J Theor Probab (2010) 23: 792–823
DOI 10.1007/s10959-010-0286-7

# On Concentration of Empirical Measures and Convergence to the Semi-circle Law

S.G. Bobkov · F. Götze · A.N. Tikhomirov

Received: 27 October 2008 / Revised: 15 December 2009 / Published online: 29 April 2010


**Abstract** Concentration properties of the general empirical distribution functions and the rate of convergence of spectral empirical distributions to the semi-circle law in the case of symmetric high-dimensional random matrices are studied under Poincaré-type inequalities.

**Keywords** Random matrices · Spectrum · Empirical distributions · Semi-circle law · Poincaré-type inequalities

**Mathematics Subject Classification (2000)** Primary 60E

## 1 Introduction

Let $\{\xi_{jk}\}_{1\le j\le k\le n}$ be a family of independent random variables on some probability space with mean $\mathbf{E}\xi_{jk} = 0$ and variance $\mathrm{Var}(\xi_{jk}) = 1$. Put $\xi_{jk} = \xi_{kj}$ for

Research partially supported by NSF grant DMS-0706866, SFB 701, RFBR grant N07-01-00583-a, and RF grant of the leading scientific schools NSh-638.2008.1.

S.G. Bobkov (✉)
School of Mathematics, University of Minnesota, Minneapolis, USA
e-mail: bobkov@math.umn.edu

F. Götze
Faculty of Mathematics, University of Bielefeld, Bielefeld, Germany
e-mail: goetze@math.uni-bielefeld.de

A.N. Tikhomirov
Department of Mathematics of Komi Research Center of RAS, Syktyvkar, Russia
e-mail: tichomir@math.uni-bielefeld.de

$1 \le k < j \le n$. Then we have a symmetric $n \times n$ random matrix

$$\mathbf{W} = \frac{1}{\sqrt{n}} \begin{pmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1n} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{n1} & \xi_{n2} & \cdots & \xi_{nn} \end{pmatrix}.$$

Introduce its (real random) eigenvalues $X_1 \le \ldots \le X_n$ and define an empirical (spectral) distribution function $F_n(x) = \frac{1}{n} \operatorname{card}\{i \le n : X_i \le x\}$ together with the average marginal distribution function of the spectrum

$$F(x) = \mathbf{E} F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{P}\{X_i \le x\}, \quad x \in \mathbf{R}.$$

Under mild moment assumptions on the entries $\xi_{jk}$, and when the dimension $n$ is large, $F$ is known to be close to the distribution function of the standard semi-circle law $G$ (Wigner's law, cf. e.g. [29, 30]). Here $G$ has density $g(x) = \frac{1}{2\pi} \sqrt{4 - x^2}$, $-2 < x < 2$. When quantifying asymptotic approximation, one often uses the Kolmogorov distance

$$\Delta_n = \|F - G\| = \sup_x |F(x) - G(x)|.$$

More information about the rate of fluctuations is encoded in the distance

$$\Delta_n^* = \mathbf{E}\,\|F_n - G\| = \mathbf{E} \sup_x |F_n(x) - G(x)|.$$

A difficult open problem concerning the Wigner theorem is to determine the rates at which $\Delta_n$ and $\Delta_n^*$ tend to zero for growing dimension. The only case where the asymptotic of $\Delta_n$ is known (to be of order $1/n$) is when the entries have a normal distribution [20]. This case is special, since here the joint distribution of the spectrum is known explicitly, and one may apply the technique of orthogonal polynomials. The general picture is however unclear, and we do not know, for example, whether the behavior of $\Delta_n$ has a universal character (as in the classical central limit theorem), or it essentially reflects the underlying distributions of the entries. Nevertheless, under various moment hypotheses, upper bounds for $\Delta_n$ and $\Delta_n^*$ are known; see for example [2, 3, 18]. The best known general result in this direction is the estimate $\Delta_n \le C/\sqrt{n}$ in case of bounded 4-th moments of the entries, and similarly for $\Delta_n^*$, however, under a slightly stronger moment assumption, cf. [3, 19]. There are also other results, quantifying in particular the closeness of $F_n$ to $G$ on very small intervals of length of order $(\log n)^\alpha/n$ using logarithmic Sobolev inequalities, cf. [16, 17].

Another interesting related problem is the concentration property of spectral empirical distributions and, as part of it, the problem of the rates of the Kolmogorov distance $\|F_n - F\|$. It represents a special case of a more general scheme, where one deals with independent or dependent observations $X_1, \ldots, X_n$, taken from a law on $\mathbf{R}^n$. The joint distribution of eigenvalues is a natural example, and it motivates us to study deviations of $F_n$ from the mean $F$ under analytical hypotheses, such as Poincaré or logarithmic Sobolev inequalities.

Let us recall that a probability measure $\mu$ on $\mathbf{R}^d$ is said to satisfy a Poincaré-type or spectral gap inequality with constant $\sigma^2$ ($\sigma \geq 0$) if for any bounded smooth function $g$ on $\mathbf{R}^d$ with gradient $\nabla g$,

$$\operatorname{Var}(g) \leq \sigma^2 \int |\nabla g|^2 \, d\mu, \tag{1.1}$$

where $\operatorname{Var}(g) = \int g^2 \, d\mu - (\int g \, d\mu)^2$ stands for the variance of $g$ under the measure $\mu$. In this case we write $\mathrm{PI}(\sigma^2)$ for short. This and related hypotheses have proved to be rather useful in the study of various forms of concentration of spectral empirical distributions, as shown, for example, in the papers by A. Guionnet and O. Zeitouni [23], S. Chatterjee and A. Bose [14], and K.R. Davidson and S.J. Szarek [15]; cf. also [8, 25, 26]. A remarkable feature of this approach in the spectral analysis is that no specific knowledge about the non-explicit mapping from a random matrix to its spectrum is required, and instead, it suffices to use only general Lipschitz properties, which are satisfied by this mapping.

In this paper we show that in the matrix model described above, the amount of concentration of the spectral empirical distribution $F_n$ around its mean $F$ (a property which may be expressed, for example, in terms of Stieltjes transforms of $F_n$) is in a certain sense responsible for the rates of $\Delta_n$ and $\Delta_n^*$. Using this observation, we prove the following.

**Theorem 1.1** *If the distributions of $\xi_{jk}$'s satisfy the Poincaré-type inequality* $\mathrm{PI}(\sigma^2)$ *on the real line*, *then*

$$\Delta_n \leq C n^{-2/3}, \tag{1.2}$$

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

$$\Delta_n^* \leq C n^{-2/3} \log^2(n+1). \tag{1.3}$$

A bound similar to (1.2) is known to hold in the case where $\xi_{jk}$ have a common distribution with a non-trivial Gaussian component [21]. In the purely Gaussian case, the right-hand side of (1.3) may be replaced with $C \log(n+1) \, n^{-2/3}$ [31]. In fact, it follows from results of [8] that this mild improvement may be obtained when the $\xi_{jk}$'s satisfy a logarithmic Sobolev inequality (a hypothesis which is stronger than PI). See also Remark 7.4 below for comments on the class of probability distributions on the line satisfying Poincaré-type inequalities.

Although the quantity $\Delta_n$ may be related to the concentration property of spectral empirical distributions, we first study deviations of $F_n$ from $F$ in terms of the Levý distance $L(F_n, F)$ and then consider the deviations from $G$ in terms of the Kolmogorov distance $\|F_n - G\|$ via $\Delta_n$. This may be done for a general model of observations $X = (X_1, \dots, X_n)$ whose joint distribution satisfies a Poincaré-type inequality. At some step we then arrive at the following.

**Theorem 1.2** *Assume the random vector $X$ has a distribution satisfying* $\mathrm{PI}(\sigma_n^2)$ *on* $\mathbf{R}^n$. *For any distribution function $G$ with finite Lipschitz semi-norm $M = \|G\|_{\mathrm{Lip}}$,*

$$\mathbf{E}\,\|F_n - G\| \le C\left[\|F - G\| + \left(\frac{\sigma_n^2}{n}\right)^{1/3}\right] \log^2\left(2 + \frac{(1 + A_n)\,n}{\sigma_n^2}\right), \tag{1.4}$$

*where* $A_n = \frac{1}{\sigma_n} \max_{j,k} |\mathbf{E}X_j - \mathbf{E}X_k|$, *and where the constant $C$ depends on $M$ only.*

The Lipschitz semi-norm is defined in the usual way as $\|G\|_{\mathrm{Lip}} = \sup_{x<y} \frac{G(y)-G(x)}{y-x}$, and its finiteness is equivalent to the property that $G$ is absolutely continuous and has a bounded density $g$, and then necessarily $\mathrm{ess\,sup}_x\, g(x) = \|G\|_{\mathrm{Lip}}$. Theorem 1.2 will be used by taking for $G$ the standard semi-circle law.

In the matrix situation of Theorem 1.1, we have $\sigma_n^2 = 2\sigma^2/n$, while $A_n$ is of order at most $\sqrt{n}$. Hence, Theorem 1.2 allows us to reduce (1.3) to the inequality (1.2). Moreover, as we will see, the proof of (1.2) itself is essentially based on the concentration of Stieltjes transforms of $F_n$, and such a property appeals to using Poincaré-type inequalities, as well as the bound (1.4) of Theorem 1.2.

In this connection, let us mention a result of S. Chatterjee and A. Bose [14], who used Fourier transforms to derive from $\mathrm{PI}(\sigma_n^2)$ a bound of a similar nature for the Wigner matrix model $\mathbf{E}\,\|F_n - G\| \le 2\,\|F - G\| + C(\frac{\sigma_n^2}{n})^{1/4}$.

The paper is organized as follows. In Sects. 2, 3 we consider general bounds on the Lévy and Kolmogorov distances between distribution functions in terms of their Stieltjes transforms. In Sects. 4, 5 we discuss applications of the Poincaré-type inequalities to linear functionals of the empirical measures and specialize them to the case of Stieltjes transforms. The results thus obtained are applied in Sect. 6 to explore the concentration property of the empirical measures in terms of the Lévy and Kolmogorov distances. In particular, Theorem 1.2 is proved. In Sect. 7 general results are specialized to the matrix case; in particular, we clarify the relationship between Theorems 1.2 and 1.1. Inequality (1.2) of Theorem 1.1 is proved under concentration assumptions in Sect. 8. The proof of a combinatorial Proposition 5.1 for the moments of Stieltjes transforms is postponed to the Appendix.

## 2 Bounds on Lévy Distance in Terms of Stieltjes Transform

Given a distribution function $F(x)$, $x \in \mathbf{R}$, its Stieltjes transform is defined as the analytic function in the upper complex half-plane

$$S_F(z) = \int_{-\infty}^{+\infty} \frac{1}{x - z}\, dF(x),$$

where $z = u + iv$, $u \in \mathbf{R}$, $v = \mathrm{Im}(z) > 0$.

In this section we derive bounds on the Lévy distance in terms of the Stieltjes transform. Given distribution functions $F$ and $G$ on the real line, the Lévy distance $L(F, G)$ between $F$ and $G$ is defined as

$$L(F, G) = \inf\{\delta \ge 0 : F(x - \delta) - \delta \le G(x) \le F(x + \delta) + \delta,\ \forall x \in \mathbf{R}\}.$$

**Proposition 2.1** *Let $F$ and $G$ be distribution functions. Given $v > 0$, let an interval $[\alpha, \beta] \subset \mathbf{R}$ be chosen to satisfy $G(\alpha) \le v$ and $1 - G(\beta) \le v$. Then*

$$L(F, G) \le \sup_{x \in [\alpha - 2v,\, \beta + 2v]} \left| \int_{-\infty}^{x} \mathrm{Im}(S_F(z) - S_G(z))\, du \right| + 4v + 50\, v \sup_{x \in \mathbf{R}} \mathrm{Im}(S_G(x + iv)).$$

Let H denote the distribution function of the standard Cauchy law, and let $\mathrm{H}_v$ denote the distribution function of the Cauchy law with parameter $v > 0$, that is, with density

$$h_v(x) = \frac{d\mathrm{H}_v(x)}{dx} = \frac{v}{\pi(x^2 + v^2)}, \qquad x \in \mathbf{R}.$$

When $F$ has density $f$, the function $u \to \frac{1}{\pi} \mathrm{Im}(S_F(u + iv))$ represents the convolution $f * h_v$ in the classical sense. In general, for any $x \in \mathbf{R}$, with previous notation $z = u + iv$ we have the relation

$$\frac{1}{\pi} \int_{-\infty}^{x} \mathrm{Im}(S_F(z))\, du = \int_{-\infty}^{+\infty} F(x - y)\, d\mathrm{H}_v(y),$$

which in the case of two distribution functions yields the identity

$$\int_{-\infty}^{x} \mathrm{Im}(S_F(u + iv) - S_G(u + iv))\, du = v \int_{-\infty}^{+\infty} \frac{F(x - y) - G(x - y)}{y^2 + v^2}\, dy. \tag{2.1}$$

As for the Lévy distance, it may be related to more accessible quantities. For $\delta \ge 0$ introduce

$$L_\delta^+(F, G) = \sup_x [F(x) - G(x + \delta)], \qquad L_\delta^-(F, G) = \sup_x [G(x) - F(x + \delta)],$$

and define

$$L_\delta(F, G) = \max\{L_\delta^+(F, G), L_\delta^-(F, G)\}.$$

As a result, we obtain a family of metric-like functionals $L_\delta$ for the space of all distribution functions on the real line, and one of them $L_0(F, G) = \sup_x |F(x) - G(x)|$ represents the (uniform) Kolmogorov distance. The relationship with the Lévy distance is given by $L(F, G) \le \delta \Leftrightarrow L_\delta(F, G) \le \delta$. In addition, for all $\delta, \delta' \ge 0$,

$$L_\delta(F, G) \le \delta' \Longrightarrow L(F, G) \le \max\{\delta, \delta'\}. \tag{2.2}$$

With these preparations, we are ready to turn to the proof of Proposition 2.1.

