Spectral Properties of Large Dimensional Random Matrices

Jack W. Silverstein

Department of Mathematics

North Carolina State University

1. Introduction. Random matrices appear in

I. Multivariate statistics

A. Sample covariance matrix

$$\frac{1}{N} T_n^{1/2} X_n X_n^* T_n^{1/2}$$

where $T_n^{1/2}$ is $n \times n$, square root of T_n, $X_n = [X_{\cdot 1}, \ldots X_{\cdot N}]$ is $n \times N$ consisting of i.i.d. entries. Since

$$\frac{1}{N} T_n^{1/2} X_n X_n^* T_n^{1/2} = \frac{1}{N} \sum_{j=1}^{N} T_n^{1/2} X_{\cdot j} (T_n^{1/2} X_{\cdot j})^*,$$

when the entries of X_n are known to have mean zero, this matrix can be viewed as the sample covariance matrix of the random vector $T_n^{1/2} X_{\cdot 1}$ which has population covariance matrix

$$\mathsf{E}(T_n^{1/2} X_{\cdot 1})(T_n^{1/2} X_{\cdot 1})^* = T_n.$$

Includes all Wishart matrices.

B. Multivariate F matrix

$$\left(\frac{1}{N}X_nX_n^*\right)\left(\frac{1}{\widehat{N}}\widehat{X}_n\widehat{X}_n^*\right)^{-1},$$

where $\widehat{X}_n$ is $n \times \widehat{N}$, $n < \widehat{N}$, independent of X_n consisting of i.i.d. entries with the same distribution as in X_n.

C. Information-plus-noise type matrix

$$\frac{1}{N}(R_n + \sigma X_n)(R_n + \sigma X_n)^*$$

where R_n, $n \times N$ is independent of X_n and $\sigma > 0$. Another matrix of sample covariance type, appearing in array signal processing, where all the information is contained in

$$\frac{1}{N}R_nR_n^*$$

but additive noise is contaminating the signal (S. and Combettes (1992)).

Work on these matrices consider n large but the number of samples, N, needed to apply standard mutivariate statistical methodology is unattainable.

II. Wireless Communications

A. Matrices used in modeling communications between N transmit antennas and n receive antennas in multiple-input-multiple-output (MIMO) systems:

$$\frac{1}{N}T_n^{1/2}X_nS_nX_n^*T_n^{1/2}$$

where S_n is $N \times N$ is nonnegative definite. S_n and T_n are the correlation matrices between, respectively the transmit antennas and the receive antennas (Chuah, et.al. (2002)).

B. Matrices used in modeling *direct-sequence code-division multiple-access* (or DS-CDMA) systems (to be discussed later).

III. Mathematical Physics

A. Discrete analogue of the one dimensional Schrödinger equation with random potential

$$A_n + \frac{1}{N} X_n^* T_n X_n$$

where A_n is $N \times N$, T_n is diagonal (Marčenko and Pastur (1967)).

B Wigner matrix, Hermitian with independent entries on and above the diagonal, used in high energy physics.

In all these models, eigenvalues play a crucial role in applications.

The Mathematics

Let $\mathcal{M}(\mathbb{R})$ denote the collection of all subprobability distribution functions on $\mathbb{R}$. We say for $\{F_n\} \subset \mathcal{M}(\mathbb{R})$, F_n converges vaguely to $F \in \mathcal{M}(\mathbb{R})$ (written $F_n \xrightarrow{v} F$) if for all $[a, b]$, a, b continuity points of F, $\lim_{n\to\infty} F_n\{[a, b]\} = F\{[a, b]\}$. We write $F_n \xrightarrow{D} F$, when F_n, F are probability distribution functions (equivalent to $\lim_{n\to\infty} F_n(a) = F(a)$ for all continuity points a of F).

For $F \in \mathcal{M}(\mathbb{R})$,

$$m_F(z) \equiv \int \frac{1}{x - z} dF(x), \quad z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : \Im z > 0\}$$

is defined as the Stieltjes transform of F.

