On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices

by

Jack W. Silverstein*
Department of Mathematics
Box 8205
North Carolina State University
Raleigh, North Carolina 27695-8205

and

Z.D. Bai
Department of Applied Mathematics
National Sun Yat-Sen University
Kaohsiung, Taiwan

Summary

A stronger result on the limiting distribution of the eigenvalues of random Hermitian matrices of the form $A+XTX^*$, originally studied in Marčenko and Pastur [4], is presented. Here, X $(N\times n)$, T $(n\times n)$, and A $(N\times N)$ are independent, with X containing i.i.d. entries having finite second moments, T is diagonal with real (diagonal) entries, A is Hermitian, and $n/N \to c > 0$ as $N \to \infty$. Under addtional assumptions on the eigenvalues of A and T, almost sure convergence of the empirical distribution function of the eigenvalues of $A + XTX^*$ is proven with the aid of Stieltjes transforms, taking a more direct approach than previous methods.

* Supported by the National Science Foundation under grant DMS-8903072

AMS 1991 subject classifications. Primary 60F15; Secondary 62H99.

Key Words and Phrases. Random matrix, empirical distribution function of eigenvalues, Stieltjes transform.

1. Introduction. Consider the random matrix XTX^*, where X is $N \times n$ containing independent columns, and T is $n \times n$ Hermitian, independent of X. Several papers have dealt with the behavior of the eigenvalues of this matrix when N and n are both large but having the same order of magnitude (Marčenko and Pastur [4], Grenander and Silverstein [2], Wachter [6], Jonsson [3], Yin and Krishnaiah [8], Yin [7]). The behavior is expressed in terms of limit theorems, as $N \to \infty$, while $n = n(N)$ with $n/N \to c > 0$, on the empirical distribution function (e.d.f.) F^{XTX^*} of the eigenvalues , (that is, $F^{XTX^*}(x)$ is the proportion of eigenvalues of $XTX^* \leq x$), the conclusion being the convergence, in some sense, of F^{XTX^*} to a nonrandom F. The spectral behavior of XTX^* is of significant importance to multivariate statistics. An example of the use of the limiting result can be found in Silverstein and Combettes [5], where it is shown to be effective in solving the detection problem in array signal processing when the (unknown) number of sources is sizable.

The papers vary in the assumptions on T, X, and the type of convergence (almost sure, or in probability), maintaining only one basic condition: F^T converges in distribution (weakly or strongly) to a nonrandom probability distribution function, denoted in this paper by H. However, the assumptions on X share a common intersection: the entries of $\sqrt{N}X$ being i.i.d. for fixed N, same distribution for all N, with unit variance (sum of the variances of real and imaginary parts in the complex case).

In Marčenko and Pastur [4] and Grenander and Silverstein [2], only convergence in probability (at continuity points of F) is established. The others prove strong convergence. It is only in Yin and Krishnaiah [8] and Yin [7] where T is considered to be something other than diagonal, although it is restricted to being nonnegative definite. The weakest assumptions on the entries of X are covered in Yin [7]. All others assume at the least a moment higher than two. A minor difference is the fact that only Marčenko and Pastur [4] and Wachter [6] allow for complex X; the proofs in the other papers can easily be extended to the complex case.

Only Marčenko and Pastur [4] considers arbitrary H. The others assume H to have all moments, relying on the method of moments to prove the limit theorem. These proofs involve intricate combinatorial arguments, some involving graph theory. On the other hand, the proof in Marčenko and Pastur [4] requires no combinatorics. It studies the limiting behavior of the Stieltjes transform

$$m_{XTX^*}(z) = \int \frac{1}{\lambda - z} dF^{XTX^*}(\lambda)$$

of F^{XTX^*}, where $z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : Im\, z > 0\}$. A function in z and $t \in [0, 1]$ is constructed which is shown to converge (in probability) to a solution of a nonrandom first order partial differential equation (p.d.e.), the solution at $t = 1$ being the limiting Stieltjes transform. Using the method of characteristics, this function is seen to be the solution to a certain algebraic equation. Before presenting this equation, it is appropriate to mention at this point that Marčenko and Pastur [4] considered a more general form of matrix, namely $A + XTX^*$, where A is $N \times N$ Hermitian, nonrandom, for which F^A converges vaguely, as $N \to \infty$, to a (possibly defective) distribution function $\mathcal{A}$. Letting $m(z)$ denote the Stieltjes transform of F, and $m_{\mathcal{A}}(z)$ the Stieltjes transform of $\mathcal{A}$, the equation is given by

$$m(z) = m_{\mathcal{A}}\left(z - c\int \frac{\tau dH(\tau)}{1 + \tau m(z)}\right). \tag{1.1}$$

It is proven in Marčenko and Pastur [4] that there is at most one solution to the p.d.e., implying (1.1) uniquely determines the limiting distribution function via a well-known inversion formula for Stieltjes transforms.

The main purpose of the present paper is to extend the result in Marčenko and Pastur [4], again with the aid of Stieltjes transforms, to almost sure convergence under the mild conditions on X assumed in Yin [7], at the same time weakening the assumptions on T (assumed in Marčenko and Pastur [4] to be formed from i.i.d. random variables with d.f. H) and A. Although some aspects require arguments of a more technical nature, the proof is more direct than those mentioned above, avoiding both extensive combinatorial arguments and the need to involve a p.d.e. By delineating the roles played by basic matrix properties and random behavior, it provides for the most part a clear understanding as to why the e.d.f. converges to a nonrandom limit satisfying (1.1).

It is remarked here that the approach taken in this paper is currently being used as a means to extend the result to arbitrary T, and to investigate the convergence of individual eigenvalues associated with boundary points in the support of F (see Silverstein and Combettes [5]).

The remainder of the paper is devoted to proving the following.

Theorem 1.1. Assume

a) For $N = 1, 2, \ldots$ $X_N = (\frac{1}{\sqrt{N}} X_{ij}^N)$, $N \times n$, $X_{ij}^N \in \mathbb{C}$, i.d. for all N, i, j, independent across i, j for each N, $\mathsf{E}|X_{1\,1}^1 - \mathsf{E}X_{1\,1}^1|^2 = 1$.

b) $n = n(N)$ with $n/N \to c > 0$ as $N \to \infty$.

c) $T_N = \operatorname{diag}(\tau_1^N, \ldots, \tau_n^N)$, $\tau_i^N \in \mathbb{R}$, and the e.d.f. of $\{\tau_1^N, \ldots, \tau_n^N\}$ converges almost surely in distribution to a probability distribution function H as $N \to \infty$.

d) $B_N = A_N + X_N T_N X_N^*$, where A_N is Hermitian $N \times N$ for which F^{A_N} converges vaguely to $\mathcal{A}$ almost surely, $\mathcal{A}$ being a (possibly defective) nonrandom d.f.

e) X_N, T_N, and A_N are independent.