*Proof of Proposition 2.1.* Our task is to show a lower bound of the integral on the right-hand side of (2.1). Let's call it $I(x)$. Given a parameter $a > 0$, which will be chosen later on, we split $I(x)$ into the two regions $|y| < av$ and $|y| > av$,

$$I_0 = v \int_{|y| < av} \frac{F(x - y) - G(x - y)}{y^2 + v^2}\, dy, \quad I_1 = v \int_{|y| > av} \frac{F(x - y) - G(x - y)}{y^2 + v^2}\, dy.$$

Changing the variable $y = vt$ and using monotonicity of the distribution functions, for the first region we get

$$\begin{aligned} I_0 &= \int_{|t|<a} \frac{F(x-vt) - G(x-vt)}{t^2+1}\,dt \\ &\geq \int_{|t|<a} \frac{F(x-av) - G(x+av)}{t^2+1}\,dt \;=\; (F(x-av) - G(x+av)) \cdot \pi\gamma, \end{aligned} \tag{2.3}$$

where

$$\gamma = \mathrm{H}(-a,a) = \frac{1}{\pi} \int_{|t|<a} \frac{1}{t^2+1}\,dt$$

is the Cauchy measure of the interval $(-a,a)$. Similarly, for the second region we may write

$$\begin{aligned} I_1 &= \int_{|t|>a} \frac{F(x-vt) - G(x-vt)}{t^2+1}\,dt \\ &= \int_{|t|>a} \frac{F(x-vt) - G(x-vt+2av)}{t^2+1}\,dt \\ &\quad + \int_{|t|>a} \frac{G(x-vt+2av) - G(x-vt)}{t^2+1}\,dt \qquad (2.4) \\ &\leq L_{2av}(F,G) \cdot \pi(1-\gamma) + J, \qquad (2.5) \end{aligned}$$

where $J$ denotes the second integral in (2.4). In order to estimate this integral, assume temporarily that $G$ has a density $g$. Then extending this integral to the whole real line, we have

$$\begin{aligned} J &\leq \iint_{x-vt<s<x-vt+2av} \frac{g(s)}{t^2+1}\,dt\,ds \\ &= \pi v \int_{-\infty}^{+\infty} g(x+vt)\,(\mathrm{H}(t) - \mathrm{H}(t-2a))\,dt. \end{aligned} \tag{2.6}$$

Now, simple calculus arguments show that

$$\mathrm{H}(t) - \mathrm{H}(t-2a) \leq C\,\frac{1}{\pi(t^2+1)}, \quad \text{for all } t \text{ real},$$

with constant $C = 2a\,(a + \sqrt{1+a^2}\,)^2$. Indeed, by the mean value theorem, we have $\mathrm{H}(t) - \mathrm{H}(t-2a) = \frac{2a}{\pi(s^2+1)}$, for some point $s \in (t-2a,t)$. Hence, one may take

$$C \;=\; 2a \sup_{t\geq 0}\ \sup_{t-2a<s<t} \frac{t^2+1}{s^2+1} \;=\; 2a \sup_{t\geq 2a} \frac{t^2+1}{(t-2a)^2+1}.$$

By direct differentiation, the last supremum is attained for $t = a + \sqrt{1+a^2}$.

Using this in (2.6), we obtain

$$J \le Cv \int_{-\infty}^{+\infty} \frac{g(x+vt)}{t^2+1}\,dt$$
$$= Cv^2 \int_{-\infty}^{+\infty} \frac{g(x-y)}{y^2+v^2}\,dy = Cv \int_{-\infty}^{+\infty} \frac{v}{(x-y)^2+v^2}\,dG(y).$$

In other words,

$$J \le Cv\,\mathrm{Im}(S_G(x+iv)). \tag{2.7}$$

Clearly (by Fatou's lemma, for example, applied to the functional $J$), the inequality (2.7) is stable under weak limits of probability measures on the line. Therefore, it may be extended to all distribution functions $G$, regardless of whether they are absolutely continuous or not.

Now, applying (2.7) in (2.5), we get

$$I_1 \le L_{2av}(F,G)\cdot 2\pi(1-\gamma) + Cv\,\mathrm{Im}(S_G(x+iv)). \tag{2.8}$$

In a similar manner,

$$-I_1 = \int_{|t|>a} \frac{G(x-vt)-G(x-vt-2av)}{t^2+1}\,dt$$
$$+ \int_{|t|>a} \frac{G(x-vt-2av)-F(x-vt)}{t^2+1}\,dt$$
$$\le L_{2av}(F,G)\cdot \pi(1-\gamma) + J'$$

with $J'$ satisfying (2.7), as well. Hence, (2.8) may be replaced with

$$|I_1| \le L_{2av}(F,G)\cdot \pi(1-\gamma) + Cv\,\mathrm{Im}(S_G(x+iv)).$$

Now, combining this with (2.3) and using $I(x) = I_0 + I_1 \ge I_0 - |I_1|$, we get

$$I(x) \ge (F(x-av) - G(x+av))\cdot \pi\gamma$$
$$- L_{2av}(F,G)\cdot \pi(1-\gamma) - Cv\,\mathrm{Im}(S_G(x+iv)). \tag{2.9}$$

The next step is to interchange the role of $F$ and $G$, still keeping the quantity $\mathrm{Im}(S_G(x+iv))$ on the right. Indeed, the analogue of (2.3) would be

$$-I_0 \ge (G(x-av) - F(x+av))\cdot \pi\gamma.$$

Using $-I(x) = -I_0 - I_1 \ge -I_0 - |I_1|$, we get another variant of (2.9),

$$-I(x) \ge (G(x-av) - F(x+av))\cdot \pi\gamma$$
$$- L_{2av}(F,G)\cdot \pi(1-\gamma) - Cv\,\mathrm{Im}(S_G(x+iv)).$$

Combining both inequalities, we obtain

$$\begin{aligned}
&\gamma \max\{F(x-av)-G(x+av), G(x-av)-F(x+av)\} \\
&\quad \le (1-\gamma)L_{2av}(F,G)+\frac{1}{\pi}\,|I(x)|+\frac{Cv}{\pi}\,\mathrm{Im}(S_G(x+iv)). \qquad (2.10)
\end{aligned}$$

By the very definition, the sup of the max in (2.10) over all $x$ equals $L_\delta(F,G)$ with $\delta = 2av$. However, we do not lose much by restricting that sup to the interval $\Delta = [\alpha - av, \beta + av]$. Indeed, if $x < \alpha - av$, by the assumption $G(\alpha) \le v$,

$$\begin{aligned}
F(x-av)-G(x+av) &\le F(\alpha-2av) = G(\alpha)+[F(\alpha-2av)-G(\alpha)] \\
&\le v + \sup_{x\in\Delta}[F(x-av)-G(x+av)].
\end{aligned}$$

In the case $x > \beta + av$, using $G(\beta) \ge 1 - v$, we have $F(x-av)-G(x+av) \le 1-G(\beta) \le v$. Hence,

$$L_\delta^+(F,G) \le v + \sup_{x\in\Delta}[F(x-av)-G(x+av)].$$

By a similar argument, $L_\delta^-(F,G) \le v + \sup_{x\in\Delta}[G(x-av)-F(x+av)]$. These two inequalities yield

$$L_\delta(F,G) \le v + \sup_{x\in\Delta} \max\{F(x-av)-G(x+av), G(x-av)-F(x+av)\}.$$

Using this bound in (2.10), we arrive at

$$\gamma\,(L_{2av}(F,G)-v) \le (1-\gamma)\,L_{2av}(F,G)+\frac{1}{\pi}\sup_{x\in\Delta}|I(x)|+\frac{Cv}{\pi}\sup_x \mathrm{Im}(S_G(x+iv)).$$

Therefore, if $\gamma > 1/2$ (equivalently, when $a > 1$),

$$\begin{aligned}
L_{2av}(F,G) &\le \frac{1}{\pi(2\gamma-1)}\sup_{x\in\Delta}|I(x)| \\
&\quad + \frac{\gamma v}{2\gamma-1}+\frac{Cv}{\pi(2\gamma-1)}\sup_x \mathrm{Im}(S_G(x+iv)).
\end{aligned}$$

Recalling property (2.2), we conclude that

$$\begin{aligned}
L(F,G) \le \max\Big\{&2av,\ \frac{1}{\pi(2\gamma-1)}\sup_{x\in[\alpha-av,\beta+av]}|I(x)| \\
&+\frac{\gamma v}{2\gamma-1}+\frac{Cv}{\pi(2\gamma-1)}\sup_x \mathrm{Im}(S_G(x+iv))\Big\}. \qquad (2.11)
\end{aligned}$$

It remains to choose a value for the parameter $a$ or $\gamma$. Note $a = \tan(\frac{\pi\gamma}{2})$. Taking, for example, $\gamma = \frac{2}{3}$, or equivalently, $a = \sqrt{3}$, we notice that the term $\frac{\gamma v}{2\gamma-1} = 2v$ is majorized by $2av < 4v$, while $\frac{1}{\pi(2\gamma-1)} = \frac{3}{\pi} < 1$. In addition,

$$C = 2a\left(a + \sqrt{1+a^2}\right)^2 = 22\sqrt{3} + 12 = 48.25\ldots$$

Thus, (2.11) gives

$$L(F,G) \le \sup_{x\in[\alpha-2v,\beta+2v]} |I(x)| + 4v + 50\,v \sup_x \operatorname{Im}(S_G(x+iv)),$$

which is the desired statement. □

*Contour integration* To estimate the Lévy distance $L(F,G)$ with the help of Proposition 2.1, it might be useful, as was noted in [19] for similar aims, to change in the integral

$$\int_{-\infty}^{x} \operatorname{Im}(S_F(z) - S_G(z))\,du$$

the contour of integration. Recall that $z = u + iv$ with $u \in \mathbf{R}$, $v = \operatorname{Im}(z) > 0$.

Note that the function $u \to \operatorname{Im}(S_F(z))$ is positive and integrable on the real line, with $\int_{-\infty}^{+\infty} \operatorname{Im}(S_F(u+iv))\,du = \pi$. Hence, for all $x$ real,

$$\int_{-\infty}^{x} \operatorname{Im}(S_F(u+iv))\,du = \operatorname{Im}\left\{\lim_{A\to+\infty}\int_{-A}^{x} S_F(u+iv)\,du\right\}.$$

Thus, we take a large positive $A > -x$, and for $v_1 > v_0 > 0$, introduce the rectangle with sides $C_1 = [-A,x] + iv_0$, $C_2 = A + i[v_0,v_1]$, $C_3 = [x,-A] + iv_1$, $C_4 = -A + i[v_1,v_0]$. By Cauchy's theorem, we have by contour integration

$$\int_{-A}^{x} (S_F(u+iv_0) - S_G(u+iv_0))\,du = -\sum_{i=2}^{4}\int_{C_i} (S_F(z) - S_G(z))\,dz, \tag{2.12}$$

and the same holds for the imaginary parts of the integrals. It is easy to see that the integral over $C_4$ in (2.12) is vanishing in the limit as $A \to +\infty$. The integral over $C_3$ may be bounded uniformly over all $x$ by $\int_{-\infty}^{+\infty} |S_F(u+iv_1) - S_G(u+iv_1)|\,du$. Hence, whenever $v_1 > v_0 > 0$, for all $x \in \mathbf{R}$,

$$\begin{aligned}\left|\int_{-\infty}^{x} \operatorname{Im}(S_F(u+iv_0) - S_G(u+iv_0))\,du\right| &\le \int_{-\infty}^{+\infty} |S_F(u+iv_1) - S_G(u+iv_1)|\,du \\ &\quad + \left|\int_{v_0}^{v_1} (S_F(x+iv) - S_G(x+iv))\,dv\right|. \end{aligned}\tag{2.13}$$

Combined with the bound of Proposition 2.1, (2.13) yields the following.