Properties:

1. m_F is an analytic function on $\mathbb{C}^+$.
2. $\Im m_F(z) > 0$.
3. $|m_F(z)| \leq \frac{1}{\Im z}$.
4. For continuity points $a < b$ of F

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

since the right hand side

$$= \frac{1}{\pi} \lim_{\eta \to 0^+} \int_a^b \int \frac{\eta}{(x-\xi)^2 + \eta^2} dF(x) d\xi$$

$$= \frac{1}{\pi} \lim_{\eta \to 0^+} \int \int_a^b \frac{\eta}{(x-\xi)^2 + \eta^2} d\xi dF(x)$$

$$= \frac{1}{\pi} \lim_{\eta \to 0^+} \int \left[\mathrm{Tan}^{-1}\left(\frac{b-x}{\eta}\right) - \mathrm{Tan}^{-1}\left(\frac{a-x}{\eta}\right) \right] dF(x)$$

$$= \int I_{[a,b]} dF(x) = F\{[a,b]\}.$$

5. If, for $x_0 \in \mathbb{R}$, $\Im m_F(x_0) \equiv \lim_{z\in\mathbb{C}^+\to x_0} \Im m_F(z)$ exists, then F is differentiable at x_0 with value $(\frac{1}{\pi})\Im m_F(x_0)$ (S. and Choi (1995)).

Proof. Fix $\epsilon > 0$. Let $\delta > 0$ be s.t. $\frac{1}{\pi}|\Im\, m_F(x+iy) - \Im\, m_F(x_0)| < \frac{\epsilon}{2}$ whenever $|x - x_0| < \delta$, $y \in (0, \delta)$. Let $x_1 < x_2$ be continuity points of F s.t. $x_1 < x_2$ and $|x_i - x_0| < \delta$, $i = 1, 2$. From 4. we can choose $y \in (0, \delta)$ s.t. $\left|F(x_2) - F(x_1) - \frac{1}{\pi}\int_{x_1}^{x_2} \Im\, m_F(x+iy)dx\right| < \frac{\epsilon}{2}(x_2 - x_1)$. For any $x \in [x_1, x_2]$, we have $|x - x_0| < \delta$. Thus

$$\left|\frac{F(x_2) - F(x_1)}{x_2 - x_1} - \frac{1}{\pi}\Im\, m_F(x_0)\right|$$

$$\leq \frac{1}{x_2 - x_1}\left|F(x_2) - F(x_1) - \frac{1}{\pi}\int_{x_1}^{x_2} \Im\, m_F(x+iy)dx\right|$$
$$+ \frac{1}{x_2 - x_1}\int_{x_1}^{x_2}\left|\frac{1}{\pi}\left(\Im\, m_F(x+iy) - \Im\, m_F(x_0)\right)\right| dx < \epsilon.$$

It follows that F is continuous at x_0, and for any sequence $\{x_n\}$ of continuity points of F converging to x_0

$$\lim_{n\to\infty} \frac{F(x_n) - F(x_0)}{x_n - x_0} = \frac{1}{\pi} Im\, m_F(x_0). \tag{1.1}$$

Let $\{x_n\}$ be any real sequence satisfying $x_n \downarrow x_0$. For each n choose continuity points $x_{cp}^{(n)-}, x_{cp}^{(n)+}$ s.t.

$$x_0 < x_{cp}^{(n)-} \le x_n \le x_{cp}^{(n)+},$$

$$(1 - \frac{1}{n})(x_n - x_0) \le x_{cp}^{(n)-} - x_0,$$

and

$$x_{cp}^{(n)+} - x_0 \le (1 + \frac{1}{n})(x_n - x_0).$$

Then

$$\left(1 - \frac{1}{n}\right) \frac{F(x_{cp}^{(n)-}) - F(x_0)}{x_{cp}^{(n)-} - x_0} \le \frac{F(x_n) - F(x_0)}{x_n - x_0} \le \left(1 + \frac{1}{n}\right) \frac{F(x_{cp}^{(n)+}) - F(x_0)}{x_{cp}^{(n)+} - x_0},$$

and we have (1.1) holding for this sequence. A similar argument can be made for $\{x_n\}$ with $x_n \uparrow x_0$. This complete the proof.

Let $S \subset \mathbb{C}^+$ be countable with a cluster point in $\mathbb{C}^+$. Using 4., the fact that $F_n \xrightarrow{v} F$ is equivalent to

$$\int f(x)dF_n(x) \to \int f(x)dF(x)$$

for all continuous f vanishing at $\pm\infty$, and the fact that an analytic function defined on $\mathbb{C}^+$ is uniquely determined by the values it takes on S, we have

$$F_n \xrightarrow{v} F \quad \Longleftrightarrow \quad m_{F_n}(z) \to m_F(z) \quad \text{for all } z \in S.$$

The fundamental connection to random matrices:

For any Hermitian $n \times n$ matrix A, we let F^A denote the *empirical distribution function* (e.d.f.) of its eigenvalues:

$$F^A(x) = \frac{1}{n}(\text{number of eigenvalues of } A \leq x).$$

Then

$$m_{F^A}(z) = \frac{1}{n}\mathsf{tr}\,(A - zI)^{-1}.$$

So, if we have a sequence $\{A_n\}$ of Hermitian random matrices, to show, with probability one, $F^{A_n} \stackrel{v}{\longrightarrow} F$ for some $F \in \mathcal{M}(\mathbb{R})$, it is equivalent to show for any $z \in \mathbb{C}^+$

$$\frac{1}{n}\mathsf{tr}\,(A_n - zI)^{-1} \to m_F(z) \quad a.s.$$

The main goal of the lectures is to show the importance of the Stieltjes transform to limiting behavior of most of the classes of random matrices introduced earlier. We will begin with an attempt at providing a systematic way to show a.s. convergence of the e.d.f.'s of the eigenvalues of three classes of large dimensional random matrices via the Stieltjes transform approach. Essential properties involved will be emphasized in order to better understand where randomness comes in and where basic properties of matrices are used.

Then it will be shown, via the Stieltjes transform, how the limiting distribution can be numerically constructed, how it can explicitly (mathematically) be derived in some cases, and, in general, how important qualitative information can be inferred. Other results will be reviewed, namely the exact separation properties of eigenvalues, and distributional behavior of linear spectral statistics.

It is hoped that with this knowledge other ensembles can be explored for possible limiting behavior.

Each theorem below corresponds to a matrix ensemble. For each one the random quantities are defined on a common probability space. They all assume:

For $n = 1, 2, \ldots$ $X_n = (X_{ij}^n)$, $n \times N$, $X_{ij}^n \in \mathbb{C}$, i.d. for all n, i, j, independent across i, j for each n, $\mathsf{E}|X_{1\,1}^1 - \mathsf{E}X_{1\,1}^1|^2 = 1$, and $N = N(n)$ with $n/N \to c > 0$ as $n \to \infty$.

THEOREM 1.1 (Marčenko and Pastur (1967), S. and Bai (1995)). *Assume:*

a) $T_n = \text{diag}(t_1^n, \ldots, t_n^n)$, $t_i^n \in \mathbb{R}$, *and the e.d.f. of* $\{t_1^n, \ldots, t_n^n\}$ *converges weakly, with probability one to a nonrandom probability distribution function* H *as* $n \to \infty$.

b) A_n *is a random* $N \times N$ *Hermitian random matrix for which* $F^{A_n} \xrightarrow{v} \mathcal{A}$ *where* $\mathcal{A}$ *is nonrandom (possibly defective).*

c) X_n, T_n, *and* A_n *are independent.*

Let $B_n = A_n + (1/N)X_n^* T_n X_n$. *Then, with probability one* $F^{B_n} \xrightarrow{v} \hat{F}$ *as* $n \to \infty$ *where for each* $z \in \mathbb{C}^+$ $m = m_{\hat{F}}(z)$ *satisfies*

$$m = m_{\mathcal{A}}\left(z - c \int \frac{t}{1 + tm} dH(t)\right). \tag{1.2}$$

It is the only solution to (1.2) with positive imaginary part.

THEOREM 1.2 (Yin (1986), S. (1995)). *Assume:*

T_n $n \times n$ is random Hermitian non-negative definite, independent of X_n with $F^{T_n} \xrightarrow{D} H$ a.s. as $n \to \infty$, H nonrandom.

Let $T_n^{1/2}$ denote any Hermitian square root of T_n, and define $B_n = (1/N)T_n^{1/2}X_nX_n^*T_n^{1/2}$. Then, with probability one $F^{B_n} \xrightarrow{D} F$ as $n \to \infty$ where for each $z \in \mathbb{C}^+$ $m = m_F(z)$ satisfies

$$m = \int \frac{1}{t(1-c-czm)-z} dH(t). \tag{1.3}$$

It is the only solution to (1.3) in the set $\{m \in \mathbb{C} : -(1-c)/z + cm \in \mathbb{C}^+\}$.

THEOREM 1.3 (Dozier and S. a)). *Assume:*

R_n $n \times N$ is random, independent of X_n, with $F^{(1/N)R_n R_n^*} \xrightarrow{D} H$ a.s. as $n \to \infty$, H nonrandom.

