Elect. Comm. in Probab. **16** (2011), 706–711

ELECTRONIC COMMUNICATIONS in PROBABILITY

ABSOLUTE CONTINUITY OF THE LIMITING EIGENVALUE DISTRIBUTION OF THE RANDOM TOEPLITZ MATRIX

ARNAB SEN [1]
Statistical Laboratory, Center for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WB, UK
email: `a.sen@statslab.cam.ac.uk`

BÁLINT VIRÁG
Departments of Mathematics and Statistics, University of Toronto, Toronto ON, M5S 2E4, Canada
email: `balint@math.toronto.edu`

Submitted September 30, 2011, *accepted in final form* October 19, 2011

AMS 2000 Subject classification: 35P20, 60B20
Keywords: Toeplitz matrix, eigenvalue distribution, spectral averaging

Abstract
We show that the limiting eigenvalue distribution of random symmetric Toeplitz matrices is absolutely continuous with density bounded by 8, partially answering a question of Bryc, Dembo and Jiang (2006). The main tool used in the proof is a spectral averaging technique from the theory of random Schrödinger operators. The similar question for Hankel matrices remains open.

1 Introduction

An $n \times n$ symmetric random Toeplitz matrix is given by

$$\mathbf{T}_n = ((a_{|j-k|}))_{0 \le j,k \le n}$$

where $(a_j)_{j \ge 0}$ is a sequence of i.i.d. random variables with $\mathrm{Var}(a_0) = 1$. For a $m \times m$ Hermitian matrix $\mathbf{A}$, we denote by

$$\mu(\mathbf{A}) := \frac{1}{m} \sum_{i=1}^{m} \delta_{\lambda_i}$$

the empirical eigenvalue distribution of $\mathbf{A}$, where $\lambda_j, 1 \le j \le m$ are the eigenvalues of $\mathbf{A}$, counting multiplicity. Bryc, Dembo and Jiang (2006) established using method of moments that with probability 1, $\mu(n^{-1/2}\mathbf{T}_n)$ converges weakly as $n \to \infty$ to a nonrandom symmetric probability measure γ which does not depend on the distribution of a_0, and has unbounded support. They conjecture (see Remark 1.1 there) that γ has a smooth density. In this note, we give a partial solution:

[1] RESEARCH SUPPORTED BY EPSRC GRANT EP/G055068/1.

Theorem 1. *The measure γ is absolutely continuous with density bounded by* 8.

The actual bound we get is $\frac{16\sqrt{2}}{\pi} = 7.20\ldots$, but we do not expect it to be optimal.
It seems that the method of moments is of little use in determining the existence of the absolute continuity of the limiting eigenvalue distribution. Indeed our proof goes along a completely different path. We make use of the fact that the spectrum of the Gaussian Toeplitz matrix can be realized as that of some diagonal matrix consisting of independent Gaussians conjugated by an appropriate projection matrix - a fact observed in a recent paper Sen and Virág (2011). The next key ingredient of our proof is a spectral averaging technique (Wegner type estimate) developed by Combes, Hislop and Mourre (1996) in connection to the problem of localization for certain families of random Schrödinger operators.
Our proof does not establish further smoothness property of γ. The absolute continuity for the limiting distribution of random Hankel matrices also remains open.

2 Connection between Toeplitz and circulant matrices

Since γ does not depend on the distribution of a_0, from now on, we will assume, without any loss, that $(a_i)_{i\geq 0}$ are i.i.d. standard Gaussian random variables. The remainder of the section we recall some facts about the connection between Toeplitz matrices and circulant matrices from Sen and Virág (2011). Let $\mathbf{T}_n^\circ$ be the symmetric Toeplitz matrix which has $\sqrt{2}a_0$ on its diagonal instead of a_0. It can be easily shown (e.g. using Hoffman-Wielandt inequality, see Bhatia (1997)) that this modification has no effect as far as the limiting eigenvalue distribution is concerned.
$\mathbf{T}_n^\circ$ is the $n\times n$ principal submatrix of a $2n\times 2n$ circulant matrix $\mathbf{C}_{2n} = (b_{j-i \bmod 2n})_{0\leq i,j\leq 2n-1}$, where $b_j = a_j$ for $0<j<n$ and $b_j = a_{2n-j}$ for $n<j<2n$, $b_0=\sqrt{2}a_0$, $b_n = \sqrt{2}a_n$. In other words,