Then, almost surely, F^{B_N}, the e.d.f. of the eigenvalues of B_N, converges vaguely, as $N \to \infty$, to a (nonrandom) d.f. F, whose Stieltjes transform $m(z)$ ($z \in \mathbb{C}^+$) satisfies (1.1).

The proof is broken up into several parts. Section 2 presents matrix results, along with results on distribution functions. The main probabilitistic arguments of the proof are contained in section 3. The proof is completed in section 4, while section 5 provides a simple proof of at most one solution $m(z) \in \mathbb{C}^+$ to (1.1) for $z \in \mathbb{C}^+$.

2. Preliminary Results. For rectangular matrix A let $\text{rank}(A)$ denote the rank of A, and for positive integers $i \le \text{rank}(A)$, let s_i^A be the i^{th} largest singular value of A. Define s_i^A to be zero for all $i > \text{rank}(A)$. When A is square having real eigenvalues, λ_i^A will denote the i^{th} largest eigenvalue of A. For $q \in \mathbb{C}^N$, $\|q\|$ will denote the Euclidean norm, and $\|A\|$ the induced spectral norm on matrices (that is, $\|A\| = s_1^A = \sqrt{\lambda_1^{AA^*}}$).

For square C with real eigenvalues, let F^C denote the e.d.f. of the eigenvalues of C. The measure induced by a d.f. G on an interval J will be denoted by $G\{J\}$.

The first three results in the following lemma are well-known. The fourth follows trivially from the fact that the rank of any matrix is the dimension of its row space.

Lemma 2.1.

a) For rectangular matrices A, B of the same size,

$$\text{rank}(A+B) \le \text{rank}(A) + \text{rank}(B).$$

b) For rectangular matrices A, B in which AB is defined,

$$\text{rank}(AB) \le \min(\text{rank}(A), \text{rank}(B)).$$

c) For Hermitian $N \times N$ matrices A, B,

$$\sum_{i=1}^{N} (\lambda_i^A - \lambda_i^B)^2 \le \mathsf{tr}\,(A-B)^2.$$

d) For rectangular A, $\text{rank}(A) \le$ the number of non-zero entries of A.

The following result can be found in Fan [1].

Lemma 2.2. Let m, n be arbitrary non-negative integers. For A, B rectangular matrices of the same size,

$$s_{m+n+1}^{A+B} \le s_{m+1}^A + s_{n+1}^B.$$

For A, B rectangular for which AB is defined

$$s_{m+n+1}^{AB} \le s_{m+1}^A s_{n+1}^B.$$

These inequalities can be expressed in terms of empirical distribution functions. For rectangular A let $\sqrt{AA^*}$ denote the matrix derived from AA^* by replacing in its spectral decomposition the eigenvalues with their square roots. Thus, $\lambda_i^{\sqrt{AA^*}} = s_i^A$.

Lemma 2.3. Let x, y be arbitrary non-negative numbers. For A, B rectangular matrices of the same size,

$$F^{\sqrt{(A+B)(A+B)^*}}\{(x+y,\infty)\} \le F^{\sqrt{AA^*}}\{(x,\infty)\} + F^{\sqrt{BB^*}}\{(y,\infty)\}.$$

If, additionally, A, B are square, then

$$F^{\sqrt{(AB)(AB)^*}}\{(xy,\infty)\} \le F^{\sqrt{AA^*}}\{(x,\infty)\} + F^{\sqrt{BB^*}}\{(y,\infty)\}.$$

Proof. Let N denote the number of rows of A, B. Let $m \ge 0$, $n \ge 0$ be the smallest integers for which $s^A_{m+1} \le x$ and $s^B_{n+1} \le y$. Then $F^{\sqrt{AA^*}}\{(x,\infty)\} = m/N$ and $F^{\sqrt{BB^*}}\{(y,\infty)\} = n/N$, so that $F^{\sqrt{(A+B)(A+B)^*}}\{(s^{A+B}_{m+n+1},\infty)\} \le F^{\sqrt{AA^*}}\{(x,\infty)\} + F^{\sqrt{BB^*}}\{(y,\infty)\}$ in the first case, and $F^{\sqrt{(AB)(AB)^*}}\{(s^{AB}_{m+n+1},\infty)\} \le F^{\sqrt{AA^*}}\{(x,\infty)\} + F^{\sqrt{BB^*}}\{(y,\infty)\}$ in the second case. Applying Lemma 2.2 we get our result.

For any bounded $f : \mathbb{R} \to \mathbb{R}$, let $\|f\| = \sup_x |f(x)|$. Using Lemma 2.2 it is straightforward to verify Lemma 3.5 of Yin [7] which states: For $N \times n$ matrices A, B

$$\|F^{AA^*} - F^{BB^*}\| \le \frac{1}{N}\operatorname{rank}(A - B). \tag{2.1}$$

This result needs to be extended.

Lemma 2.4. For $N \times N$ Hermitian matrices A, B

$$\|F^{A} - F^{B}\| \le \frac{1}{N}\operatorname{rank}(A - B).$$

Proof. Let I denote the $N \times N$ identity matrix and c be any real number for which both $A + cI$ and $B + cI$ are non-negative definite. For any $x \in \mathbb{R}$, $F^A(x) - F^B(x) = F^{(A+cI)^2}((x+c)^2) - F^{(B+cI)^2}((x+c)^2)$. thus, $\|F^A - F^B\| = \|F^{(A+cI)^2} - F^{(B+cI)^2}\|$, and we get our result from (2.1).

The next result follows directly from Lemma 2.1 a), b) and Lemma 2.4.

Lemma 2.5 Let A be $N \times N$ Hermitian, Q, $\overline{Q}$ both $N \times n$, and T, $\overline{T}$ both $n \times n$ Hermitian. Then

a) $$\|F^{A+QTQ^*} - F^{A+\overline{Q}T\overline{Q}^*}\| \le \frac{2}{N}\operatorname{rank}(Q - \overline{Q})$$

and

b) $$\|F^{A+QTQ^*} - F^{A+Q\overline{T}Q^*}\| \le \frac{1}{N}\operatorname{rank}(T - \overline{T}).$$

The next lemma relies on the fact that for $N \times N$ B, $\tau \in \mathbb{C}$, and $q \in \mathbb{C}^N$ for which B and $B + \tau qq^*$ are invertible,

$$q^*(B + \tau qq^*)^{-1} = \frac{1}{1 + \tau q^* B^{-1} q} q^* B^{-1}, \tag{2.2}$$

which follows from $q^*B^{-1}(B + \tau qq^*) = (1 + \tau q^*B^{-1}q)q^*$.