**Corollary 2.2** *Let $F$ and $G$ be arbitrary distribution functions. With some universal constant $c > 0$, for all $v_1 > v_0 > 0$,*

$$c\,L(F,G) \le v_0 + v_0 \sup_{x\in\mathbf{R}} \mathrm{Im}(S_G(x+iv_0)) + \int_{-\infty}^{+\infty} |S_F(u+iv_1) - S_G(u+iv_1)|\,du + \sup_{x\in[\alpha-2v_0,\,\beta+2v_0]} \left| \int_{v_0}^{v_1} (S_F(x+iv) - S_G(x+iv))\,dv \right|, \quad (2.14)$$

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

## 3 Remarks on Kolmogorov Distance

In general, the behavior of the function $v \to \mathrm{Im}(S_G(x+iv))$ near zero reflects local smoothness properties of the distribution function $G$ at a given point $x$. In particular, it may be related to its concentration function $Q_G(h) = \sup_x (G(x+h) - G(x))$, or to, what is more convenient for our purposes,

$$\delta_G(v_0) = \sup_{|x-y|\ge v_0} \frac{|G(x) - G(y)|}{|x-y|}, \qquad v_0 > 0.$$

To see this, let us represent the Stieltjes transform of $G$ (or its imaginary part) as

$$\mathrm{Im}(S_G(x+iv)) - \pi G(x) = \frac{1}{v} \int_{-\infty}^{+\infty} \frac{2t}{(1+t^2)^2}\,(G(x+vt) - G(x))\,dt,$$

which implies

$$\mathrm{Im}(S_G(x+iv)) \le \pi + \frac{2}{v} \int_0^{+\infty} \frac{2t}{(1+t^2)^2}\,Q_G(vt)\,dt.$$

If $\xi$ denotes a positive random variable with density $p(t) = \frac{2t}{(1+t^2)^2}$, we obtain from the definition that

$$U_G(v_0) \equiv \sup_x \sup_{v\ge v_0} \mathrm{Im}(S_G(x+iv)) \le \pi + 2\,\mathbf{E}\,\xi\,\delta_G(v_0\xi). \quad (3.1)$$

For example, if $G$ has a finite Lipschitz semi-norm $M = \|G\|_{\mathrm{Lip}}$ (equivalently, when $G$ has a density $g$ bounded by the constant $M$), then both $\delta_G$ and $U_G$ are bounded functions in $v_0$. More precisely,

$$\mathrm{Im}(S_G(x+iv)) = v \int_{-\infty}^{+\infty} \frac{g(u)}{|(x-u)+iv|^2}\,du \le M\pi.$$

Note that in this case the Lévy and Kolmogorov distances are equivalent in the sense that $L(F,G) \le \|F-G\| \le (1+M)\,L(F,G)$. Hence, inequality (2.14) of Corollary 2.2 will now take the form

$$c\,\|F-G\| \le v_0 + \int_{-\infty}^{+\infty} |S_F(u+iv_1) - S_G(u+iv_1)|\,du \\ + \sup_{x\in[\alpha-2v_0,\,\beta+2v_0]} \left| \int_{v_0}^{v_1} (S_F(x+iv) - S_G(x+iv))\,dv \right| \tag{3.2}$$

with constant $c$, depending on $M$ only.

We will use this bound in the case where $G$ is the distribution function of the semi-circle law with density $g(x) = G'(x) = \frac{1}{2\pi}\sqrt{4-x^2}\,I_{\{|x|\le 2\}}$.

**Corollary 3.1** *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 $c > 0$,*

$$c\,\|F-G\| \le v_0 + \sup_{x\in[-2,2]} \left| \int_{v_0}^{v_1} (S_F(x+iv) - S_G(x+iv))\,dv \right| \\ + \int_{-\infty}^{+\infty} |S_F(u+iv_1) - S_G(u+iv_1)|\,du.$$

## 4 Empirical Poincaré Inequalities

Assume that the random variables $X_1,\dots,X_n$ have a joint distribution $\mu$ on $\mathbf{R}^n$, satisfying the Poincaré-type inequality (1.1). Given a bounded smooth complex-valued function $f$ on the real line, one may apply (1.1) to

$$g(x_1,\dots,x_n) = \frac{f(x_1)+\cdots+f(x_n)}{n} = \int f\,dF_n, \tag{4.1}$$

where $F_n$ is the empirical measure, defined for "observations" $X_1 = x_1,\dots,X_n = x_n$. Since

$$|\nabla g(x_1,\dots,x_n)|^2 = \frac{|f'(x_1)|^2+\cdots+|f'(x_n)|^2}{n^2} = \frac{1}{n}\int |f'|^2\,dF_n, \tag{4.2}$$

we obtain an integro-differential inequality, which may be viewed as an empirical Poincaré-type inequality for the measure $\mu$:

**Proposition 4.1** *Under* $\mathrm{PI}(\sigma^2)$, *for any smooth $F$-integrable function $f : \mathbf{R} \to \mathbf{C}$, such that $f'$ belongs to $L^2(\mathbf{R}, dF)$,*

$$\mathbf{E}\left| \int f\,dF_n - \int f\,dF \right|^2 \le \frac{\sigma^2}{n}\int |f'|^2\,dF. \tag{4.3}$$

Recall that $F = \mathbf{E}F_n$ denotes the mean of the empirical distribution functions under the measure $\mu$. The inequality holds for all locally Lipschitz functions with the generalized modulus of the derivative $|f'(x)| = \limsup_{y\to x} \frac{|f(x)-f(y)|}{|x-y|}$. As long as $\int |f'|^2\, dF$ is finite, $\int |f|\, dF$ is finite, and (4.3) holds.

Proposition 4.1 may be extended to all $L^p$-spaces by applying the following general property of Poincaré-type inequalities.

**Lemma 4.2** *Under* $\mathrm{PI}(\sigma^2)$, *any Lipschitz function* $g$ *on* $\mathbf{R}^n$ *has a finite exponential moment. More precisely, if* $\int g\, d\mu = 0$ *and* $\|g\|_{\mathrm{Lip}} \le 1$, *then*

$$\mu\{g > t\} \le 3e^{-t/\sigma}, \quad t > 0. \tag{4.4}$$

*Moreover, for any locally Lipschitz* $g$ *on* $\mathbf{R}^n$ *with* $\mu$*-mean zero*,

$$\|g\|_p \le \sigma p\, \|\nabla g\|_p, \quad p \ge 2. \tag{4.5}$$

Here $\|g\|_{\mathrm{Lip}} = \sup_{y\neq x} \frac{|g(x)-g(y)|}{|x-y|}$ with respect to the Euclidean distance in $\mathbf{R}^n$, and the generalized modulus of the gradient is defined by $|\nabla g(x)| = \limsup_{y\to x} \frac{|g(x)-g(y)|}{|x-y|}$.

For a proof of (4.4), which is known as a variant of the well-known Gromov–Milman concentration inequality [13, 22], we refer to [11], see also [1] or [25]. As for (4.5), in a more general scheme of Poincaré-type inequalities with weights, it may be found in [12]. Applying (4.5) to $g = \int f\, dF_n$ as in (4.1), we obtain the following.

**Proposition 4.3** *Under* $\mathrm{PI}(\sigma^2)$, *for any smooth* $F$*-integrable function* $f : \mathbf{R} \to \mathbf{C}$, *for all* $p \ge 2$,

$$\mathbf{E}\left|\int f\, dF_n - \int f\, dF\right|^p \le \frac{(\sigma p)^p}{n^{p/2}}\, \mathbf{E}\left(\int |f'|^2\, dF_n\right)^{p/2}. \tag{4.6}$$

*In particular*,

$$\mathbf{E}\left|\int f\, dF_n - \int f\, dF\right|^p \le \frac{(\sigma p)^p}{n^{p/2}} \int |f'|^p\, dF. \tag{4.7}$$

## 5 Moments of Stieltjes Transforms

The inequality (4.6) may be much sharper than (4.7), especially for $f$'s, which behave like a delta function. This may be seen by looking at the function $f(x) = \frac{1}{x-z}$ with complex parameter $z = u + iv$, where $u \in \mathbf{R}$ is arbitrary and $v > 0$ is small. In that case we deal with the Stieltjes transforms

$$s_n(z) = \int_{-\infty}^{+\infty} \frac{1}{x-z}\, dF_n(x) = \frac{1}{n}\sum_{j=1}^{n} \frac{1}{x_j - z}, \quad s(z) = \int_{-\infty}^{+\infty} \frac{1}{x-z}\, dF(x)$$

of the empirical measures $F_n$ and their mean $F = \mathbf{E} F_n$ with respect to the measure $\mu$ on $\mathbf{R}^n$. As discussed above, the closeness of $s_n$ to $s$ for small $v$ bears important information on the closeness of $F_n$ to $F$ in the sense of the Lévy distance, for example.

Note that the empirical Poincaré-type inequality (4.3) gives

$$\mathbf{E}\,|s_n(z) - s(z)|^2 \le \frac{\sigma^2}{n} \int \frac{1}{|x-z|^4}\,dF(x) \le \frac{\sigma^2}{nv^3}\,\mathrm{Im}(s(z)), \tag{5.1}$$

where we combined the identity

$$\mathrm{Im}(s(z)) = v \int \frac{1}{|x-z|^2}\,dF(x)$$

with the bound $|x - z| \ge v$. Therefore, if $\mathrm{Im}(s(z))$ happens to be bounded for small $v$ (which is the case, e.g., when $F$ has a bounded density), then the variance of $s_n(z)$ can be estimated by a quantity of order $\frac{\sigma^2}{nv^3}$.

In order to get more information on the fluctuations $s_n(z) - s(z)$, a natural step could be to try to extend (5.1) to $L_p$-norms

$$\|s_n(z) - s(z)\|_p = \big(\mathbf{E}\,|s_n(z) - s(z)|^p\big)^{1/p}$$

by proving similar bounds of order $\frac{\sigma}{(nv^3)^{1/2}}$ as in the case $p = 2$. Let, for example, $p = 4$. Appealing to (4.7) would then give

$$\mathbf{E}\,|s_n(z) - s(z)|^4 \le \frac{(4\sigma)^4}{n^2} \int \frac{1}{|x-z|^8}\,dF(x) \le \frac{(4\sigma)^4}{n^2 v^7}\,\mathrm{Im}(s(z)).$$

That is, $\|s_n(z) - s(z)\|_4$ is bounded by a quantity of order $\frac{\sigma}{\sqrt{n}\,v^{7/4}}$, which is worse than what we have in the case $p = 2$. On the other hand, (4.6) gives

$$\mathbf{E}\,|s_n(z) - s(z)|^4 \le \frac{(4\sigma)^4}{n^2}\,\mathbf{E}\bigg(\int \frac{1}{|x-z|^4}\,dF_n(x)\bigg)^2. \tag{5.2}$$

The expectation on the right-hand side may be further estimated by the empirical Poincaré inequality (4.3) with $f(x) = \frac{1}{|x-z|^4}$. Using $|f'(x)| \le \frac{4}{|x-z|^5}$, we get

$$\begin{aligned}\mathbf{E}\bigg(\int \frac{1}{|x-z|^4}\,dF_n(x)\bigg)^2 &\le \bigg(\int \frac{1}{|x-z|^4}\,dF(x)\bigg)^2 + \frac{(4\sigma)^2}{n}\int \frac{1}{|x-z|^{10}}\,dF(x)\\ &\le \bigg(\frac{1}{v^3}\,\mathrm{Im}(s(z))\bigg)^2 + \frac{(4\sigma)^2}{nv^9}\,\mathrm{Im}(s(z)).\end{aligned}$$

Applying this bound in (5.2), we obtain that, up to some numerical constant $c > 0$,

$$c\,\mathbf{E}\,|s_n(z) - s(z)|^4 \le \frac{\sigma^4}{n^2 v^6}\,\mathrm{Im}^2(s(z)) + \frac{\sigma^6}{n^3 v^9}\,\mathrm{Im}(s(z)). \tag{5.3}$$

Therefore, $\|s_n(z) - s(z)\|_4$ is bounded by a similar quantity of order $\frac{\sigma}{\sqrt{n}\,v^{3/2}}$ as in the case $p = 2$, provided that $\frac{\sigma}{\sqrt{n}\,v^{3/2}}$ is of order at most 1.

By a routine combinatorial argument, the inequality (5.3) may be extended to higher powers of 2 by means of the following statement.