$$\mathbf{Q}_{2n}\mathbf{C}_{2n}\mathbf{Q}_{2n} = \begin{pmatrix} \mathbf{T}_n^\circ & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{0}_n \end{pmatrix}, \quad \text{where } \mathbf{Q}_{2n} = \begin{pmatrix} \mathbf{I}_n & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{0}_n \end{pmatrix}. \tag{1}$$

The circulant matrix can be easily diagonalized as $(2n)^{-1/2}\mathbf{C}_{2n} = \mathbf{U}_{2n}\mathbf{D}_{2n}\mathbf{U}_{2n}^*$ where $\mathbf{U}_{2n}$ is the discrete Fourier transform, i.e. a unitary matrix given by

$$\mathbf{U}_{2n}(j,k) = \frac{1}{\sqrt{2n}}\exp\left(\frac{2\pi ijk}{2n}\right), 0\leq j,k\leq 2n-1$$

and $\mathbf{D}_{2n} = \operatorname{diag}(d_0, d_1, \ldots, d_{2n-1})$, where

$$d_j = \frac{1}{\sqrt{2n}}\sum_{k=0}^{2n-1} b_k \exp\left(\frac{2\pi ijk}{2n}\right) = \frac{1}{\sqrt{2n}}\left[\sqrt{2}a_0 + (-1)^n\sqrt{2}a_n + 2\sum_{k=1}^{n-1} a_k\cos\left(\frac{2\pi jk}{2n}\right)\right].$$

Clearly, $d_j = d_{2n-j}$ for all $n<j<2n$. Also, $(d_j)_{0\leq j\leq n}$ are independent mean zero Gaussian random variables with $\operatorname{Var}(d_j) = 1$ if $0<j<n$ and $\operatorname{Var}(d_j)=2$ if $j\in\{0,n\}$. Define $\mathbf{P}_{2n} := \mathbf{U}_{2n}^*\mathbf{Q}_{2n}\mathbf{U}_{2n}$ so that

$$(2n)^{-1/2}\mathbf{U}_{2n}^*\mathbf{Q}_{2n}\mathbf{C}_{2n}\mathbf{Q}_{2n}\mathbf{U}_{2n} = \mathbf{P}_{2n}\mathbf{D}_{2n}\mathbf{P}_{2n}. \tag{2}$$

Check that $\mathbf{P}_{2n}$ is a Hermitian projection matrix with $\mathbf{P}_{2n}(j,j) = 1/2$ for all j. For notational simplification, we will drop the subscript $2n$ from the relevant matrices unless we want to emphasize the dependence on n.

3 Proof of the main theorem

For a vector $\mathbf{u} \in \mathbb{C}^m$ and a Hemitian matrix $\mathbf{A}$, let $\sigma(\mathbf{A}, \mathbf{u}) := \sum_{i=1}^m |\langle \mathbf{v}_i, \mathbf{u}\rangle|^2 \delta_{\lambda_i}$ be the spectral measure of $\mathbf{A}$ at $\mathbf{u}$, where $\mathbf{A} = \sum_{i=1}^m \lambda_i \mathbf{v}\mathbf{v}_i^*$ is a spectral decomposition of $\mathbf{A}$. For a finite measure ν on $\mathbb{R}$, its Cauchy-Stieltjes transform is given by

$$s(z;\nu) = \int_{\mathbb{R}} \frac{1}{x-z}\nu(dx), \quad z \in \mathbb{C}, \mathrm{Im}(z) > 0.$$

Let $\mathbb{E}\mu(n^{-1/2}\mathbf{T}_n^\circ)$ denote the expected empirical eigenvalue distribution of $n^{-1/2}\mathbf{T}_n^\circ$ which is defined by $\mathbb{E}\mu(n^{-1/2}\mathbf{T}_n^\circ)(B) = \mathbb{E}[\mu(n^{-1/2}\mathbf{T}_n^\circ)(B)]$ for all Borel sets B.