Lemma 2.6. Let $z \in \mathbb{C}^+$ with $v = Im\, z$, A and B $N \times N$ with B Hermitian, $\tau \in \mathbb{R}$, and $q \in \mathbb{C}^N$. Then

$$\left|\mathsf{tr}\left((B - zI)^{-1} - (B + \tau qq^* - zI)^{-1}\right)A\right| \leq \frac{\|A\|}{v}.$$

Proof. Since $(B - zI)^{-1} - (B + \tau qq^* - zI)^{-1} = \tau(B - zI)^{-1}qq^*(B + \tau qq^* - zI)^{-1}$, we have by (2.2)

$$\left|\mathsf{tr}\left((B - zI)^{-1} - (B + \tau qq^* - zI)^{-1}\right)A\right| = \left|\frac{\tau \mathsf{tr}\,(B - zI)^{-1}qq^*(B - zI)^{-1}A}{1 + \tau q^*(B - zI)^{-1}q}\right|$$

$$= \left|\tau \frac{q^*(B - zI)^{-1}A(B - zI)^{-1}q}{1 + \tau q^*(B - zI)^{-1}q}\right| \leq \|A\|\,|\tau| \frac{\|(B - zI)^{-1}q\|^2}{|1 + \tau q^*(B - zI)^{-1}q|}.$$

Write $B = \sum \lambda_i^B e_i e_i^*$ where the e_i's are the orthonormal eigenvectors of B. Then

$$\|(B - zI)^{-1}q\|^2 = \sum \frac{|e_i^* q|^2}{|\lambda_i^B - z|^2},$$

and

$$|1 + \tau q^*(B - zI)^{-1}q| \geq |\tau|\, Im\, q^*(B - zI)^{-1}q = |\tau| v \sum \frac{|e_i^* q|^2}{|\lambda_i^B - z|^2}.$$

The result follows.

Lemma 2.7. Let $z_1, z_2 \in \mathbb{C}^+$ with $\max(Im\, z_1, Im\, z_2) \geq v > 0$, A and B $N \times N$ with A Hermitian, and $q \in \mathbb{C}^N$. Then

$$|\mathsf{tr}\, B((A - z_1 I)^{-1} - (A - z_2 I)^{-1})| \leq |z_2 - z_1| N \|B\| \frac{1}{v^2}, \text{ and}$$

$$|q^* B(A - z_1 I)^{-1}q - q^* B(A - z_2 I)^{-1}q| \leq |z_2 - z_1|\, \|q\|^2 \|B\| \frac{1}{v^2}.$$

Proof. The first inequality follows easily from the fact that for $N \times N$ matrices C, D,

$$|\mathsf{tr}\, CD| \leq (\mathsf{tr}\, CC^* \mathsf{tr}\, DD^*)^{1/2} \leq N\|C\|\,\|D\|,$$

and the fact that $\|(A - z_i I)^{-1}\| \leq 1/v$, $i = 1, 2$. The second inequality follows from the latter observation.

Let $\mathcal{M}(\mathbb{R})$ denote the collection of all sub-probability distribution functions on $\mathbb{R}$. Vague convergence in $\mathcal{M}(\mathbb{R})$ will be denoted by $\xrightarrow{v}$ (that is, $F_N \xrightarrow{v} G$ as $N \to \infty$ means $\lim_{N\to\infty} F_N\{[a, b]\} = G\{[a, b]\}$ for all a, b continuity points of G). We write $F_N \xrightarrow{D} G$ if F_N and G are probability d.f.'s. We denote the d.f. corresponding to the zero measure simply by 0.

Lemma 2.8. For $\{F_N\}_{N=1}^{\infty} \subset \mathcal{M}(\mathbb{R})$, $F_N \neq 0$, such that no subsequence converges vaguely to 0, there exists a positive m such that

$$\inf_N F_N\{[-m, m]\} > 0.$$

Proof. Suppose not. Then a sequence $m_i \to \infty$ and a subsequence $\{F_{N_i}\}_{i=1}^{\infty}$ can be found satisfying $F_{N_i}\{[-m_i, m_i]\} \to 0$, which implies $F_{N_i} \xrightarrow{v} 0$, a contradiction.

Let $\{f_i\}$ be an enumeration of all continuous functions that take a constant $\frac{1}{m}$ value (m a positive integer) on $[a, b]$, where a, b are rational, 0 on $(-\infty, a - \frac{1}{m}] \cup [b + \frac{1}{m}, \infty)$, and linear on each of $[a - \frac{1}{m}, a]$, $[b, b + \frac{1}{m}]$. Standard arguments will yield the fact that for $F_1, F_2 \in \mathcal{M}(\mathbb{R})$

$$D(F_1, F_2) \equiv \sum_{i=1}^{\infty} \left| \int f_i dF_1 - \int f_i dF_2 \right| 2^{-i}$$

is a metric on $\mathcal{M}(\mathbb{R})$ inducing the topology of vague convergence (a variation of this metric has been used in Wachter [6] and Yin [7] on the space of probability d.f.'s). Using the Helly selection theorem, it follows that for $F_N, G_N \in \mathcal{M}(\mathbb{R})$

$$\lim_{N\to\infty} \|F_N - G_N\| = 0 \implies \lim_{N\to\infty} D(F_N, G_N) = 0. \tag{2.3}$$

Since for all i and $x, y \in \mathbb{R}$, $|f_i(x) - f_i(y)| \leq |x - y|$ it follows that for e.d.f.'s F, G on the (respective) sets $\{x_1, \ldots, x_N\}, \{y_1, \ldots, y_N\}$

$$D^2(F, G) \leq \left(\frac{1}{N} \sum_{j=1}^{N} |x_j - y_j| \right)^2 \leq \frac{1}{N} \sum_{j=1}^{N} (x_j - y_j)^2. \tag{2.4}$$

Finally, since for $G \in \mathcal{M}(\mathbb{R})$ the Stieltjes transform $m_G(z) = \int \frac{1}{\lambda - z} dG(\lambda)$ ($z \in \mathbb{C}^+$) possesses the well-known inversion formula

$$G\{[a, b]\} = \frac{1}{\pi} \lim_{\eta \to 0^+} \int_a^b Im\, m_G(\xi + i\eta) d\xi$$

(a, b continuity points of G), it follows that for any countable set $S \subset \mathbb{C}^+$ for which $\mathbb{R} \subset \overline{S}$ (the closure of S), and $F_N, G \in \mathcal{M}(\mathbb{R})$

$$\lim_{N\to\infty} m_{F_N}(z) = m_G(z) \ \forall z \in S \implies F_N \xrightarrow{v} G \ \text{ as } N \to \infty. \tag{2.5}$$

3. Truncation, Centralization, and an Important Lemma. Following along similar lines as Yin [7], we proceed to replace X_N and T_N by matrices suitable for further analysis. To avoid confusion, the dependency of most of the variables on N will occasionally be dropped from the notation. All convergence statements will be as $N \to \infty$.