**Proposition 5.1** *Under* $\mathrm{PI}(\sigma^2)$, *for any integer* $p = 2^k$ $(k \geq 1)$ *and any* $z = u + iv$, $v > 0$,

$$\frac{p^{-2p}}{(12)^p}\,\mathbf{E}\,|s_n(z) - s(z)|^p \;\leq\; \sum_{j=1}^{k} \left(\frac{\sigma^2}{nv^3}\right)^{(1-\frac{1}{2^j})p} [\mathrm{Im}(s(z))]^{\frac{p}{2^j}}. \tag{5.4}$$

Details of the proof of Proposition 5.1 are given in the Appendix.

On the basis of (5.4), we derive a bound for the moments $\|s_n(z) - s(z)\|_p$ with arbitrary real $p \geq 2$. Put $S = \mathrm{Im}(s(z))$ and $c = \frac{\sigma}{\sqrt{nv^3}}$. If $c \leq 1$, it follows from (5.4) that

$$\frac{p^{-2p}}{(12)^p}\,\mathbf{E}\,|s_n(z) - s(z)|^p \;\leq c^p \sum_{j=1}^{k} S^{\frac{p}{2^j}} \;\leq c^p (1+S)^{p/2}.$$

Hence, $\|s_n(z) - s(z)\|_p \leq 12\,cp^2\sqrt{1+S}$. For the range $2^k < p < 2^{k+1}$, the previous bound, applied to $2^{k+1}$, immediately yields

$$\|s_n(z) - s(z)\|_p \leq 48\,cp^2\sqrt{1+S}, \qquad c = \frac{\sigma}{\sqrt{nv^3}},\;\; S = \mathrm{Im}(s(z)), \tag{5.5}$$

which is thus valid for all $p \geq 1$ (the case $1 \leq p < 2$ may be settled just via (5.1)). Using Chebyshev's inequality and optimizing over $p$, one derives from (5.5) bounds on large deviations, such as

$$\mathbf{P}\big\{|s_n(z) - s(z)| > 96\,c\sqrt{1+S}\,t\big\} \leq 3\,e^{-\sqrt{t}}, \quad t > 0.$$

For further applications, however, it will be more convenient to work with the "vertical" integrals, appearing in Corollary 2.2, that is, integrals of the form

$$J(x) = \int_{v_0}^{v_1} (s_n(x+iv) - s(x+iv))\,dv,$$

where $x \in \mathbf{R}$ and $v_1 > v_0 > 0$ are fixed parameters. Note that the value $S$ in (5.5) may be bounded by the quantity

$$U_F(v_0) \;=\; \sup_{x\in\mathbf{R}}\sup_{v\geq v_0} \mathrm{Im}(s(x+iv)).$$

Hence, if $\frac{\sigma}{\sqrt{nv_0^3}} \leq 1$, we get from (5.5) that the $L^p$-norms $\|J(x)\|_p = (\mathbf{E}\,|J(x)|^p)^{1/p}$, $p \geq 1$, satisfy

$$\|J(x)\|_p \leq \int_{v_0}^{v_1} \|s_n(x+iv) - s(x+iv)\|_p\,dv \leq C\,p^2 \tag{5.6}$$

with $C = 96\,\frac{\sigma}{\sqrt{nv_0}}\,\sqrt{1+U_F(v_0)}$. This bound on the growth of moments allows us to control the maxima

$$J_N = \max\{|J(x_1)|, \dots, |J(x_N)|\}$$

over a large number of points $x_1, \dots, x_N$ on the line. Indeed, for any $p \geq 1$, we can write $J_N^p \leq |J(x_1)|^p + \dots + |J(x_N)|^p$, so that by (5.6),

$$\mathbf{E}\,J_N \,\leq\, \big(\mathbf{E}\,|J(x_1)|^p + \cdots + \mathbf{E}\,|J(x_N)|^p\big)^{1/p} \,\leq\, CN^{1/p}\,p^2.$$

Optimizing over all $p$ (which should be of order $\log N$) and choosing $v_0 = (\frac{\sigma^2}{n})^{1/3}$, so that $\frac{\sigma}{\sqrt{nv_0^3}} = 1$, we arrive at the following.

**Corollary 5.2** *Under* $\mathrm{PI}(\sigma^2)$, *we have for all* $x_1, \dots, x_N \in \mathbf{R}$ *and* $v_1 > v_0 = (\frac{\sigma^2}{n})^{1/3}$, *up to some numerical constant* $C$,

$$\mathbf{E}\,\max_{1\leq k\leq N}\left|\int_{v_0}^{v_1}(s_n(x_k+iv) - s(x_k+iv))\,dv\right| \leq Cv_0\sqrt{1+U_F(v_0)}\,\log^2(1+N).$$

## 6 Concentration of Empirical Distribution Functions in Lévy and Kolmogorov Metrics

We shall now proceed to the final steps in the study of rates of approximation of empirical distribution functions $F_n$ by the mean measure $F = \mathbf{E}F_n$ in terms of the Lévy and Kolmogorov metrics. To this aim, we apply Proposition 2.1 to the pair $(F_n, F)$ in place of $(F, G)$ and then combine it with Corollary 5.2.

Thus, again, let $X = (X_1, \dots, X_n)$ be a random vector in $\mathbf{R}^n$ with joint distribution $\mu$, satisfying the Poincaré-type inequality with constant $\sigma^2$. Note that the measure $\mu$ is restricted by a parameter $A$, defined by

$$|\mathbf{E}X_j - \mathbf{E}X_k| \leq A\sigma, \quad 1 \leq j,\ k \leq n.$$

Under these hypotheses, we prove the following theorem.

**Theorem 6.1** *Let* $v_0 = (\frac{\sigma^2}{n})^{1/3}$. *With some absolute constant C we have*

$$\mathbf{E}\,L(F_n, F) \,\leq\, C\,v_0(1+U_F(v_0))\log^2\left(2+\frac{1+A}{v_0}\right), \tag{6.1}$$

*where* $U_F(v_0) = \sup_{x\in\mathbf{R}}\sup_{v\geq v_0}\mathrm{Im}(s(x+iv))$.

The quantity $U_F(v_0)$ was discussed in Sect. 3, where it was related to uniform smoothness properties of $F$, cf. (3.1). Also, recall that, for $z = u + iv$ ($u \in \mathbf{R}$, $v > 0$),

$$s(z) = \mathbf{E}\,s_n(z) = \int \frac{1}{x-z}\,dF(x)$$

stands for the Stieltjes transform of $F$, and we use $s_n(z) = \int \frac{1}{x-z}\, dF_n(x)$ to denote the Stieltjes transform of empirical distribution functions $F_n$.

If $F$ has a finite Lipschitz semi-norm $M = \|F\|_{\mathrm{Lip}}$, then $\|F_n - F\| \leq (1 + M)L(F_n, F)$ and $U_F(v) \leq M\pi$, for all $v > 0$, so the bound (6.1) may be simplified.

**Corollary 6.2** *If $M = \|F\|_{\mathrm{Lip}}$ is finite,*

$$\mathbf{E}\,\|F_n - F\| \,\leq\, C\,(1+M)\left(\frac{\sigma^2}{n}\right)^{1/3} \log^2\left(2 + \frac{(1+A)\,n}{\sigma^2}\right), \tag{6.2}$$

*where $C$ is an absolute constant.*

Thus, if $\sigma^2$ is of order 1 and $A = O(n^k)$ with some finite $k$, then

$$\mathbf{E}\,\|F_n - F\| \,\leq\, C\,(1+M)\,\frac{\log^2 n}{n^{1/3}}.$$

In fact, in the matrix model (cf. next section), $\sigma^2$ will be of order $1/n$, while $A$ will be of order at most $\sqrt{n}$. Hence, in that case the bound (6.2) would give $E\,\|F_n - F\| \leq C\,\frac{\log^2 n}{n^{2/3}}$.

If we do not know whether $F$ is Lipschitz, it might be reasonable to replace $F$ in (6.2) by a different distribution $G$ at the expense of an additional error term $\|F - G\|$. Indeed, one may apply a general bound

$$U_F(v_0) \leq \frac{\pi\,\|F - G\|}{v_0} + M\pi,$$

where now $M = \|G\|_{\mathrm{Lip}}$. Using this in (6.1) we arrive at

$$\mathbf{E}\,\|F_n - G\| \,\leq\, C\left[\,\|F - G\| + \left(\frac{\sigma^2}{n}\right)^{1/3}\right] \log^2\left(2 + \frac{(1+A)\,n}{\sigma^2}\right)$$

with a constant $C$, depending on $M$ only. This is exactly the inequality of Theorem 1.2, stated in the Introduction.

Let us turn to the proofs. In order to bound the vertical integral in (2.14), we need the following.

**Lemma 6.3** *Under* $\mathrm{PI}(\sigma^2)$, *for all $v > 0$,*

$$\mathbf{E}\int_{-\infty}^{+\infty} |s_n(u+iv) - s(u+iv)|\,du \,\leq\, \frac{\sigma\pi}{\sqrt{n}}\,\big(v^{-1} + Dv^{-2}\big),$$

*where $D$ is the variance of a random variable with distribution $F$.*

*Proof* A similar argument was used in [9], and we include it for completeness. Put $\varphi(u) = \mathbf{E}\,|s_n(u+iv) - s(u+iv)|$. By the empirical Poincaré-type inequality (2.3), applied to the function $f(z)=\frac{1}{x-z}$, cf. (3.1), we have

$$\varphi(u)\le\frac{\sigma}{\sqrt{n}}\left(\int\frac{1}{((x-u)^2+v^2)^2}\,dF(x)\right)^{1/2}.$$

Hence, for any $a$ real, we obtain by Cauchy's inequality that

$$\begin{aligned}\left(\int\varphi(u)\,du\right)^2&\le\int\varphi(u)^2\big((a-u)^2+v^2\big)\,du\int\frac{1}{(a-u)^2+v^2}\,du\\&\le\frac{\sigma^2\pi}{nv}\int\!\!\int\frac{(a-u)^2+v^2}{((x-u)^2+v^2)^2}\,du\,dF(x)\\&=\frac{\sigma^2\pi}{nv}\int\!\!\int\frac{(u^2+v^2)+(x-a)^2}{(u^2+v^2)^2}\,du\,dF(x)\\&=\frac{\sigma^2\pi}{nv}\left(\frac{\pi}{v}+\frac{c}{v^3}\int(x-a)^2\,dF(x)\right),\end{aligned}$$

where $c=\int\frac{1}{(1+t^2)^2}\,dt<\pi$. It remains to choose $a=\int x\,dF(x)$. □

*Proof of Theorem 6.1.* Since $L(F_n,F)\le 1$, we may assume that $v_0$ does not exceed a small numerical constant.

Given $\beta>0$ to be specified later, put $\alpha=-\beta$, $\alpha_0=-\beta_0$, where $\beta_0=\beta+2v_0$. By Proposition 2.1, with some universal constant $c>0$ we have

$$\begin{aligned}c\,L(F_n,F)&\le v_0+v_0\,U_F(v_0)\\&\quad+\sup_{x\in[\alpha_0,\beta_0]}\left|\int_{-\infty}^{x}\operatorname{Im}(s_n(u+iv_0)-s(u+iv_1))\,du\right|,\end{aligned}\tag{6.3}$$

provided that $F(\alpha)\le v_0$ and $1-F(\beta)\le v_0$. Divide the symmetric interval $[\alpha_0,\beta_0]$ into $2N$ subintervals $[x_k,x_{k+1}]$ with endpoints $x_k=\frac{k}{2N}(\beta_0-\alpha_0)$, $k=-N,\dots,N$. Since in general $|S_F(z)|\le\frac{1}{\operatorname{Im}(z)}$, we get

$$\left|\int_{x_k}^{x_{k+1}}\operatorname{Im}(s_n(u+iv_0)-s(u+iv_0))\,du\right|\le\frac{2}{v_0}(x_{k+1}-x_k)=\frac{\beta_0-\alpha_0}{Nv_0}.$$

Hence, up to the summand $\frac{\beta_0-\alpha_0}{Nv_0}$, the supremum in (6.3) may be restricted to the points $x_k$, $k=-N,\dots,N-1$. Applying the inequality (2.13) to each such point with arbitrary $v_1>v_0$, we get

$$\begin{aligned}c\,L(F_n,F)&\le v_0+v_0\,U_F(v_0)+\int_{-\infty}^{+\infty}|s_n(u+iv_1)-s(u+iv_1)|\,du\\&\quad+\frac{\beta_0-\alpha_0}{Nv_0}+\max_{-N\le k\le N-1}\left|\int_{v_0}^{v_1}(s_n(x_k+iv)-s(x_k+iv))\,dv\right|.\end{aligned}$$

Here, by Lemma 6.3 with $v=v_1$, the mean of this horizontal integral is vanishing as $v_1\to+\infty$. The mean of the maximum may be bounded by virtue of Corollary 5.2, which should be used with $2N$ points $x_k$ ($k = -N, \dots, N-1$). We thus obtain for the expectation (with some other numerical constant)

$$c\,\mathbf{E}L(F_n, F) \le v_0 + v_0 U_F(v_0) + \frac{\beta_0 - \alpha_0}{N v_0} + v_0 \sqrt{1 + U_F(v_0)}\,\log^2(2N+1),$$

which can be simplified as

$$c\,\mathbf{E}L(F_n, F) \le v_0(1 + U_F(v_0))\bigg(1 + \frac{\beta_0 - \alpha_0}{N v_0^2} + \log^2(N)\bigg).$$

To optimize over $N$, we do not lose much by taking $N = 1 + [B]$ with $B = (\beta_0 - \alpha_0)/v_0^3$. Using, say, $v_0 \le 1$, we then get

$$c\,\mathbf{E}L(F_n, F) \le v_0(1 + U_F(v_0))\bigg[2 + \log^2\bigg(1 + \frac{\beta_0 - \alpha_0}{v_0^3}\bigg)\bigg]. \tag{6.4}$$

It remains to choose $\beta$ to satisfy the requirements $F(\alpha) \le v_0$ and $1 - F(\beta) \le v_0$.