Lemma 2. *Let $(\mathbf{e}_j)_{0\le j\le 2n-1}$ be the coordinate vectors of $\mathbb{R}^{2n}$. Then*

$$s(z;\mathbb{E}\mu(n^{-1/2}\mathbf{T}_n^\circ)) = \frac{\sqrt{2}}{n}\sum_{j=0}^{2n-1} \mathbb{E}\langle \mathbf{P}\mathbf{e}_j, (\mathbf{PDP} - z\mathbf{I})^{-1}\mathbf{P}\mathbf{e}_j\rangle \quad z \in \mathbb{C}, \mathrm{Im}(z) > 0.$$

Before we start proving the above lemma, we state a simple fact about spectral measures of Hermitian matrices.

Lemma 3. *Let $\mathbf{A}$ be an $m \times m$ Hermitian matrix. Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ and $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_\ell$ be vectors in $\mathbb{C}^m$ satisfying $\sum_{i=1}^k \mathbf{u}_i\mathbf{u}_i^* = \sum_{j=1}^\ell \mathbf{v}_j\mathbf{v}_j^*$. Then*

$$\sum_{i=1}^k \sigma(\mathbf{A}, \mathbf{u}_i) = \sum_{j=1}^\ell \sigma(\mathbf{A}, \mathbf{v}_j). \tag{3}$$

Proof of Lemma 3. Let $\mathbf{A} = \sum_{r=1}^m \lambda_r \mathbf{w}_r\mathbf{w}_r^*$ be a spectral decomposition of $\mathbf{A}$. Now for each r, it follows from the definition of the spectral measure that the probability masses at λ_r for the both side of (3) are equal. This completes the proof of the lemma. □

Proof of Lemma 2. By (1), we have

$$s(z;\mu(n^{-1/2}\mathbf{T}_n^\circ)) = \frac{1}{n}\sum_{j=0}^{n-1}\langle \mathbf{e}_j, (n^{-1/2}\mathbf{QCQ} - z\mathbf{I})^{-1}\mathbf{e}_j\rangle,$$

Changing basis as in (2), we can rewrite this as

$$\frac{\sqrt{2}}{n}\sum_{j=0}^{n-1}\langle \mathbf{U}^*\mathbf{e}_j, (\mathbf{PDP} - z\mathbf{I})^{-1}\mathbf{U}^*\mathbf{e}_j\rangle = \frac{\sqrt{2}}{n}\sum_{j=0}^{n-1} s(z;\sigma(\mathbf{PDP}, \mathbf{U}^*\mathbf{e}_j)).$$

Now by Lemma 3 and the fact that $\sum_{j=0}^{n-1}\mathbf{U}^*\mathbf{e}_j\mathbf{e}_j^*\mathbf{U} = \sum_{j=0}^{2n-1}\mathbf{P}\mathbf{e}_j\mathbf{e}_j^*\mathbf{P}$, we deduce

$$s(z;\mu(n^{-1/2}\mathbf{T}_n^\circ)) = \frac{\sqrt{2}}{n}\sum_{j=0}^{2n-1}\langle \mathbf{P}\mathbf{e}_j, (\mathbf{PDP} - z\mathbf{I})^{-1}\mathbf{P}\mathbf{e}_j\rangle. \tag{4}$$

The lemma now follows by taking expectation on both sides of (4) and by observing that for a fixed $z \in \mathbb{C}, \mathrm{Im}(z) \neq 0$, the map $\nu \mapsto s(z;\nu)$ is linear and hence commutes with the expectation. □

Next we will prove a key lemma about the uniform bound on the Stieltjes transform of the expected empirical eigenvalue distribution of Toeplitz matrices.