Let $\widehat{X}_{ij} = X_{ij} I_{(|X_{ij}|<\sqrt{N})}$ and $\widehat{B}_N = A + \widehat{X} T \widehat{X}^*$, where $\widehat{X} = (\frac{1}{\sqrt{N}} \widehat{X}_{ij})$. Using Lemmas 2.5a and 2.1d, it follows as in Yin [7] pp. 58-59 that

$$\|F^{B_N} - F^{\widehat{B}_N}\| \xrightarrow{a.s.} 0. \tag{3.1}$$

Let $\widetilde{B}_N = A + \widetilde{X} T \widetilde{X}^*$ where $\widetilde{X} = \widehat{X} - \mathsf{E}\widehat{X}$ ($\widetilde{X}_{ij} = \widehat{X}_{ij} - \mathsf{E}\widehat{X}_{ij}$). Since $\operatorname{rank}(\mathsf{E}\widehat{X}) \leq 1$, we have from Lemma 2.5a

$$\|F^{\widehat{B}_N} - F^{\widetilde{B}_N}\| \longrightarrow 0. \tag{3.2}$$

For $\alpha > 0$ define $T_\alpha = \operatorname{diag}(\tau_1 I_{(|\tau_1| \leq \alpha)}, \ldots, \tau_n I_{(|\tau_n| \leq \alpha)})$, and let Q be any $N \times n$ matrix. If α and $-\alpha$ are continuity points of H, we have by Lemma 2.5b and assumptions b) and c)

$$\|F^{A+QTQ^*} - F^{A+QT_\alpha Q^*}\| \leq \frac{1}{N} \operatorname{rank}(T - T_\alpha) = \frac{1}{N} \sum_{i=1}^{n} I_{(|\tau_n| > \alpha)} \xrightarrow{a.s.} cH\{[-\alpha, \alpha]^c\}.$$

It follows that if $\alpha = \alpha_N \to \infty$ then

$$\|F^{A+QTQ^*} - F^{A+QT_\alpha Q^*}\| \xrightarrow{a.s.} 0 \tag{3.3}$$

Choose $\alpha = \alpha_N \uparrow \infty$ so that

$$\alpha^4 (\mathsf{E}|X_{1\,1}|^2 I_{(|X_{1\,1}| \geq \ln N)} + \frac{1}{N}) \to 0 \text{ and } \sum_{N=1}^{\infty} \frac{\alpha^8}{N^2} (\mathsf{E}|X_{1\,1}|^4 I_{(|X_{1\,1}| < \sqrt{N})} + 1) < \infty. \tag{3.4}$$

(Note: It is easy to verify $\frac{1}{N^2} \mathsf{E}|X_{1\,1}|^4 I_{(|X_{1\,1}| < \sqrt{N})}$ is summable.)

Let $\overline{X}_{ij} = \widetilde{X}_{ij} I_{(|X_{ij}| < \ln N)} - \mathsf{E}\widetilde{X}_{ij} I_{(|X_{ij}| < \ln N)}$, $\overline{X} = (\frac{1}{\sqrt{N}} \overline{X}_{ij})$, $\overline{\overline{X}}_{ij} = \widetilde{X}_{ij} - \overline{X}_{ij}$, and $\overline{\overline{X}} = (\frac{1}{\sqrt{N}} \overline{\overline{X}}_{ij})$. Then, from (2.4), Lemma 2.1c, and simple applications of the Cauchy-Schwarz inequality we have

$$D^2(F^{A+\widetilde{X} T_\alpha \widetilde{X}^*}, F^{A+\overline{X} T_\alpha \overline{X}^*}) \leq \frac{1}{N} \mathsf{tr}\, (\widetilde{X} T_\alpha \widetilde{X}^* - \overline{X} T_\alpha \overline{X}^*)^2$$

$$= \frac{1}{N} \left[\mathsf{tr}\, (\overline{\overline{X}} T_\alpha \overline{\overline{X}}^*)^2 + \mathsf{tr}\, (\overline{\overline{X}} T_\alpha \overline{X}^* + \overline{X} T_\alpha \overline{\overline{X}}^*)^2 + 2 \mathsf{tr}\, (\overline{\overline{X}} T_\alpha \overline{X}^* + \overline{X} T_\alpha \overline{\overline{X}}^*) \overline{\overline{X}} T_\alpha \overline{\overline{X}}^* \right]$$

$$\leq \frac{1}{N}\left[\mathsf{tr}\,(\overline{\overline{X}}T_\alpha\overline{\overline{X}}^{\,*})^2+4\mathsf{tr}\,(\overline{\overline{X}}T_\alpha\overline{X}^{\,*}\overline{X}T_\alpha\overline{\overline{X}}^{\,*})+4\biggl(\mathsf{tr}\,(\overline{\overline{X}}T_\alpha\overline{X}^{\,*}\overline{X}T_\alpha\overline{\overline{X}}^{\,*})\mathsf{tr}\,(\overline{\overline{X}}T_\alpha\overline{\overline{X}}^{\,*})^2\biggr)^{1/2}\right].$$

It is straightforward to show

$$\mathsf{tr}\,(\overline{\overline{X}}T_\alpha\overline{\overline{X}}^{\,*})^2\leq\alpha^2\mathsf{tr}\,(\overline{\overline{X}}\,\overline{\overline{X}}^{\,*})^2 \text{ and } \mathsf{tr}\,(\overline{\overline{X}}T_\alpha\overline{X}^{\,*}\overline{X}T_\alpha\overline{\overline{X}}^{\,*})\leq(\alpha^4\mathsf{tr}\,(\overline{\overline{X}}\,\overline{\overline{X}}^{\,*})^2\mathsf{tr}\,(\overline{X}\,\overline{X}^{\,*})^2)^{1/2}.$$

Therefore, in order to show

$$D(F^{A+\widetilde{X}T_\alpha\widetilde{X}^{\,*}},F^{A+\overline{X}T_\alpha\overline{X}^{\,*}})\xrightarrow{a.s.}0 \tag{3.5}$$

it is sufficient to verify

$$\alpha^4\frac{1}{N}\mathsf{tr}\,(\overline{\overline{X}}\,\overline{\overline{X}}^{\,*})^2\xrightarrow{a.s.}0 \text{ and } \frac{1}{N}\mathsf{tr}\,(\overline{X}\,\overline{X}^{\,*})^2=O(1)\text{ a.s.} \tag{3.6}$$