Without loss of generality we may assume $-A\sigma \le \mathbf{E}X_j \le A\sigma$, for all $j \le n$. By Lemma 4.2 with $g(x) = x_i$, cf. (4.4), for all $h > 0$,

$$\mathbf{P}\{X_i - \mathbf{E}X_i \ge \sigma h\} \le 3e^{-h}, \quad \mathbf{P}\{X_i - \mathbf{E}X_i \le -\sigma h\} \le 3e^{-h}.$$

Therefore, $\mathbf{P}\{X_i \ge h\} \le 3e^{-(h-A)}$ and $\mathbf{P}\{X_i \le -h\} \le 3e^{-(h-A)}$, whenever $h \ge A$. Averaging over all $i$'s, we obtain similar bounds for the mean distribution function $F = \mathbf{E}F_n$, that is,

$$1 - F(h) \le 3e^{-(h-A)}, \quad F(-h) \le 3e^{-(h-A)} \quad (h > A).$$

Hence, one may take $\alpha = -\beta$ with $\beta = A + \log\frac{3}{v_0}$, and then $\beta_0 = A + 2v_0 + \log\frac{3}{v_0}$. Since $v_0$ does not exceed a small numerical constant, the expression under the log sign may be bounded, up to a (universal) factor, by $1 + \frac{1+A}{v_0^4}$. Therefore, the expression in the square brackets in (6.4) does not exceed, up to a factor, $\log^2(2 + \frac{1+A}{v_0})$.

Thus, Theorem 6.1 is proved. □

## 7 High-Dimensional Random Matrices

We can now apply the bound, obtained in Theorem 1.2, to spectral empirical distributions. Let $\{\xi_{jk}\}_{1 \le j \le k \le n}$ be a family of independent random variables on some probability space with mean $\mathbf{E}\xi_{jk} = 0$ and variance $\mathrm{Var}(\xi_{jk}) = 1$. Put $\xi_{jk} = \xi_{kj}$, for $k < j$. Then we have a symmetric $n \times n$ random matrix $\mathbf{W}$ with entries

$$W_{jk} = \frac{1}{\sqrt{n}}\xi_{j,k}, \quad 1 \le j,\ k \le n.$$

Arrange its (real random) eigenvalues in the increasing order: $X_1 \le \cdots \le X_n$. With particular values $X_1 = x_1, \dots, X_n = x_n$ one associates an empirical (spectral) distribution function $F_n$ with mean (expected) measure $F = \mathbf{E}F_n$.

The joint distribution $P$ of the collection $\{\xi_{jk}\}_{1\le j\le k\le n}$ represents a product probability measure on the Euclidean space $\mathbf{R}^N$ of dimension $N=n(n+1)/2$, while the joint distribution $\mu$ of the spectral values $X_j$ is a probability measure on $\mathbf{R}^n$, obtained from $P$ as the image under the map $T$ from symmetric matrices to their eigenvalues. We will apply the following elementary lemma.

**Lemma 7.1** *Let $\mu_1,\dots,\mu_N$ be probability measures on $\mathbf{R}$, satisfying* $\mathrm{PI}(\sigma^2)$. *The image $\mu$ of the product measure $P=\mu_1\otimes\cdots\otimes\mu_N$ under any Lipschitz map $T:\mathbf{R}^N\to\mathbf{R}^n$ satisfies* $\mathrm{PI}(\sigma^2\|T\|_{\mathrm{Lip}}^2)$.

In our specific matrix situation, the map $T:\mathbf{R}^{n(n+1)/2}\to\mathbf{R}^n$ has Lipschitz semi-norm $\|T\|_{\mathrm{Lip}}=\frac{\sqrt{2}}{\sqrt{n}}$. To be more precise, by Hoffman-Wielandt's theorem (cf., e.g., [4], p. 165), the Hilbert–Schmidt norm satisfies

$$\sum_{i=1}^n |X_i-X_i'|^2\le \|\mathbf{W}-\mathbf{W}'\|_{HS}^2=\frac{1}{n}\sum_{j,k=1}^n|\xi_{jk}-\xi_{jk}'|^2\le\frac{2}{n}\sum_{1\le j\le k\le n}|\xi_{jk}-\xi_{jk}'|^2,$$

for any collections $\{\xi_{jk}\}_{j\le k}$ and $\{\xi_{jk}\}_{j\le k}'$ with eigenvalues $(X_1,\dots,X_n)$, $(X_1',\dots,X_n')$, respectively. This is a well-known fact, used in concentration problems, cf., e.g., [8, 15, 23, 26].

Therefore, if the distributions of $\xi_{jk}$'s satisfy a one-dimensional Poincaré-type inequality with common constant $\sigma^2$, then $\mu$ satisfies a Poincaré-type inequality on $\mathbf{R}^n$ with constant $\sigma_n^2=\frac{2\sigma^2}{n}$ (which is asymptotically much better). Note that necessarily $\sigma^2\ge\mathbf{E}\xi_{j,k}^2=1$. In addition, since $\max_j|\mathbf{E}X_j|$ is of order $\sigma$, the other involved parameter of the measure $A_n=\frac{1}{\sigma_n}\max_{j,k}|\mathbf{E}X_j-\mathbf{E}X_k|$ is of order at most $\sqrt{n}$. Hence, Theorem 1.2 yields the following for our matrix model.

**Corollary 7.2** *If the distributions of $\xi_{jk}$ satisfy* $\mathrm{PI}(\sigma^2)$ *on the real line, then for any distribution function $G$ with finite Lipschitz semi-norm $M=\|G\|_{\mathrm{Lip}}$,*

$$\mathbf{E}\,\|F_n-G\|\le C\left[\|F-G\|+\left(\frac{\sigma}{n}\right)^{2/3}\right]\log^2(n+1),\tag{7.1}$$

*where the constant $C$ depends on $M$, only.*

The statement may be applied to $G=F$ itself, but it is more natural to consider for $G$ the standard semi-circle law. The problem is then to estimate $\|F-G\|$.

Introduce the resolvent function of the matrix $\mathbf{W}$,

$$\mathbf{R}(z)=(\mathbf{W}-z\mathbf{I})^{-1},\quad z=u+iv,\ u\in\mathbf{R},\ v>0,$$

where $\mathbf{I}$ denotes the identity matrix (of size $n\times n$), and the Stieltjes transforms

$$s_n(z)=\frac{1}{n}\operatorname{Tr}\mathbf{R}(z)=\int\frac{1}{x-z}\,dF_n(x),$$

$$s(z) = \mathbf{E}\, s_n(z) = \int \frac{1}{x-z}\, dF(x).$$

Since under the PI($\sigma^2$)-hypothesis, imposed on the distributions of the random variables $\xi_{jk}$, the joint distribution of the eigenvalues of $\mathbf{W}$ satisfies PI($\sigma_n^2$) on $\mathbf{R}^n$ with $\sigma_n^2 = \frac{2\sigma^2}{n}$, we may apply the moment estimates for $\|s_n - s\|_p$ developed in Sect. 5. In particular, inequalities (5.1) and (5.3) for the even moments $p = 2$ and $p = 4$ take the form

$$\mathbf{E}\,|s_n(z) - s(z)|^2 \le \frac{C_1}{n^2 v^3}\, \mathrm{Im}(s(z)), \tag{7.2}$$

$$\mathbf{E}\,|s_n(z) - s(z)|^4 \le \frac{C_2}{n^4 v^6}\, \mathrm{Im}^2(s(z)) + \frac{C_3}{n^6 v^9}\, \mathrm{Im}(s(z)) \tag{7.3}$$

with constants $C_i$'s depending on $\sigma$, only. As it turns out, these are the concentration inequalities, which control the distance from $F$ to the semi-circle law $G$.

**Theorem 7.3** *Assume that the random variables $\xi_{jk}$ are independent with $\mathbf{E}\,\xi_{jk} = 0$, $\mathbf{E}\,|\xi_{jk}|^2 = 1$, $\mathbf{E}\,|\xi_{jk}|^4 \le M_4 < +\infty$, and that they satisfy the conditions* (7.2)–(7.3). *Then*

$$\Delta_n = \|F - G\| \le C n^{-2/3},$$

*where $G$ denotes the distribution function of the standard semi-circle law, and where the constant $C$ depends on $C_i$'s and $M_4$, only.*

Combined with Corollary 7.2, this assertion implies the main Theorem 1.1. The proof of Theorem 7.3 requires some preparations and is given in the next section.

*Remark 7.4* It is well known that Poincaré-type inequalities on the real line may be reduced to Hardy-type inequalities with weights. Necessary and sufficient conditions for a measure on the positive half-axis to satisfy a Hardy-type inequality with general weights were found by M.G. Kac and I.S. Krein [24]. We refer the interested reader to [28] and [27] for a full characterization, and here just recall a principal result (see also [7] for a different approach).

Let $\mu$ be a probability measure on the line with median $m$, that is, $\mu(-\infty, m) \le \frac{1}{2}$ and $\mu(m, +\infty) \le \frac{1}{2}$. Define

$$A_0(\mu) = \sup_{x<m} \left[ \mu(-\infty, x) \int_{-\infty}^{x} \frac{dt}{p_\mu(t)} \right],$$

$$A_1(\mu) = \sup_{x>m} \left[ \mu(x, +\infty) \int_{x}^{+\infty} \frac{dt}{p_\mu(t)} \right],$$

where $p_\mu$ denotes the density of the absolutely continuous component of $\mu$ with respect to the Lebesgue measure, and where we set $A_0 = 0$, respectively $A_1 = 0$, if $\mu(-\infty, m) = 0$ or $\mu(m, +\infty) = 0$. Then $\mu$ satisfies PI($\sigma^2$) with some finite constant if and only if both $A_0(\mu)$ and $A_1(\mu)$ are finite. Moreover, the optimal value of $\sigma^2$ satisfies

$$c_0(A_0(\mu)+A_1(\mu)) \le \sigma^2 \le c_1(A_0(\mu)+A_1(\mu)),$$

where $c_0$ and $c_1$ are positive universal constants.

Necessarily, $\mu$ must have a non-trivial absolutely continuous part with a density which is positive almost everywhere on the supporting interval. To roughly describe the whole picture in the case where $\mu$ is absolutely continuous and has a positive, continuous well-behaving density, one may note that the Poincaré constant is finite, as long as the measure has a finite exponential moment. In particular, any probability measure with a logarithmically concave density satisfies PI($\sigma^2$) with a finite $\sigma$, cf. [5].

Let us also note that the measure $\mu$ may have a non-trivial discrete component in order to satisfy PI($\sigma^2$). However, purely discrete measures do not satisfy PI($\sigma^2$). For example, Theorem 1.1 may not be applied to the Bernoulli measure (with two atoms).

*Remark 7.5* It should be clear that an attempt to obtain a good estimate for $\Delta_n^*$ includes in particular a certain concentration result on empirical measures. One can argue whether the PI($\sigma^2$)-hypothesis is artificial or close to being necessary in statements on the rates of $\Delta_n^*$ such as in Theorem 1.1. In this context let us mention the concentration result concerning the functionals $M_n = \max\{\xi_1,\dots,\xi_n\}$ and $m_n = \min\{\xi_1,\dots,\xi_n\}$, generated by a sequence of i.i.d. random variables with given distribution $\mu$ on the line. It turns out (cf. [6, 10]) that the variances $\mathrm{Var}(M_n)$ and $\mathrm{Var}(m_n)$ are bounded for the growing dimension $n$ if and only if the measure $\mu$ satisfies a Poincaré-type inequality in the class of convex functions. Already this simple example suggests that some kinds of Poincaré-type inequalities may appear naturally also in the context of spectral empirical measures.

## 8 Proof of Theorem 7.3

Keeping the previous notation, we start with some auxiliary lemmas. Denote by $t(z) = \int \frac{1}{x-z}\,dG(x)$ the Stieltjes transform of the standard semi-circle law.

