore, 2005), pp. 181–193
23. F. Götze, A.N. Tikhomirov, The rate of convergence for spectra of GUE and LUE matrix ensembles. Cent. Eur. J. Math. **3**(4), 666–704 (2005)
24. F. Götze, A. Tikhomirov, Limit theorems for spectra of positive random matrices under dependence. Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **311**, Veroyatn. i Stat. 7, 92–123, 299 (2004); translation in J. Math. Sci. (N. Y.) **133**(3), 1257–1276 (2006)
25. F. Götze, A.N. Tikhomirov, Limit theorems for spectra of random matrices with martingale structure. Teor. Veroyatn. Primen. **51**(1), 171–192 (2006); translation in Theor. Probab. Appl. **51**(1), 42–64 (2007)
26. F. Götze, A.N. Tikhomirov, The rate of convergence of spectra of sample covariance matrices. Teor. Veroyatn. Primen. **54**(1), 196–206 (2009); translation in Theor. Probab. Appl. **54**(1), 129–140 (2010)
27. F. Götze, A.N. Tikhomirov, On the Rate of Convergence to the Marchenko–Pastur Distribution. Preprint, arXiv:1110.1284
28. F. Götze, A.N. Tikhomirov, The rate of convergence to the semi-circular law, in *High Dimensional Probability, VI*. Progress in Probability, vol. 66 (Birkhaeuser, Basel, 2013), pp. 141–167. arXiv:1109.0611
29. V. Marchenko, L. Pastur, The eigenvalue distribution in some ensembles of random matrices. Math. USSR Sbornik **1**, 457–483 (1967)
30. G. Pan, W. Zhou, Circular law, extreme singular values and potential theory. J. Multivariate Anal. **101**(3), 645–656 (2010). arXiv:0705.3773

31. M. Rudelson, Invertibility of random matrices: norm of the inverse. Ann. Math. (2) **168**(2), 575–600 (2008). http://arXiv:math/0507024.
32. M. Rudelson, R. Vershynin, The Littlewood–Offord problem and invertibility of random matrices. Adv. Math. **218**(2), 600–633 (2008). http://arXiv.org/abs/math.PR/0703307.
33. T. Tao, V. Vu, Random matrices: the circular law. Comm. Contemp. Math. **10**(2), 261–307 (2008)
34. T. Tao, V. Vu, Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. Ann. Math. **169**(2), 595–632 (2009). http://arXiv:math/0511215.
35. T. Tao, V. Vu, Random matrices: universality of ESDs and the circular law. With an appendix by Manjunath Krishnapur. Ann. Probab. **38**(5), 2023–2065 (2010)
36. A.N. Tikhomirov, On the rate of convergence of the expected spectral distribution function of a Wigner matrix to the semi-circular law. Siberian Adv. Math. **19**(3), 211–223 (2009)
37. A.N. Tikhomirov, The rate of convergence of the expected spectral distribution function of a sample covariance matrix to the Marchenko-Pastur distribution. Siberian Adv. Math. **19**(4), 277–286 (2009)
38. D.A. Timushev, A.N. Tikhomirov, A.A. Kholopov, On the accuracy of the approximation of the GOE spectrum by the semi-circular law. (Russian) Teor. Veroyatn. Primen. **52**(1), 180–185 (2007); translation in Theor. Probab. Appl. **52**(1), 171–177 (2008)
39. E.P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions. Ann. Math. (2) **62**, 548–564 (1955)
40. E.P. Wigner, On the distribution of the roots of certain symmetric matrices. Ann. Math. (2) **67**, 325–327 (1958)
41. J. Wishart, The generalised product moment distribution in samples from a normal multivariate population. Biometrika **20A**(1–2), 32–52 (1928)