For any $N\times n$ matrix $Y=(Y_{ij})$ with $Y_{ij}\in\mathbb{C}$ i.i.d., mean zero, and finite eighth moment, it is straightforward to show for all N

$$\mathsf{E}(\mathsf{tr}\,(YY^*)^2)=Nn\mathsf{E}|Y_{1\,1}|^4+Nn(N+n-2)\mathsf{E}^2|Y_{1\,1}|^2 \tag{3.7}$$

and

$$\begin{aligned}&\mathsf{Var}(\mathsf{tr}\,(YY^*)^2)\\ &\quad\leq K(N^2\mathsf{E}|Y_{1\,1}|^8+N^3(\mathsf{E}|Y_{1\,1}|^6\mathsf{E}|Y_{1\,1}|^2+\mathsf{E}^2|Y_{1\,1}|^4)+N^4(\mathsf{E}|Y_{1\,1}|^4\mathsf{E}^2|Y_{1\,1}|^2+\mathsf{E}^4|Y_{1\,1}|^2)),\end{aligned} \tag{3.8}$$

where K depends only on the maximum of $\frac{n}{N}$. The verification of (3.8) can be facilitated by writing the variance as

$$\sum_{\substack{i\,j\,k\,l\\ \underline{i}\,\underline{j}\,\underline{k}\,\underline{l}}}\mathsf{E}(Y_{i\,j}\bar{Y}_{k\,j}Y_{k\,l}\bar{Y}_{i\,l}Y_{\underline{i}\,\underline{j}}\bar{Y}_{\underline{k}\,\underline{j}}Y_{\underline{k}\,\underline{l}}\bar{Y}_{\underline{i}\,\underline{l}})-\mathsf{E}(Y_{i\,j}\bar{Y}_{k\,j}Y_{k\,l}\bar{Y}_{i\,l})\mathsf{E}(Y_{\underline{i}\,\underline{j}}\bar{Y}_{\underline{k}\,\underline{j}}Y_{\underline{k}\,\underline{l}}\bar{Y}_{\underline{i}\,\underline{l}}) \tag{3.9}$$

(where $\bar{Y}_{a\,b}$ is the complex conjugate of $Y_{a\,b}$) and using the following facts:

1) For any non-zero term in (3.9), at least one of the ordered pairs of indices $(i,j),(k,j),(k,l),(i,l)$ must match up with one of $(\underline{i},\underline{j}),(\underline{k},\underline{j}),(\underline{k},\underline{l}),(\underline{i},\underline{l})$, and none of the eight random variables appear alone.
2) $\mathsf{E}Z^a\,\mathsf{E}Z^b\leq\mathsf{E}Z^{a+b}$ for nonnegative random variable Z and nonnegative a,b.

Since $\mathsf{E}\overline{\overline{X}}_{1\,1}=0$ and $\overline{\overline{X}}_{1\,1}=\widetilde{X}_{1\,1}I_{(|X_{1\,1}|\geq\ln N)}+\mathsf{E}(\widetilde{X}_{1\,1}I_{(|X_{1\,1}|<\ln N)})$, we have

$$\mathsf{E}|\overline{\overline{X}}_{1\,1}|^2=\mathsf{Var}(Re\,\overline{\overline{X}}_{1\,1})+\mathsf{Var}(Im\,\overline{\overline{X}}_{1\,1}) \tag{3.10}$$

$$=\mathsf{Var}(Re\,\widetilde{X}_{1\,1}I_{(|X_{1\,1}|\geq\ln N)})+\mathsf{Var}(Im\,\widetilde{X}_{1\,1}I_{(|X_{1\,1}|\geq\ln N)})\leq\mathsf{E}|\widetilde{X}_{1\,1}|^2I_{(|X_{1\,1}|\geq\ln N)}$$

$$\leq 2(\mathsf{E}|X_{1\,1}|^2 I_{(\ln N \leq |X_{1\,1}| < \sqrt{N})} + |\mathsf{E}X_{1\,1} I_{(|X_{1\,1}| < \sqrt{N})}|^2 \mathsf{P}(|X_{1\,1}| \geq \ln N))$$

$$\leq K \mathsf{E}|X_{1\,1}|^2 I_{(|X_{1\,1}| \geq \ln N)} \to 0.$$

For $m \geq 4$

$$\mathsf{E}|\overline{\overline{X}}_{1\,1}|^m \leq 2^{m-1}(\mathsf{E}|\widetilde{X}_{1\,1}|^m I_{(|X_{1\,1}| \geq \ln N)} + |\mathsf{E}\widetilde{X}_{1\,1} I_{(|X_{1\,1}| < \ln N)}|^m)$$

$$\leq 2^{2(m-1)} \mathsf{E}|X_{1\,1}|^m I_{(\ln N \leq |X_{1\,1}| < \sqrt{N})} + (2^{2(m-1)} + 2^{2m-1})(\mathsf{E}|X_{1\,1}|)^m$$

$$\leq K_m(N^{\frac{m-4}{2}} \mathsf{E}|X_{1\,1}|^4 I_{(\ln N \leq |X_{1\,1}| < \sqrt{N})} + 1).$$

Therefore, using (3.7), (3.10), and (3.4) we find

$$\mathsf{E}\frac{1}{N}\alpha^4 \mathsf{tr}\,(\overline{\overline{X}}\,\overline{\overline{X}}^{\,*})^2 \leq \alpha^4 K(\frac{1}{N}\mathsf{E}|X_{1\,1}|^4 I_{(\ln N \leq |X_{1\,1}| < \sqrt{N})} + \frac{1}{N} + \mathsf{E}^2|\overline{\overline{X}}_{1\,1}|^2)$$

$$\leq \alpha^4 K'(\mathsf{E}|X_{1\,1}|^2 I_{(|X_{1\,1}| \geq \ln N)} + \frac{1}{N}) \to 0,$$

and from (3.8) (using again $\mathsf{E}|X_{1\,1}|^4 I_{(|X_{1\,1}| < \sqrt{N})} \leq N\mathsf{E}|X_{1\,1}|^2$)

$$\mathsf{Var}(\alpha^4 \frac{1}{N}\mathsf{tr}\,(\overline{\overline{X}}\,\overline{\overline{X}}^{\,*})^2) \leq K\frac{\alpha^8}{N^2}(\mathsf{E}|X_{1\,1}|^4 I_{(|X_{1\,1}| < \sqrt{N})} + 1)$$

which, from (3.4), is summable. Therefore, $\alpha^4 \frac{1}{N}\mathsf{tr}\,(\overline{\overline{X}}\,\overline{\overline{X}}^{\,*})^2 \xrightarrow{a.s.} 0$.

Simple applications of the dominated convergence theorem will yield $\mathsf{E}|\overline{X}_{1\,1}|^2 \to \mathsf{E}|X_{1\,1} - \mathsf{E}X_{1\,1}|^2 = 1$. Moreover, it is straightforward to verify for $m \geq 4$

$$\mathsf{E}|\overline{X}_{1\,1}|^m \leq K_m(\ln N)^{m-2}.$$

Thus, from (3.7) we find

$$\mathsf{E}\frac{1}{N}\mathsf{tr}\,(\overline{X}\,\overline{X}^{\,*})^2 \to c(1+c),$$

and from (3.8)

$$\mathsf{Var}(\frac{1}{N}\mathsf{tr}\,(\overline{X}\,\overline{X}^{\,*})^2) \leq K\frac{(\ln N)^2}{n^2}$$

which is summable. Therefore, $\frac{1}{N}\mathsf{tr}\,(\overline{X}\,\overline{X}^{\,*})^2 \xrightarrow{a.s.} c(1+c)$, so that (3.6) holds, which implies (3.5). This, together with (2.3), (2.5), and (3.1-3.3), shows that, in order to prove $F^{B_N} \xrightarrow{v} F_{c,H}$, it is sufficient to verify for any $z \in \mathbb{C}^+$