**Lemma 8.1** *Write for $z = u + iv$, $v > 0$,*

$$s(z)^2 + zs(z) + 1 = \delta_n(z) \tag{8.1}$$

*with some function $\delta_n(z)$. Then, with some absolute constant $C$, for all $|u| \le 2$,*

$$|s(z) - t(z)| \le \frac{C\,|\delta_n(z)|}{\max\{\mathrm{Im}(s(z)),\ v^{1/4}\,\mathrm{Im}^{\frac{1}{2}}(s(z)),\ \sqrt{v}\,\}}.$$

*Proof* Applying the well-known identity $t(z)^2 + zt(z) + 1 = 0$, we may write

$$|t(z) - s(z)|\,|t(z) + s(z) + z| = |\delta_n(z)|.$$

Note that $|t(z)+s(z)+z| \geq \mathrm{Im}(s(z))$ and, on the other hand,

$$|t(z)+s(z)+z| \geq \mathrm{Im}\{z+t(z)\} \geq \frac{\mathrm{Im}\{\sqrt{z^2-4}\}}{2}.$$

Since $\mathrm{Re}\{z^2-4\}<0$ in $|u|\leq 2$, we have $\mathrm{Im}\{\sqrt{z^2-4}\} \geq \frac{1}{\sqrt{2}}|\sqrt{z^2-4}|$. But

$$\left|\sqrt{z^2-4}\right| = \max\left\{\sqrt{|z-2|}, \sqrt{|z+2|}\right\} \min\left\{\sqrt{|z-2|}, \sqrt{|z+2|}\right\},$$

so $|\sqrt{z^2-4}| \geq \max\{\sqrt{2-u}, \sqrt{2+u}, \sqrt{v}\}$. This completes the proof of the lemma. □

The next step is to obtain a representation like (8.1) and derive a bound on $\delta_n(z)$. Given an $n\times n$ matrix $\mathbf{A}$, we denote by $\mathbf{A}^{(j)}$ the principal submatrix of order $n-1$, i.e., $\mathbf{A}^{(j)}$ is obtained from $\mathbf{A}$ by deleting the $j$-th row and the $j$-th column. Let $\sum^{(j)}$ denote the sum over all indices from $\{1,\dots,n\}\setminus\{j\}$. Let

$$\mathbf{R}^{(j)}(z) = \left(\mathbf{W}^{(j)} - z\mathbf{I}\right)^{-1},$$

where $\mathbf{I}$ denotes (with subindex or without) the identity matrix of a corresponding size.

In the next statements we assume the conditions of Theorem 7.3 are satisfied and that all constants may depend on the values of $C_k$ $(k=1,2,3)$ and $M_4$.

**Lemma 8.2** *There exists a constant $C$ such that*

$$s(z) = -\frac{1}{z+s(z)} + \frac{\delta_n(z)}{z+s(z)}, \tag{8.2}$$

*where*

$$|\delta_n(z)| \leq \frac{C\,\mathrm{Im}^{\frac{1}{2}}(s(z))}{nv^{\frac{3}{2}}} + \frac{C\,\mathrm{Im}(s(z))}{n^2v^3} + \frac{C\,\mathrm{Im}^{\frac{1}{2}}(s(z))}{n^{\frac{3}{2}}v^{\frac{3}{2}}\,|z+s(z)|} + \frac{C\,\mathrm{Im}(s(z))}{n^{\frac{3}{2}}v^2\,|z+s(z)|}. \tag{8.3}$$

*Proof* We use a well-known representation (cf. (4.6) in [19], p. 236)

$$R_{jj}(z) = -\frac{1}{z+s(z)} + \frac{1}{z+s(z)}\,(\varepsilon_{j1}+\cdots+\varepsilon_{j4})\,R_{jj}(z), \tag{8.4}$$

where

$$\varepsilon_{j1} = \frac{1}{\sqrt{n}}\,\xi_{jj}, \quad \varepsilon_{j2} = -\frac{1}{n}\left(\mathrm{Tr}\mathbf{R} - \mathrm{Tr}\mathbf{R}^{(j)}\right),$$

$$\varepsilon_{j3} = \frac{1}{n}\sum\nolimits_{l,k}^{(j)} (\xi_{jk}\xi_{jl} - \delta_{kl})R_{kl}^{(j)},$$

$$\varepsilon_{j4} \equiv \Delta(\mathbf{R}) = \frac{1}{n}\,(\mathrm{Tr}\,\mathbf{R} - \mathbf{E}\,\mathrm{Tr}\,\mathbf{R}).$$

Put $\delta_n(z) = \delta_{n1} + \cdots + \delta_{n4}(z)$, where

$$\delta_{n\nu}(z) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,\varepsilon_{j\nu} R_{jj}(z).$$

By (8.4), $\delta_{n1}(z) = \delta_{n1}^{(1)}(z) + \cdots + \delta_{n1}^{(4)}(z)$ with

$$\delta_{n1}^{(\nu)}(z) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,\varepsilon_{j1}\varepsilon_{j\nu} R_{jj}(z).$$

Applying Cauchy's inequality and a simple bound $\frac{1}{n}\sum_{j=1}^{n} |R_{jj}(z)|^2 \le \frac{1}{v}\,\mathrm{Im}(s_n(z))$, we get

$$\left|\delta_{n1}^{(1)}(z)\right| \le \left(\frac{1}{n^3}\sum_{j=1}^{n} \mathbf{E}\,|\xi_{jj}|^4\right)^{\frac{1}{2}} \left(\frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|R_{jj}(z)|^2\right)^{\frac{1}{2}} \le \frac{C\,\mathrm{Im}^{\frac{1}{2}}(s(z))}{n\sqrt{v}}. \tag{8.5}$$

Using

$$\left|\mathrm{Tr}\,\mathbf{R} - \mathrm{Tr}\,\mathbf{R}^{(j)}\right| \le v^{-1}, \tag{8.6}$$

we also have $|\delta_{n1}^{(2)}(z)| \le \frac{C\,\mathrm{Im}^{\frac{1}{2}}(s(z))}{n^{\frac{3}{2}} v^{\frac{3}{2}}}$. Similarly,

$$\left|\delta_{n1}^{(3)}(z)\right| \le \left(\frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|\varepsilon_{j1}|^2|\varepsilon_{j3}|^2\right)^{\frac{1}{2}} \left(\frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|R_{jj}(z)|^2\right)^{\frac{1}{2}} \le \frac{C\,\mathrm{Im}(s(z))}{nv}. \tag{8.7}$$

Simple calculations show that $|\delta_{n1}^{(4)}(z)| \le \frac{C\,\mathrm{Im}(s(z))}{n^{\frac{3}{2}} v^2}$. By (8.6), $|\delta_{n2}(z)| \le \frac{1}{nv^{\frac{3}{2}}}\,\mathrm{Im}^{\frac{1}{2}}(s(z))$. Furthermore, using the representation

$$R_{jj} = \frac{1}{-z + \frac{1}{\sqrt{n}}\xi_{jj} - \mathrm{Tr}\,\mathbf{R}^{(j)}}\,\big(1 + \varepsilon_{j3}R_{jj}(z)\big),$$

we get

$$|\delta_{n3}(z)| \le \frac{1}{n}\sum_{j=1}^{n}\left|\mathbf{E}\,\frac{\varepsilon_{j3}^2}{-z + \frac{1}{\sqrt{n}}\xi_{jj} - \mathrm{Tr}\,\mathbf{R}^{(j)}}\,R_{jj}(z)\right|.$$

Once more, applying Hölder's inequality, we obtain

$$\begin{aligned}|\delta_{n3}(z)| &\le \left(\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,\frac{\frac{1}{n^4}(\mathrm{Tr}\,|\mathbf{R}^{(j)}|^2)^2}{|-z + \frac{1}{\sqrt{n}}\xi_{jj} - \frac{1}{n}\mathrm{Tr}\,\mathbf{R}^{(j)}|^2}\right)^{\frac{1}{2}} \left(\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|R_{jj}(z)|^2\right)^{\frac{1}{2}} \\ &\le \frac{C\,\mathrm{Im}^{\frac{1}{2}}(s(z))}{nv^{\frac{3}{2}}}.\end{aligned} \tag{8.8}$$

By the assumptions (7.2)–(7.3), we have $|\delta_{n4}(z)| \le \frac{C\,\mathrm{Im}(s(z))}{n^2 v^3}$. This completes the proof. $\square$

Put $v_0 = \max\{Cn^{-\frac{2}{3}}, \gamma\Delta_n\}$ for sufficiently small $\gamma > 0$.

**Corollary 8.3** *There exist positive constants $C_1$ and $C_2$ such that for any $v \ge v_0$, we have $|s(z)| \le C_1$ and*

$$|z + s(z)| \ge C_2. \tag{8.9}$$

*Proof* As before, let $F(x) = \mathbf{E}F_n(x)$. Since $s(z) - t(z) = \int_{-\infty}^{\infty} \frac{F(x)-G(x)}{(x-z)^2}\,dx$, we have $|s(z) - t(z)| \le \frac{\pi\Delta_n}{v} \le \frac{\pi\Delta_n}{v_0} \le \frac{\pi}{\gamma}$, so

$$|s(z)| \le 1 + \frac{\pi}{\gamma}. \tag{8.10}$$

Relations (8.2), (8.3), and inequality (8.10) together imply (8.9). $\square$

Now, define the quantity $\Delta(\mathbf{R}) = \frac{1}{n}|\mathrm{Tr}\,\mathbf{R} - \mathbf{E}\,\mathrm{Tr}\,\mathbf{R}|$ and introduce the event $\mathcal{A} = \{\Delta(\mathbf{R}) \le \eta\,|z + s(z)|\}$, where $\eta$ is a sufficiently small positive absolute constant. We use $I_{\{\mathcal{A}\}}$ to denote the corresponding indicator random variable.

**Lemma 8.4** *There exists a constant $C$, such that for any $v \ge v_0$,*

$$\frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|R_{jj}(z)|^2\, I_{\{\mathcal{A}\}} \le C.$$

*Proof* Let $B = \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|R_{jj}(z)|^2\, I_{\{\mathcal{A}\}}$. The representation (8.4) yields

$$\begin{aligned} B \le \frac{C}{|z+s(z)|^2} \Bigg(1 &+ \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|\varepsilon_{j1}|^2\,|R_{jj}(z)|^2\, I_{\{\mathcal{A}\}} \\ &+ \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|\varepsilon_{j2}|^2\,|R_{jj}(z)|^2\, I_{\{\mathcal{A}\}} + \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|\varepsilon_{j3}|^2|\,R_{jj}(z)|^2\, I_{\{\mathcal{A}\}}\Bigg) \\ &+ \frac{C}{|z+s(z)|^2}\,\mathbf{E}\,|\Delta(\mathbf{R})|^2 \left(\frac{1}{n}\sum_{j=1}^{n} |R_{jj}(z)|^2\right) I_{\{\mathcal{A}\}}. \end{aligned}$$

From this,

$$\begin{aligned} B \le \frac{C_1}{|z+s(z)|^2} \Bigg(1 &+ \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|\varepsilon_{j1}|^2\,|R_{jj}(z)|^2\, I_{\{\mathcal{A}\}} \\ &+ \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|\varepsilon_{j2}|^2\,|R_{jj}(z)|^2\, I_{\{\mathcal{A}\}} + \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\,|\varepsilon_{j3}|^2\,|R_{jj}(z)|^2\, I_{\{\mathcal{A}\}}\Bigg). \end{aligned} \tag{8.11}$$

Applying Cauchy's inequality, we get

$$\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_{j1}|^2\,|R_{jj}(z)|^2\,I_{\{\mathcal{A}\}} \le \frac{C_1\max_j \mathbf{E}^{\frac{1}{2}}|\xi_{jj}|^4}{nv}\left(\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|R_{jj}(z)|^2\,I_{\{\mathcal{A}\}}\right)^{\frac{1}{2}}$$
$$= \frac{C_1\max_j \mathbf{E}^{\frac{1}{2}}|\xi_{jj}|^4}{nv}\,B^{\frac{1}{2}}. \tag{8.12}$$

By the inequality (8.6) and Corollary 8.3, we get

$$\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_{j2}|^2\,|R_{jj}(z)|^2\,I_{\{\mathcal{A}\}} \;\le\; \frac{C_1}{n^2v^3}\,B.$$