Lemma 4. *For all n, we have*

$$\sup_{z:\mathrm{Im}(z)>0} |s(z;\mathbb{E}\mu(n^{-1/2}\mathbf{T}_n^{\circ}))| \le 16\sqrt{2}.$$

The above lemma will be a direct consequence of the following result of Combes et al. (1996) on the spectral averaging for one parameter family self-adjoining operators.

Proposition 5 (Combes et al. (1996))**.** *Let $H_\lambda, \lambda \in \mathbb{R}$ be a C^2-family of self-adjoint operators such that $D(H_\lambda) = D_0 \subset \mathscr{H}\ \forall \lambda \in \mathbb{R}$, and such that $(H_\lambda - z)^{-1}$ is twice strongly differentiable in λ for all $z, \mathrm{Im}(z) \neq 0$. Assume that there exist a finite positive constant c_0, and a positive bounded self-adjoint operator B such that, on D_0*

$$\dot{H}_\lambda := \frac{dH_\lambda}{d\lambda} \ge c_0 B^2. \tag{5}$$

Also assume H_λ is linear in λ, i.e., $\ddot{H}_\lambda := \frac{d^2 H_\lambda}{d\lambda^2} = 0$. Then for all $E \in \mathbb{R}$ and twice continuously differentiable function g such that $g, g', g'' \in L^1(\mathbb{R})$ and for all $\varphi \in \mathscr{H}$,

$$\sup_{\delta>0}\left|\int_{\mathbb{R}} g(\lambda)\langle \varphi, B(H_\lambda - E - i\delta)^{-1}B\varphi\rangle d\lambda\right| \le c_0^{-1}(\|g\|_1 + \|g'\|_1 + \|g''\|_1)\|\varphi\|^2. \tag{6}$$

Proposition 5 is an immediate corollary of Theorem 1.1 of Combes et al. (1996) where instead of $\ddot{H}_\lambda = 0$, it was assumed that $|\ddot{H}_\lambda| \le c_1 \dot{H}_\lambda$. The vanishing second derivative assumption shortens the the proof by a considerable amount. We have included a proof of the above proposition in the appendix to make this paper self-contained and also to make constant in the bound (6) explicit.

Proof of Lemma 4. Set $\mathbf{E}_j = \mathbf{e}_j\mathbf{e}_j^* + \mathbf{e}_{2n-j}\mathbf{e}_{2n-j}^*$ for $1 \le j < n$, and $\mathbf{E}_j = \mathbf{e}_j\mathbf{e}_j^*$ for $j \in \{0, n\}$. Take

$$\mathbf{B}_j = \mathbf{P}\mathbf{e}_j\mathbf{e}_j^*\mathbf{P} \text{ or } \mathbf{P}\mathbf{e}_{2n-j}\mathbf{e}_{2n-j}^*\mathbf{P} \text{ for } 1 \le j < n \text{ and } \mathbf{B}_j = \mathbf{P}\mathbf{e}_j\mathbf{e}_j^*\mathbf{P} \text{ for } j \in \{0, n\}. \tag{7}$$

Fix $0 \le j \le n$. We apply Theorem 5 with $H_\lambda = \mathbf{P}\big(\mathbf{D} + (\lambda - d_j)\mathbf{E}_j\big)\mathbf{P}$. In words, we replace d_j and d_{2n-j} by λ in **PDP** to get H_λ. Note that H_λ is random self-adjoint operator which is a function of $\{d_k : 0 \le k \le n, k \neq j\}$. Also, H_λ is linear in λ and so, $\ddot{H}_\lambda = 0$. Since $\dot{H}_\lambda = \mathbf{P}\mathbf{E}_j\mathbf{P} \ge \mathbf{B}_j = \mathbf{P}(j,j)^{-1}\mathbf{B}_j^2$, the condition (5) is satisfied with $B = \mathbf{B}_j$ and $c_0 = 2$ since $\mathbf{P}(j,j) = 1/2$. Take $g = \phi_j$ where ϕ_j be the density of Z for $0 < j < n$ or the density of $\sqrt{2}Z$ for $j \in \{0, n\}$, Z being a standard Gaussian random variable. It is easy to check that $\|g\|_1 = 1, \|g'\|_1 \le \sqrt{\frac{2}{\pi}}, \|g''\|_1 \le 2$. Then plugging $\varphi = \mathbf{e}_j$ or $\mathbf{e}_{2n-j}$ and $\mathbf{B}_j = \mathbf{P}\mathbf{e}_j\mathbf{e}_j^*\mathbf{P}$ or $\mathbf{P}\mathbf{e}_{2n-j}\mathbf{e}_{2n-j}^*\mathbf{P}$ in (6) and taking expectation w.r.t. the remaining randomness $\{d_k : 0 \le k \le n, k \neq j\}$, we obtain