$$m_{F^{A+\overline{X}T\overline{X}^*}}(z) \xrightarrow{a.s.} m_{F_{c,H}}(z).$$

Notice the matrix $\mathrm{diag}(\mathsf{E}|\overline{X}_{1\,1}|^2\tau_1^N, \ldots, \mathsf{E}|\overline{X}_{1\,1}|^2\tau_n^N)$ also satisfies assumption c) of Theorem 1.1. We will substitute this matrix for T, and replace $\overline{X}$ by $\frac{1}{\sqrt{\mathsf{E}|\overline{X}_{1\,1}|^2}}\overline{X}$ (at least for N sufficiently large so that $\mathsf{E}|\overline{X}_{1\,1}|^2 > 0$). Since $|\widetilde{X}_{ij} I_{(|X_{ij}|<\ln N)}| \le \ln N + \mathsf{E}|X_{1\,1}|$ we have (for N sufficiently large) $\frac{1}{\sqrt{\mathsf{E}|\overline{X}_{1\,1}|^2}}|\overline{X}_{ij}| \le a \ln N$ for some $a > 2$. Let $\log N$ denote the logarithm of N with base $e^{1/a}$ (so that $a \ln N = \log N$). Simplifying notation, we write $B_N = A + XTX^*$ with $X = (\frac{1}{\sqrt{N}} X_{ij})$ where,

1) X_{ij} are i.i.d. for fixed N,
2) $|X_{1\,1}| \le \log N$,
3) $\mathsf{E}X_{ij} = 0$, $\mathsf{E}|X_{1\,1}|^2 = 1$,

and, along with assumptions b), c), d) of Theorem 1.1, proceed to show for any $z \in \mathbb{C}^+$

$$m_{F^{B_N}}(z) \xrightarrow{a.s.} m_{F_{c,H}}(z).$$

As will be seen in the next section, the following lemma and (2.2) contribute the most to the truth of Theorem 1.1.

Lemma 3.1. Let $C = (c_{ij})$, $c_{ij} \in \mathbb{C}$, be an $N \times N$ matrix with $\|C\| \le 1$, and $Y = (X_1, \ldots, X_N)^T$, $X_i \in \mathbb{C}$, where the X_i's are i.i.d. satisfying conditions 2) and 3). Then

$$\mathsf{E}|Y^*CY - \mathsf{tr}\, C|^6 \le KN^3 \log^{12} N \tag{3.11}$$

where the constant K does not depend on N, C, nor on the distribution of X_1.

Proof: We first consider C real. Since $1 \ge \|C\| = (\lambda_1^{CC^T})^{1/2}$, it follows that $|c_{ii}| \le 1$ for each i. We have

$$\mathsf{E}|Y^*CY - \mathsf{tr}\, C|^6 \le 32\bigg(\mathsf{E}\big(\sum_i c_{ii}(|X_i|^2 - 1)\big)^6 + \mathsf{E}\Big|\sum_{i\ne j} c_{ij}\bar{X}_i X_j\Big|^6\bigg).$$

For the first sum the expansion is straightforward:

$$\mathsf{E}\big(\sum_i c_{ii}(|X_i|^2 - 1)\big)^6 \le K\log^{12} N\bigg(\sum_i |c_{ii}|^6 + \sum_{i\ne j}(c_{ii}^2 c_{jj}^4 + |c_{ii}c_{jj}|^3)$$

$$+ \sum_{\substack{i\,j\,k \\ \text{distinct}}} (c_{ii}c_{jj}c_{kk})^2\bigg) \le KN^3 \log^{12} N$$

For the second sum we have the expression

$$\sum_{\substack{i_1\ne j_1 \\ \vdots \\ i_6 \ne j_6}} c_{i_1\,j_1}c_{i_2\,j_2}c_{i_3\,j_3}c_{j_4\,i_4}c_{j_5\,i_5}c_{j_6\,i_6}\mathsf{E}\bar{X}_{i_1}X_{j_1}\bar{X}_{i_2}X_{j_2}\bar{X}_{i_3}X_{j_3}X_{i_4}\bar{X}_{j_4}X_{i_5}\bar{X}_{j_5}X_{i_6}\bar{X}_{j_6}.$$

Notice a term will be zero if any X_k appears alone. The sum can be further decomposed into sums where each one corresponds to a partitioning of the 12 indices, with each set in the partition containing at least 2 indices, none containing any pair i_l, j_l. Consider one such sum. The summation is performed by 1) restricting the indices in the same partition set to take on the same value, and 2) not allowing indices from different partition sets to take on the same value. The expected value part will be the same for each term and can be factored out. It is bounded in absolute value by $\log^{12} N$. The sum can be further decomposed, using an inclusion-exclusion scheme, where each resulting sum only satisfies 1). By Lemma 3.4 of [Yin], each of these sums is bounded in absolute value by

$$\Big(\sum_{i_1 \neq j_1} c_{i_1\,j_1}^2 \cdots \sum_{i_6 \neq j_6} c_{i_6\,j_6}^2\Big)^{1/2} \leq N^3$$

since $\sum c_{i_1\,j_1}^2 = \mathsf{tr}\, CC^T \leq N$. Thus we get (3.11).

For arbitrary C we write $C = C_1 + iC_2$ with C_1 and C_2 real. It is a simple matter to verify $\max(\|C_1\|, \|C_2\|) \leq \|C\|$. Using this, the inequality

$$|Y^*CY - \mathsf{tr}\, C| \leq |Y^*C_1Y - \mathsf{tr}\, C_1| + |Y^*C_2Y - \mathsf{tr}\, C_2|$$

and the truth of (3.11) for real matrices, we get our result.

4. Completing the Proof of Theorem 1.1. Fix $z = u + iv \in \mathbb{C}^+$. We begin by separating the convergence of F^{A_N} into two cases. Consider first the behavior of $F^{X_N T_N X_N^*}$. It is straightforward to verify $\frac{1}{n}\mathsf{tr}\, X_N^* X_N \xrightarrow{a.s.} 1$, either by using Lemma 3.1 (with $C = I$), or by just taking the fourth moment of $\frac{1}{n}\mathsf{tr}\, X_N^* X_N - 1$. This implies that, almost surely, any vaguely convergent subsequence of F^{X^*X} must be proper, or, in other words, the sequence $\{F^{X_N^* X_N}\}$ is almost surely tight. Since $F^{T_N} \xrightarrow{D} H$ a.s., it follows from the second inequality in Lemma 2.3 that, almost surely, $\{F^{X_N T_N X_N^*}\}$ is tight.