Furthermore, using an analogue of Rosenthal's inequality for quadratic forms (considered in [19], p. 274), for all $p\ge 1$, we have with some constants $C=C(p)$, depending on $p$,

$$\mathbf{E}\,|\varepsilon_{j3}|^{2p} \le \frac{C}{n^p}\,\mathbf{E}\left(\frac{1}{n}\operatorname{Tr}\left|\mathbf{R}^{(j)}\right|^2\right)^p \le \frac{C}{n^pv^p}\,\mathbf{E}\left(\operatorname{Im}\left\{\frac{1}{n}\operatorname{Tr}R^{(j)}\right\}\right)^p$$
$$\le \frac{C}{n^pv^p}\left(\operatorname{Im}\left\{\mathbf{E}\,\frac{1}{n}\operatorname{Tr}R^{(j)}\right\}\right)^p + \frac{C}{n^pv^p}\left(\mathbf{E}\,|\Delta(\mathbf{R})|^p + \frac{1}{n^pv^p}\right). \tag{8.13}$$

This inequality and the conditions (7.2)–(7.3) together imply that, for $p=1,\dots,4$,

$$\mathbf{E}\,|\varepsilon_{j3}|^{2p} \le \frac{C_p\,|s(z)|^p}{n^p\,v^p} + \frac{C}{n^{2p}\,v^{5p/2}}.$$

Define $\xi^2=\frac{1}{n}\sum_{j=1}^{n}|\varepsilon_{j3}|^2\,|R_{jj}(z)|^2\,I_{\{\mathcal{A}\}}$. The representation (8.4) implies that

$$\mathbf{E}\xi^2 \le \frac{C}{|z+s(z)|^2}\Bigg(\mathbf{E}\,|\varepsilon_{j3}|^2 + \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_{j3}|^2|\varepsilon_{j1}|^2\,|R_{jj}(z)|^2\,I_{\{\mathcal{A}\}}$$
$$+\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_{j3}|^2|\varepsilon_{j2}|^2\,|R_{jj}(z)|^2I_{\{\mathcal{A}\}} + \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_{j3}|^4\,|R_{jj}(z)|^2\,I_{\{\mathcal{A}\}}\Bigg)$$
$$+\frac{C}{|z+s(z)|^2}\,\mathbf{E}\,|\Delta(\mathbf{R})|^2\,\xi^2.$$

This inequality implies

$$\mathbf{E}\,\xi^2 \le \frac{C}{|z+s(z)|^2}\Bigg(\mathbf{E}\,|\varepsilon_{j3}|^2 + \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_{j1}|^4\,|R_{jj}(z)|^2\,I_{\{\mathcal{A}\}}$$
$$+\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_{j2}|^4\,|R_{jj}(z)|^2\,I_{\{\mathcal{A}\}} + \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_{j3}|^4\,|R_{jj}(z)|^2\,I_{\{\mathcal{A}\}}\Bigg).$$

Using (8.13) with $p=4$ and Hölder's inequality, after simple calculations we get

$$\mathbf{E}\xi^2 \le \frac{C}{|z+s(z)|^2}\left(\frac{\operatorname{Im}(s(z))}{nv}+\frac{1}{n^2v}+\frac{1}{n^4v^6}+\frac{(\operatorname{Im}(s(z)))^2}{n^2v^3}\right)B^{\frac{1}{2}}. \tag{8.14}$$

Inequalities (8.11), (8.12), and (8.14) together imply that there exists a constant $c>0$, such that for $v \ge v_0 = \max\{cn^{-\frac{2}{3}}, \gamma\Delta_n\}$,

$$B \le \frac{\alpha}{|z+s(z)|}B^{\frac{1}{2}}+\frac{C}{|z+s(z)|^2}$$

for sufficiently small absolute $\alpha>0$. This inequality completes the proof of Lemma 8.4. □

Now we may obtain an improved bound for $\delta_n(z)$.

**Lemma 8.5** *There exist positive constants $C_1$ and $C_2$, such that for any $v \ge v_0$,*

$$|\delta_n(z)| \le \frac{C_1 \operatorname{Im}(s(z))}{n^2v^3}+\frac{C_2}{nv}.$$

*Proof* Introduce a new quantity

$$\widetilde{\delta}_{nv}(z) = \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\varepsilon_{jv}R_{jj}(z)I_{\{\mathcal{A}\}}.$$

It is straightforward to check that

$$\left|s(z)-\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}R_{jj}(z)I_{\{\mathcal{A}\}}\right| \le \frac{1+|s(z)|}{\gamma\,|s(z)+z|}\,\mathbf{E}\left|\frac{1}{n}(\operatorname{Tr}\mathbf{R}(z)-\mathbf{E}\operatorname{Tr}\mathbf{R}(z))\right|^2$$
$$\le \frac{C\operatorname{Im}(s(z))}{n^2v^3}.$$

This inequality and equality (8.4) imply that

$$s(z) = -\frac{1}{z+s(z)}+\frac{C\theta\operatorname{Im}(s(z))}{n^2v^3}+\widetilde{\delta}_n(z),$$

where $\widetilde{\delta}_n(z)=\widetilde{\delta}_{n1}(z)+\cdots+\widetilde{\delta}_{n4}(z)$. Define

$$\widetilde{\delta}_{n1}^{(\nu)}(z) = \frac{1}{n(z+s(z))}\sum_{j=1}^{n}\mathbf{E}\,\varepsilon_{j1}\varepsilon_{j\nu}\,R_{jj}(z)I_{\{\mathcal{A}\}}.$$

Note that, by inequality (8.5) and Lemma 8.5,

$$\left|\widetilde{\delta}_{n1}^{(1)}(z)\right| \le \frac{C}{n}. \tag{8.15}$$

Applying Lemma 8.4 and Cauchy's inequality, we obtain

$$\big|\widetilde{\delta}_{n1}^{(2)}(z)\big| \le \frac{C}{nv}. \tag{8.16}$$

From inequality (8.7) and Lemma 8.4 it follows that

$$\big|\widetilde{\delta}_{n1}^{(3)}(z)\big| \le \frac{C\,\mathrm{Im}^{\frac{1}{2}}(s(z))}{n\sqrt{v}} \quad \text{and} \quad |\widetilde{\delta}_{n1}^{(4)}(z)| \le \frac{C\,\mathrm{Im}^{\frac{1}{2}}(s(z))}{n^{\frac{3}{2}}v^{\frac{3}{2}}}. \tag{8.17}$$

Furthermore,

$$\big|\widetilde{\delta}_{n2}(z)\big| \le \frac{1}{n^2v}\sum_{j=1}^{n}\mathbf{E}\,|R_{jj}(z)|\,I_{\{\mathcal{A}\}} \le \frac{1}{nv}\left(\sum_{j=1}^{n}\mathbf{E}\,|R_{jj}(z)|^2\,I_{\{\mathcal{A}\}}\right)^{\frac{1}{2}} \le \frac{C}{nv}. \tag{8.18}$$

Inequality (8.8) and Lemma 8.4 imply $|\widetilde{\delta}_{n3}(z)| \le \frac{C}{nv}$, while by the conditions (7.2)–(7.3), we have

$$\big|\widetilde{\delta}_{n4}\big| \le |\delta_{n4}(z)| \le \frac{C\,\mathrm{Im}(s(z))}{n^2v^3}. \tag{8.19}$$

It remains to combine the inequalities (8.15)–(8.19). The proof of the lemma is complete. □

*Proof of Theorem 7.3* We put $v_1 = 1$ and apply Corollary 3.1. As shown in [19] (cf. the inequality (4.30) therein), there exists a constant $C > 0$ such that

$$\int_{-\infty}^{\infty} |s(u+iv_1) - t(u+iv_1)|\,du \le \frac{C}{n}.$$

Hence, we need to bound the "vertical" integral in Corollary 3.1 only. According to Lemma 8.1,

$$|s(z) - t(z)| \le |\delta_n(z)| \min\left\{\frac{1}{\mathrm{Im}(s(z))}, v^{\frac{1}{2}}\right\}.$$

By Lemma 8.5,

$$|\delta_n(z)| \le \frac{C\,\mathrm{Im}(s(z))}{n^2v^3} + \big|\widetilde{\delta}_n(z)\big| \le \frac{C\,\mathrm{Im}(s(z))}{n^2v^3} + \frac{C}{nv}.$$

These inequalities together imply that, for all $v \ge v_0$, we have $|s(z) - t(z)| \le \frac{C}{nv^{3/2}}$. After integration we get

$$\int_{v_0}^{v_1} |s(z) - t(z)|\,dv \le \frac{C}{n\sqrt{v_0}} \le Cn^{-\frac{2}{3}}.$$

This completes the proof of Theorem 7.3 and thus of Theorem 1.1. □

## Appendix: Proof of Proposition 5.1

By Proposition 4.3, for any smooth $F$-integrable function $f:\mathbf{R}\to\mathbf{C}$, for all $p\geq 2$,

$$\mathbf{E}\left|\int f\,dF_n\right|^p \leq 2^p\left|\int f\,dF\right|^p + 2^p\,(cp)^p\,\mathbf{E}\left(\int |f'|^2\,dF_n\right)^{p/2}. \tag{9.1}$$

Here and below we set $c=\sigma/\sqrt{n}$. The same inequality may be applied to the function $|f'|^2$ with power $p/2$ (when $p\geq 4$), and the argument can be repeated.

More generally, consider the case $p=2^k$ ($k\geq 1$ integer) and, starting with $f_0=f$, assume we have a collection of non-negative smooth functions $f_j$, $1\leq j\leq k$, such that

$$f_1\geq |f_0'|^2,\quad f_2\geq |f_1'|^2,\ \dots,\ f_k\geq |f_{k-1}'|^2. \tag{9.2}$$

Then by the repeated use of (9.1),

$$\begin{aligned}\mathbf{E}\left|\int f\,dF_n\right|^p &\leq 2^p\left|\int f_0\,dF\right|^p + 2^p\,(cp)^p\,\mathbf{E}\left(\int f_1\,dF_n\right)^{\frac{p}{2}}\\ &\leq 2^p\left|\int f_0\,dF\right|^p + 2^{p+\frac{p}{2}}\,(cp)^p\left(\int f_1\,dF\right)^{\frac{p}{2}}\\ &\quad + 2^{p+\frac{p}{2}}\,(cp)^p\left(\frac{cp}{2}\right)^{\frac{p}{2}}\mathbf{E}\left(\int f_2\,dF_n\right)^{\frac{p}{4}}.\end{aligned}$$

At the end of this iteration we will arrive at the bound

$$\begin{aligned}\mathbf{E}\left|\int f\,dF_n\right|^p &\leq 2^p\left|\int f\,dF\right|^p\\ &\quad+\sum_{j=1}^{k-1}2^{p+\frac{p}{2}+\cdots+\frac{p}{2^j}}\,(cp)^p\left(\frac{cp}{2}\right)^{\frac{p}{2}}\cdots\left(\frac{cp}{2^{j-1}}\right)^{\frac{p}{2^{j-1}}}\left(\int f_j\,dF\right)^{\frac{p}{2^j}}\\ &\quad+2^{p+\frac{p}{2}+\cdots+\frac{p}{2^{k-1}}}\,(cp)^p\left(\frac{cp}{2}\right)^{\frac{p}{2}}\cdots\left(\frac{cp}{2^{k-1}}\right)^{\frac{p}{2^{k-1}}}\int f_k\,dF.\end{aligned}$$

To simplify, first note that $p+\frac{p}{2}+\cdots+\frac{p}{2^j}\leq 2p$, which may be used for the powers of 2, while it is important to maintain the correct power of $c$:

$$p+\frac{p}{2}+\cdots+\frac{p}{2^{j-1}}=\left(2-\frac{1}{2^{j-1}}\right)p.$$

As for the products of the powers of $\frac{p}{2^j}$, all of them are bounded by the last product (when $j=k-1$), which is equal to

$$p^p\left(\frac{p}{2}\right)^{\frac{p}{2}}\left(\frac{p}{4}\right)^{\frac{p}{4}}\cdots 2^2=\prod_{j=1}^{k}(2^j)^{2^j}=2^{\sum_{j=1}^{k}j\,2^j}=2^{(k-1)\,2^{k+1}+2}.$$

Since $(k-1)\,2^{k+1}+2 \le 2k\cdot 2^k$, one may bound the above product by $2^{2k\cdot 2^k}=p^{2p}$. Using these remarks and bounds, we obtain that

$$\mathbf{E}\left|\int f\,dF_n\right|^p \le 2^p \left|\int f\,dF\right|^p + (2p)^{2p}\sum_{j=1}^{k-1} c^{(2-\frac{1}{2^{j-1}})p}\left(\int f_j\,dF\right)^{\frac{p}{2^j}} \\ + (2p)^{2p}\, c^{(2-\frac{1}{2^{k-1}})p}\int f_k\,dF.$$

Replacing $2^p$ with $(2p)^{2p}$ here, we can summarize the result in a separate statement.

Thus, assume as before that a probability measure $\mu$ on $\mathbf{R}^n$ satisfies a Poincaré-type inequality with constant $\sigma^2$. Let $F_n$ denote the empirical measures with mean $F=\mathbf{E}F_n$ with respect to $\mu$.