$$\sup_{z:\mathrm{Im}(z)>0} \mathbf{P}(j,j)^2\left|\mathbb{E}\langle \mathbf{P}\mathbf{e}_j, (\mathbf{PDP} - z\mathbf{I})^{-1}\mathbf{P}\mathbf{e}_j\rangle\right| \le c_0^{-1}(\|g\|_1 + \|g'\|_1 + \|g''\|_1) \le 2. \tag{8}$$

The lemma is now immediate from (8) and Lemma 2. □

Proof of Theorem 1. From the inversion formula, $\nu\{(x,y)\} = \lim_{\delta\downarrow 0}\frac{1}{\pi}\int_x^y \mathrm{Im}(s(E+i\delta;\nu))dE$ for all $x<y$ continuity points of ν, it follows that if for some probability measure μ, $\sup_{z:\mathrm{Im}(z)>0}\mathrm{Im}(s(z;\mu)) \le K$ then μ is absolutely continuous w.r.t. Lebesgue measure and its density is bounded by $\pi^{-1}K$. Note that $s(z;\mathbb{E}\mu(n^{-1/2}\mathbf{T}_n^\circ)) \to s(z;\gamma)$ as $n\to\infty$ for each $z\in\mathbb{C}, \mathrm{Im}(z)>0$ since $\mathbb{E}\mu(n^{-1/2}\mathbf{T}_n^\circ)$ converges weakly to γ (see Bryc et al. (2006)). So by Lemma 4, it follows that

$$\sup_{z:\mathrm{Im}(z)>0}|s(z;\gamma)| \le 16\sqrt{2} < 8\pi$$

which completes the proof of the theorem. □

Appendix

Proof of Proposition 5. Define for $\epsilon>0$ and $0<\delta<1$,

$$R(\lambda,\epsilon,\delta) := (H_\lambda - E + i\delta + i\epsilon\dot{H}_\lambda)^{-1} \tag{9}$$

and set

$$K(\lambda,\epsilon,\delta) := BR(\lambda,\epsilon,\delta)B. \tag{10}$$

Note that from assumption (5) ,

$$-\mathrm{Im}\langle\varphi, K(\lambda,\epsilon,\delta)\varphi\rangle = \langle\varphi, BR(\lambda,\epsilon,\delta)^*(\delta+\epsilon\dot{H}_\lambda)R(\lambda,\epsilon,\delta)B\varphi\rangle \ge c_0\epsilon\|K(\lambda,\epsilon,\delta)\varphi\|^2,$$

which, coupled with Cauchy-Schwarz inequality, implies that $\forall\varphi\in\mathscr{H}, \|\varphi\|=1$,

$$\|K(\lambda,\epsilon,\delta)\varphi\| \ge -\mathrm{Im}\langle\varphi,K(\lambda,\epsilon,\delta)\varphi\rangle \ge c_0\epsilon\|K(\lambda,\epsilon,\delta)\varphi\|^2. \tag{11}$$

Now define

$$F(\epsilon,\delta) := \int_{\mathbb{R}} g(\lambda)\langle\varphi,K(\lambda,\epsilon,\delta)\varphi\rangle d\lambda.$$