Now suppose $\mathcal{A} = 0$, that is, almost surely, only $o(N)$ eigenvalues of A_N remain bounded. Let, for Hermitian C, $|C|$ denote the matrix derived from C by replacing in its spectral decomposition the eigenvalues with their absolute values (the singular values of C). Writing $A_N = B_N - X_N T_N X_N^*$, we have from the first inequality in Lemma 2.2

$$F^{|A_N|}\{(x+y,\infty)\} \le F^{|B_N|}\{(x,\infty)\} + F^{|X_N T_N X_N^*|}\{(y,\infty)\}$$

for non-negative x, y. It follows that for every $x \ge 0$, $F^{|B_N|}\{(x,\infty)\} \xrightarrow{a.s.} 1$, that is, $F^{B_N} \xrightarrow{v} 0$ almost surely. Thus, the Stieltjes transforms of F^{A_N} and F^{B_N} converge almost surely to zero, the limits obviously satisfying (1.1).

We assume for the remainder of the proof $\mathcal{A} \ne 0$. Using again the first inequality in Lemma 2.2 it follows that, whenever $\{F^{X_N T_N X_N^*}\}$ is tight, and $F^{B_{N_i}} \xrightarrow{v} 0$ on a subsequence $\{N_i\}$, we have $F^{A_{N_i}} \xrightarrow{v} 0$. Thus, $\{F^{B_N}\}$ almost surely satisfies the conditions of Lemma 2.8. Therefore, the quantity

$$\delta = \inf_N Im(m_{F^{B_N}}) \ge \inf \int \frac{v\, dF^{B_N}(x)}{2(x^2+u^2)+v^2}$$

is positive almost surely.

For $i = 1, 2, \ldots, n$, let q_i $(= q_i^N)$ denote the i^{th} column of X_N, $B_{(i)} = B_{(i)}^N = B_N - \tau_i q_i q_i^*$,

$$x = x_N = \frac{1}{N}\sum_{i=1}^{n} \frac{\tau_i}{1+\tau_i m_{F^{B_N}}(z)} \quad \text{and} \quad x_{(i)} = x_{(i)}^N = \frac{1}{N}\sum_{j=1}^{n} \frac{\tau_j}{1+\tau_j m_{F^{B_{(i)}}}(z)}.$$

Notice both $Im\, x$ and $Im\, x_{(i)}$ are non-positive. Write $B_N - zI = A_N - (z-x)I + X_N T_N X_N^* - xI$. Then

$$(A_N - (z-x)I)^{-1} = (B_N - zI)^{-1} + (A_N - (z-x)I)^{-1}(X_N T_N X_N^* - xI)(B_N - zI)^{-1},$$

and, using (2.2)

$$\text{(4.1)} \quad m_{A_N}(z-x) - m_{B_N}(z) = \frac{1}{N}\mathsf{tr}\,(A_N - (z-x)I)^{-1}\left(\sum_{i=1}^{n} \tau_i q_i q_i^* - xI\right)(B_N - zI)^{-1}$$

$$= \frac{1}{N}\mathsf{tr}\,(A_N-(z-x)I)^{-1}\left(\sum_{i=1}^n \frac{\tau_i}{1+\tau_i q_i^*(B_{(i)}-zI)^{-1}q_i} q_i q_i^*(B_{(i)}-zI)^{-1} - x(B_N-zI)^{-1}\right)$$

$$= \frac{1}{N}\sum_{i=1}^n \frac{\tau_i}{1+\tau_i m_{F^{B_N}}(z)} d_i,$$

where $d_i = d_i^N =$

$$\frac{1+\tau_i m_{F^{B_N}}(z)}{1+\tau_i q_i^*(B_{(i)}-zI)^{-1}q_i} q_i^*(B_{(i)}-zI)^{-1}(A_N-(z-x)I)^{-1}q_i - \frac{1}{N}\mathsf{tr}\,(B_N-zI)^{-1}(A_N-(z-x)I)^{-1}.$$

Notice the norms of $(B_N - zI)^{-1}$, $(B_{(i)} - zI)^{-1}$, $(A_N - (z-x)I)^{-1}$, and $(A_N - (z - x_{(i)})I)^{-1}$ are all bounded by $1/v$. Using the fact that q_i is independent of both $B_{(i)}$ and $x_{(i)}$, we have by Lemma 3.1

$$\mathsf{E}|\,\|q_i\|^2 - 1|^6 \le K\frac{\log^{12} N}{N^3}, \quad \mathsf{E}|q_i^*(B_{(i)}-zI)^{-1}q_i - \frac{1}{N}\mathsf{tr}\,(B_{(i)}-zI)^{-1}|^6 \le \frac{K}{v^6}\frac{\log^{12} N}{N^3}, \text{ and}$$

$$\mathsf{E}|q_i^*(B_{(i)}-zI)^{-1}(A_N-(z-x_{(i)})I)^{-1}q_i - \frac{1}{N}\mathsf{tr}\,(B_{(i)}-zI)^{-1}(A_N-(z-x_{(i)})I)^{-1}|^6 \le \frac{K}{v^{12}}\frac{\log^{12} N}{N^3}.$$

This is enough to ensure that, almost surely,

$$\begin{aligned}(4.2)\quad & \max_{i\le n}\max\big[|\,\|q_i\|^2-1|, |q_i^*(B_{(i)}-zI)^{-1}q_i - m_{F^{B_{(i)}}}(z)|, \\ & |q_i^*(B_{(i)}-zI)^{-1}(A_N-(z-x_{(i)})I)^{-1}q_i - \frac{1}{N}\mathsf{tr}\,(B_{(i)}-zI)^{-1}(A_N-(z-x_{(i)})I)^{-1}|\big] \to 0.\end{aligned}$$

We concentrate now on a realization for which (4.2) holds, $\delta > 0$, $F^{T_N} \xrightarrow{D} H$, and $F^{A_N} \xrightarrow{v} \mathcal{A}$. Lemma 2.6 gives us

$$(4.3)\qquad \max_{i\le n}\max\left[|m_{F^{B_N}}(z) - m_{F^{B_{(i)}}}(z)|, |m_{F^{B_N}}(z) - q_i^*(B_{(i)}-zI)^{-1}q_i|\right] \to 0.$$

For N large enough so that

$$\max_{i\le n}\max\left[|Im\, m_{F^{B_N}}(z) - Im\, m_{F^{B_{(i)}}}(z)|, |Im\, m_{F^{B_N}}(z) - Im\, q_i^*(B_{(i)}-zI)^{-1}q_i|\right] < \frac{\delta}{2},$$

we have for i, $j \le n$

$$\left|\frac{1+\tau_i m_{F^{B_N}}(z)}{1+\tau_i q_i^*(B_{(i)}-zI)^{-1}q_i} - 1\right| < \frac{2}{\delta}|m_{F^{B_N}}(z) - q_i^*(B_{(i)}-zI)^{-1}q_i|,$$

and

$$\left|\frac{\tau_j}{1+\tau_j m_{F^{B_N}}(z)} - \frac{\tau_j}{1+\tau_j m_{F^{B_{(i)}}}(z)}\right| \le \frac{2}{\delta^2}|m_{F^{B_N}}(z) - m_{F^{B_{(i)}}}(z)|.$$