**Lemma 9.1** *Given a smooth $F$-integrable function $f:\mathbf{R}\to\mathbf{C}$, for any collection of non-negative smooth functions $(f_j)_{1\le j\le k}$, satisfying (9.2) we have*

$$(2p)^{-2p}\,\mathbf{E}\left|\int f\,dF_n\right|^p \le \left|\int f\,dF\right|^p + \sum_{j=1}^{k} c^{(2-\frac{1}{2^{j-1}})p}\left(\int f_j\,dF\right)^{\frac{p}{2^j}},$$

*where $p=2^k$ and $c=\sigma/\sqrt{n}$.*

Let us now apply this observation to the functions of the form $f(x)=|x-z|^{-q}$ with parameters $z=u+iv$ ($u\in\mathbf{R}$, $v>0$) and $q>0$. Since

$$|f'(x)| = \frac{q\,|x-u|}{|x-z|^{q+2}} \le q\,|x-z|^{-q-1},$$

one may take $f_1(x)=q^2\,|x-z|^{-(2q+2)}$, which is of the same form (up to a factor). Hence, by the same argument, one may take

$$f_2(x)=q^4\,(2q+2)^2\,|x-z|^{-(2(2q+2)+2)} = q^4\,(2q+2)^2\,|x-z|^{-(4q+6)}.$$

Similarly, in the next step,

$$\begin{aligned} f_3(x) &= q^8\,(2q+2)^4\,(4q+6)^2\,|x-z|^{-(2(4q+6)+2)} \\ &= q^8\,(2q+2)^4\,(4q+6)^2\,|x-z|^{-(8q+14)}. \end{aligned}$$

More generally, on the $j$-th step,

$$f_j(x)=a_0^{2^j}\,a_1^{2^{j-1}}\ldots a_{j-1}^{2^1}\,|x-z|^{-a_j},$$

where the sequence $a_j$ is defined recursively by the relation $a_{j+1}=2a_j+2$ with initial value $a_0=q$. The solution to this recursive problem is

$$a_j=(2+q)\,2^j-2.$$

Consequently, using $a_j \le (2+q)\,2^j$,

$$C_j \equiv a_0^{2^j}\, a_1^{2^{j-1}} \ldots a_{j-1}^{2^1} \le (2+q)^{2^j+2^{j-1}+\cdots+2^1}\; 2^{0\cdot 2^j+1\cdot 2^{j-1}+2\cdot 2^{j-2}+\cdots+(j-1)\,2^1}.$$

We have $2^j + 2^{j-1} + \cdots + 2^1 = 2^{j+1} - 2 < 2^{j+1}$. Furthermore,

$$1\cdot 2^{j-1} + 2\cdot 2^{j-2} + \cdots + (j-1)\cdot 2^1 = \sum_{\ell=1}^{j-1} \ell\cdot 2^{j-\ell} = 2^j \sum_{\ell=1}^{j-1} \ell\cdot 2^{-\ell} < 2^{j+1}.$$

With these two bounds, we get $C_j \le (2\,(2+q))^{2^{j+1}}$. Therefore,

$$f_j(x) \le (2(2+q))^{2^{j+1}}\, |x-z|^{-((2+q)\,2^j-2)},$$

and using $|x-z| \ge v$,

$$\begin{aligned}\int f_j\, dF &\le (2\,(2+q))^{2^{j+1}} \int |x-z|^{-((2+q)\,2^j-2)}\, dF(x)\\ &\le (2\,(2+q))^{2^{j+1}}\, \frac{1}{v^{(2+q)\,2^j-3}} \int \frac{v}{|x-z|^2}\, dF(x).\end{aligned}$$

Therefore, recalling that $\int \frac{v}{|x-z|^2}\, dF(x) = \operatorname{Im}(S_F(z))$,

$$\left(\int f_j\, dF\right)^{\frac{p}{2^j}} \le (2\,(2+q))^{2p}\, \frac{1}{v^{((2+q)-3\cdot 2^{-j})\,p}}\, (\operatorname{Im}(S_F(z)))^{p/2^j}.$$

Similarly, when $j=0$,

$$\int f\, dF = \int \frac{1}{|x-z|^q}\, dF \le \frac{1}{v^{q-1}} \int \frac{v}{|x-z|^2}\, dF(x) = \frac{1}{v^{q-1}} \operatorname{Im}(S_F(z)).$$

As a result, we obtain the following from Lemma 9.1.

**Corollary 9.2** *For any integer $k \ge 1$ and $q > 0$, for all $z = u + iv$, $u \in \mathbf{R}$, $v > 0$,*

$$\begin{aligned}&(2p)^{-2p}\, \mathbf{E}\left(\int \frac{1}{|x-z|^q}\, dF_n(x)\right)^p\\ &\quad\le \frac{1}{v^{(q-1)p}}\, (\operatorname{Im}(S_F(z)))^p\\ &\qquad + (2\,(2+q))^{2p} \sum_{j=1}^{k} c^{(2-\frac{1}{2^{j-1}})p}\, \frac{1}{v^{((2+q)-3\cdot 2^{-j})\,p}}\, [\operatorname{Im}(S_F(z))]^{p/2^j},\end{aligned}$$

*where $p = 2^k$ and $c = \sigma/\sqrt{n}$.*

Now, let us return to Proposition 4.3 and apply (4.6) to the particular function $f(x) = \frac{1}{x-z}$. Writing $s_n(z) = S_{F_n}(z)$, $s(z) = S_F(z)$, the inequality (4.6) yields, for all $p \geq 2$,

$$\mathbf{E}\,|s_n(z) - s(z)|^p \leq (cp)^p\, \mathbf{E}\left(\int \frac{1}{|x-z|^4}\, dF_n(x)\right)^{p/2}. \tag{9.3}$$

But by Corollary 9.2 with $q = 4$ and $p/2$ in place of $p$ (where $p = 2^k$, $k \geq 2$),

$$\begin{aligned} &p^{-p}\, \mathbf{E}\left(\int \frac{1}{|x-z|^4}\, dF_n(x)\right)^{p/2} \\ &\quad\leq \frac{1}{v^{3p/2}}\, [\mathrm{Im}(s(z))]^{p/2} \\ &\qquad + (12)^p \sum_{j=1}^{k-1} c^{(1-\frac{1}{2^j})p}\, \frac{1}{v^{3\,(1-2^{-(j+1)})\,p}}\, [\mathrm{Im}(s(z))]^{p/2^{j+1}}. \end{aligned}$$

Hence, with (9.3),

$$\begin{aligned} &p^{-2p}\, \mathbf{E}\,|s_n(z) - s(z)|^p \\ &\quad\leq \frac{c^p}{v^{3p/2}}\, (\mathrm{Im}(s(z)))^{p/2} \\ &\qquad + (12)^p \sum_{j=1}^{k-1} c^{(2-\frac{1}{2^j})p}\, \frac{1}{v^{3\,(1-2^{-(j+1)})\,p}}\, [\mathrm{Im}(s(z))]^{p/2^{j+1}}. \end{aligned}$$

This may be simplified by replacing $j+1$ with $j$, which yields the desired inequality

$$\frac{p^{-2p}}{(12)^p}\, \mathbf{E}\,|s_n(z) - s(z)|^p \leq \sum_{j=1}^{k} \left(\frac{c^2}{v^3}\right)^{(1-\frac{1}{2^j})p} [\mathrm{Im}(s(z))]^{p/2^j}.$$

Proposition 5.1 is proved.

## References


1. Aida, S., Stroock, D.: Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. Math. Res. Lett. **1**, 75–86 (1994)
2. Bai, Z.D.: Convergence rate of expected spectral distributions of large random matrices. Part I. Wigner matrices. Ann. Probab. **21**, 625–648 (1993)
3. Bai, Z.D., Miao, B., Tsay, J.: Convergence rates of the spectral distributions of large Wigner matrices. Int. Math. J. **1**(1), 65–90 (2002)
4. Bhatia, R.: Matrix Analysis. Graduate Texts in Mathematics, vol. 169. Springer, New York (1997) xii+347 pp
5. Bobkov, S.G.: Isoperimetric and analytic inequalities for log-concave probability measures. Ann. Probab. **27**(4), 1903–1921 (1999)
6. Bobkov, S.G., Götze, F.: Discrete isoperimetric and Poincare-type inequalities. Probab. Theory Relat. Fields **114**(2), 245–277 (1999)

7. Bobkov, S.G., Götze, F.: Hardy-type inequalities via Riccati and Sturm-Liouville equations. In: Maz'ya, V. (ed.) Intern. Math. Series. Sobolev Spaces in Mathematics I, Sobolev Type Inequalities, vol. 8, pp. 69–86. Springer, Berlin (2008)
8. Bobkov, S.G., Götze, F.: Concentration of empirical distribution functions with applications to non i.i.d. models. Preprint (2008). To appear in: Bernoulli
9. Bobkov, S.G., Götze, F., Tikhomirov, A.N.: On the concentration of high-dimensioanal matrices with randomly signed entries. J. Math. Sci. **163**(4), 328–351 (2009)
10. Bobkov, S.G., Houdré, C.: Weak dimension-free concentration of measure. Bernoulli **6**(4), 621–632 (2000)
11. Bobkov, S.G., Ledoux, M.: Poincare's inequalities and Talagrand's concentration phenomenon for the exponential distribution. Probab. Theory Relat. Fields **107**, 383–400 (1997)
12. Bobkov, S.G., Ledoux, M.: Weighted Poincaré-type inequalities for Cauchy and other convex measures. Ann. Probab. **37**(2), 403–427 (2009)
13. Borovkov, A.A., Utev, S.A.: On an inequality and a characterization of the normal distribution connected with it. Probab. Theory Appl. **28**, 209–218 (1983)
14. Chatterjee, S., Bose, A.: A new method for bounding rates of convergence of empirical spectral distributions. J. Theor. Probab. **17**(4), 1003–1019 (2004)
15. Davidson, K.R., Szarek, S.J.: Local operator theory, random matrices and Banach spaces. In: Handbook of the Geometry of Banach spaces, vol. I, pp. 317–366. North-Holland, Amsterdam (2001)
16. Erdös, L., Schlein, B., Yau, H.-T.: Local semicircle law and complete delocalization for Wigner random matrices. Commun. Math. Phys. **287**(2), 641–655 (2009)
17. Erdös, L., Schlein, B., Yau, H.-T.: Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. Ann. Probab. **37**(3), 815–852 (2009)
18. Girko, V.L.: Extended proof of the statement: convergence rate of the expected spectral functions of symmetric random matrices $\Xi_n$ is equal to $O(n^{-1/2})$ and the method of critical steepest descent. Random Oper. Stoch. Equ. **10**, 253–300 (2002)
19. Götze, F., Tikhomirov, A.N.: Rate of convergence to the semi-circular law. Probab. Theory Relat. Fields **127**(2), 228–276 (2003)
20. Götze, F., Tikhomirov, A.N.: The rate of convergence for spectra of GUE and LUE matrix ensembles. Cent. Eur. J. Math. **3**(4), 666–704 (2005) (electronic)
21. Götze, F., Tikhomirov, A.N., Timushev, D.A.: Rate of convergence to the semi-circle law for the deformed Gaussian unitary ensemble. Cent. Eur. J. Math. **5**(2), 305–334 (2007) (electronic)
22. Gromov, M., Milman, V.D.: A topological application of the isoperimetric inequality. Am. J. Math. **105**, 843–854 (1983)
23. Guionnet, A., Zeitouni, O.: Concentration of the spectral measure for large matrices. Electron. Commun. Probab. **5**, 119–136 (2000)
24. Kac, I.S., Krein, M.G.: Criteria for the discreteness of the spectrum of a singular string. Izv. Vysš. Učebn. Zaved. Mat. **2**(3), 136–153 (1958) (in Russian)
25. Ledoux, M.: The concentration of measure phenomenon. Math. Surveys and Monogr aphs, vol. 89. AMS (2001)
26. Ledoux, M.: Deviation inequalities on largest eigenvalues. In: Geom. Aspects of Funct. Anal., Israel Seminar 2004–2005. Lecture Notes in Math., vol. 1910, pp. 167–219. Springer, Berlin (2007)
27. Maz'ya, V.G.: Sobolev spaces. Springer, Berlin and New York (1985)
28. Muckenhoupt, B.: Hardy's inequality with weights. Studia Math. **XLIV**, 31–38 (1972)
29. Pastur, L.A.: The spectrum of random matrices. Teoret. Mat. Fiz. **10**(1), 102–112 (1972) (in Russian)
30. Pastur, L.A.: Spectra of random selfadjoint operators. Usp. Mat. Nauk **28**(1), 3–64 (1973) (169) (in Russian)
31. Timushev, D.A.: On the rate of convergence in probability of the spectral distribution function of a random matrix. Theory Probab. Appl. **51**(3), 546–549 (2007)