Inequality (11) implies the bound

$$F(\epsilon,\delta) \le (\epsilon c_0)^{-1}\|g\|_1. \tag{12}$$

Now differentiating F w.r.t. ϵ, we obtain

$$\begin{aligned} i\frac{dF(\epsilon,\delta)}{d\epsilon} &= \int_{\mathbb{R}} g(\lambda)\langle\varphi, BR(\lambda,\epsilon,\delta)\dot{H}_\lambda R(\lambda,\epsilon,\delta)B\varphi\rangle d\lambda \\ &= -\int_{\mathbb{R}} g(\lambda)\frac{d}{d\lambda}\langle\varphi,K(\lambda,\epsilon,\delta)\varphi\rangle d\lambda. \end{aligned}$$

where the last equality follows from the fact $\ddot{H}_\lambda = 0$. Therefore, from (11) and by integration of parts,

$$\left|\frac{dF(\epsilon,\delta)}{d\epsilon}\right| = \left|\int_{\mathbb{R}} g'(\lambda)\langle\varphi,K(\lambda,\epsilon,\delta)\varphi\rangle d\lambda\right| \le (\epsilon c_0)^{-1}\|g'\|_1. \tag{13}$$

By integrating the differential inequality (13) and using the bound (12), we can improve the bound for F as

$$|F(\epsilon,\delta)| \le c_0^{-1}\|g'\|_1\cdot|\log\epsilon| + |F(1,\delta)| \le c_0^{-1}\|g'\|_1\cdot|\log\epsilon| + c_0^{-1}\|g\|_1, \quad \forall\epsilon\in(0,1). \tag{14}$$

Now if we consider the function $\tilde{F}(\epsilon,\delta) := \int_{\mathbb{R}} g'(\lambda)\langle\varphi, K(\lambda,\epsilon,\delta)\varphi\rangle d\lambda$, then by replacing the function g by its derivative g' in (14), we deduce that

$$|\tilde{F}(\epsilon,\delta)| \le c_0^{-1}\|g''\|_1 \cdot |\log\epsilon| + c_0^{-1}\|g'\|_1, \quad \forall\epsilon\in(0,1)$$

which further implies that

$$\left|\frac{dF(\epsilon,\delta)}{d\epsilon}\right| \le c_0^{-1}\|g''\|_1 \cdot |\log\epsilon| + c_0^{-1}\|g'\|_1, \quad \forall\epsilon\in(0,1). \tag{15}$$

Again integrating (15), we get

$$|F(\epsilon,\delta)| \le c_0^{-1}(\|g''\|_1 + \|g'\|_1) + |F(1,\delta)| \le c_0^{-1}(\|g''\|_1 + \|g'\|_1 + \|g\|_1), \tag{16}$$

which holds for all $\epsilon,\delta \in (0,1)$. The proof of the Proposition now follows from the fact that $R(\lambda,\epsilon,\delta)$ converges weakly to $(H_\lambda - E + i\delta)^{-1}$ as $\epsilon \to 0+$ provided $\delta > 0$, and the dominated convergence theorem since $\left|\int_{\mathbb{R}} g(\lambda)\langle\varphi, K(\lambda,\epsilon,\delta)\varphi\rangle d\lambda\right| \le C$, by (16). □

Acknowledgments

Most of the research has been conducted while A.S. visited Technical University of Budapest in July 2011.

References

Bhatia, R. (1997). *Matrix analysis*, Vol. 169 of *Graduate Texts in Mathematics*, Springer-Verlag, New York.

Bryc, W., Dembo, A. and Jiang, T. (2006). Spectral measure of large random Hankel, Markov and Toeplitz matrices, *Ann. Probab.* **34**(1): 1–38.

Combes, J. M., Hislop, P. D. and Mourre, E. (1996). Spectral averaging, perturbation of singular spectra, and localization, *Trans. Amer. Math. Soc.* **348**(12): 4883–4894.

Sen, A. and Virág, B. (2011). The top eigenvalue of the random toeplitz matrix and the sine kernel. available at Arxiv math 1109.5494.