Therefore,

$$\max_{i\leq n}\max\left[\left|\frac{1+\tau_i m_{F^{B_N}}(z)}{1+\tau_i q_i^*(B_{(i)}-zI)^{-1}q_i}-1\right|, |x-x_{(i)}|\right]\to 0. \tag{4.4}$$

Using Lemmas 2.6, 2.7, (4.2)-(4.4), we have $\max_{i\leq n} d_i \to 0$, and since

$$\left|\frac{\tau_i}{1+\tau_i m_{F^{B_N}}(z)}\right| \leq \frac{1}{\delta},$$

we conclude from (4.1) that

$$m_{A_N}(z-x) - m_{B_N}(z) \to 0$$

Consider a subsequence $\{N_i\}$ on which $m_{F^{B_{N_i}}}(z)$ converges to a number m. Since

$$f(\tau) = \frac{\tau}{1+\tau m}$$

is bounded, and

$$\left|\frac{\tau}{1+\tau m_{F^{B_{N_i}}}(z)} - f(\tau)\right| \leq \frac{1}{\delta^2}|m_{F^{B_{N_i}}}(z) - m|,$$

it follows that, along $\{N_i\}$,

$$x\ (= x_{N_i}) \to c\int \frac{\tau dH(\tau)}{1+\tau m},$$

so that

$$m = m_{\mathcal{A}}\big(z - c\int \frac{\tau dH(\tau)}{1+\tau m}\big).$$

Therefore, m is unique (Marčenko and Pastur [4] or Section 5 below), and we must have $m_{F^{B_N}}(z) \to m$, an event which occurs with probability 1. Therefore, using (2.5), the proof is complete.

5. A Proof of Uniqueness. The following lemma renders the proof of Theorem 1.1 to be free of any dependence on arguments involving p.d.e.'s.

Lemma 5.1. For $z = z_1 + iz_2 \in \mathbb{C}^+$, there exists at most one $m \in \mathbb{C}^+$ s.t.

$$m = \int \frac{d\mathcal{A}(\tau)}{\tau - \left(z - c\int \frac{\lambda dH(\lambda)}{1+\lambda m}\right)}. \tag{5.1}$$

Proof. For $m = m_1 + im_2 \in \mathbb{C}^+$ satisfying (5.1) we have

$$m = \int \frac{d\mathcal{A}(\tau)}{\tau - z_1 + c\int \frac{\lambda(1+\lambda m_1)dH(\lambda)}{(1+\lambda m_1)^2+\lambda^2 m_2^2} - i\left(z_2 + c\int \frac{\lambda^2 m_2 dH(\lambda)}{(1+\lambda m_1)^2+\lambda^2 m_2^2}\right)}.$$

Therefore

$$m_2 = \left(z_2 + c\int \frac{\lambda^2 m_2 dH(\lambda)}{|1+\lambda m|^2}\right)\int \frac{d\mathcal{A}(\tau)}{\left|\tau - z + c\int \frac{\lambda dH(\lambda)}{1+\lambda m}\right|^2}. \tag{5.2}$$

Suppose $\underline{m} = \underline{m}_1 + i\underline{m}_2 \in \mathbb{C}^+$ also satisfies (5.1). Then

$$m - \underline{m} = \tag{5.3}$$

$$c\int \frac{\left(\int \frac{\lambda dH(\lambda)}{1+\lambda \underline{m}} - \int \frac{\lambda dH(\lambda)}{1+\lambda m}\right) d\mathcal{A}(\tau)}{\left(\tau - z + c\int \frac{\lambda dH(\lambda)}{1+\lambda m}\right)\left(\tau - z + c\int \frac{\lambda dH(\lambda)}{1+\lambda \underline{m}}\right)} =$$

$$(m - \underline{m})c\int \frac{\lambda^2 dH(\lambda)}{(1+\lambda m)(1+\lambda \underline{m})}\int \frac{d\mathcal{A}(\tau)}{\left(\tau - z + c\int \frac{\lambda dH(\lambda)}{1+\lambda m}\right)\left(\tau - z + c\int \frac{\lambda dH(\lambda)}{1+\lambda \underline{m}}\right)}.$$

Using Hölder's inequality and (5.2) we have

$$\left| c\int \frac{\lambda^2 dH(\lambda)}{(1+\lambda m)(1+\lambda \underline{m})}\int \frac{d\mathcal{A}(\tau)}{\left(\tau - z + c\int \frac{\lambda dH(\lambda)}{1+\lambda m}\right)\left(\tau - z + c\int \frac{\lambda dH(\lambda)}{1+\lambda \underline{m}}\right)} \right|$$

$$\leq \left(c\int \frac{\lambda^2 dH(\lambda)}{|1+\lambda m|^2}\int \frac{d\mathcal{A}(\tau)}{\left|\tau - z + c\int \frac{\lambda dH(\lambda)}{1+\lambda m}\right|^2}\right)^{1/2}\left(c\int \frac{\lambda^2 dH(\lambda)}{|1+\lambda \underline{m}|^2}\int \frac{d\mathcal{A}(\tau)}{\left|\tau - z + c\int \frac{\lambda dH(\lambda)}{1+\lambda \underline{m}}\right|^2}\right)^{1/2}$$

$$= \left(c\int \frac{\lambda^2 dH(\lambda)}{|1+\lambda m|^2}\frac{m_2}{\left(z_2 + c\int \frac{\lambda^2 m_2 dH(\lambda)}{|1+\lambda m|^2}\right)}\right)^{1/2}\left(c\int \frac{\lambda^2 dH(\lambda)}{|1+\lambda \underline{m}|^2}\frac{\underline{m}_2}{\left(z_2 + c\int \frac{\lambda^2 \underline{m}_2 dH(\lambda)}{|1+\lambda \underline{m}|^2}\right)}\right)^{1/2} < 1.$$

Therefore, from (5.3) we must have $m = \underline{m}$.

REFERENCES

[1] Fan, K. (1951). Maximum properties and inequalities for the eigenvalues of completely continuous operators. *Proc. Nat. Acad. Sci. U.S.A.* **37** 760-766.

[2] Grenander, U., and Silverstein, J.W. (1977). Spectral analysis of networks with random topologies. *SIAM J. Appl. Math.* **32** 499-519.

[3] Jonsson, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1-38.

[4] Marčenko, V.A. and Pastur, L.A. (1967). Distribution of eigenvalues for some sets of random matrices. Math. USSR-Sb. **1** 457-483.

[5] Silverstein, J.W. and Combettes, P.L. (1992). Signal detection via spectral theory of large dimensional random matrices. *IEEE Trans. Signal Processing* **40** 2100-2105.

[6] Wachter, K.W. (1978). The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probab.* **6** 1-18.

[7] Yin, Y.Q. (1986). Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.* **20** 50-68.

[8] Yin, Y.Q., and Krishnaiah, P.K. (1984). A limit theorem for the eigenvalues of product of two random matrices. *J. Multivariate Anal.* **13** 489-507.