Let $B_n = (1/N)(R_n + \sigma X_n)(R_n + \sigma X_n)^*$ where $\sigma > 0$, nonrandom. Then, with probability one $F^{B_n} \xrightarrow{D} F$ as $n \to \infty$ where for each $z \in \mathbb{C}^+$ $m = m_F(z)$ satisfies

$$(1.4). \qquad m = \int \frac{1}{\frac{t}{1+\sigma^2 cm} - (1 + \sigma^2 cm)z + \sigma^2(1-c)} dH(t)$$

It is the only solution to (1.4) in the set $\{m \in \mathbb{C}^+ : \Im(mz) \geq 0\}$.

THEOREM 1.4 (Paul and S., in progress). *Assume:*

S_n *$N \times N$ is random Hermitian with* $F^{S_n} \xrightarrow{D} G$ a.s. *as* $n \to \infty$, *G nonrandom.*

Define $B_n = (1/N)T_n^{1/2} X_n S_n X_n^* T_n^{1/2}$. *Then, with probability one* $F^{B_n} \xrightarrow{D} F$ *as* $n \to \infty$ *where, for each* $z \in \mathbb{C}^+$

$$m = m_F(z) = \int \frac{1}{t \int \frac{s}{1+cse} dG(s) - z} dH(t),$$

where e has positive imaginary part and satisfies

$$e = \int \frac{t}{t \int \frac{s}{1+cse} dG(s) - z} dH(t).$$

It is the only solution with positive imaginary part.

Remark: In Theorem 1.1 if $A_n = 0$ for all n large, then $m_{\mathcal{A}}(z) = -1/z$ and we find that m_F has an inverse

$$z = -\frac{1}{m} + c\int \frac{t}{1+tm} dH(t). \tag{1.5}$$

Since

$$F^{(1/N)X_n^* T_n X_n}(x) = \left(1 - \frac{n}{N}\right) I_{[0,\infty)}(x) + \frac{n}{N} F^{(1/N)T_n^{1/2} X_n X_n^* T_n^{1/2}}(x)$$

we have

(1.6)

$$m_{F^{(1/N)X_n^* T_n X_n}}(z) = -\frac{1-n/N}{z} + \frac{n}{N} m_{F^{(1/N)T_n^{1/2} X_n X_n^* T_n^{1/2}}}(z) \quad z \in \mathbb{C}^+,$$

so we have

$$m_{\hat{F}}(z) = -\frac{1-c}{z} + c m_F(z). \tag{1.7}$$

Using this identity, it is easy to see that (1.3) and (1.5) are equivalent.

2. Why these theorems are true. We begin with three facts which account for most of why the limiting results are true, and the appearance of the limiting equations for the Stieltjes transforms.

LEMMA 2.1 *For $n \times n$ A, $q \in \mathbb{C}^n$, and $t \in \mathbb{C}$ with A and $A + tqq^*$ invertible, we have*

$$q^*(A + tqq^*)^{-1} = \frac{1}{1 + tq^*A^{-1}q} q^*A^{-1}$$

(since $q^*A^{-1}(A + tqq^*) = (1 + tq^*A^{-1}q)q^*$).

COROLLARY 2.1 *For $q = a + b$, $t = 1$ we have*

$$a^*(A+(a+b)(a+b)^*)^{-1} = a^*A^{-1} - \frac{a^*A^{-1}(a+b)}{1 + (a+b)^*A^{-1}(a+b)}(a+b)^*A^{-1}$$

$$= \frac{1 + b^*A^{-1}(a+b)}{1 + (a+b)^*A^{-1}(a+b)} a^*A^{-1} - \frac{a^*A^{-1}(a+b)}{1 + (a+b)^*A^{-1}(a+b)} b^*A^{-1}.$$

Proof: Using Lemma 2.1 we have

$$\begin{aligned}(A+(a+b)(a+b)^*)^{-1}-A^{-1} &= -(A+(a+b)(a+b)^*)^{-1}(a+b)(a+b)^*A^{-1} \\ &= -\frac{1}{1+(a+b)^*A^{-1}(a+b)}A^{-1}(a+b)(a+b)^*A^{-1}\end{aligned}$$

Multiplying both sides on the left by a^* gives the result.

LEMMA 2.2 *For $n \times n$ A and B, with B Hermitian, $z \in \mathbb{C}^+$, $t \in \mathbb{R}$, and $q \in \mathbb{C}^n$, we have*

$$|\mathsf{tr}\,[(B-zI)^{-1}-(B+tqq^*-zI)^{-1}]A| = \left|t\frac{q^*(B-zI)^{-1}A((B-zI)^{-1}q}{1+tq^*(B-zI)^{-1}q}\right| \le \frac{\|A\|}{\Im z}.$$

Proof. The identity follows from Lemma 2.1. We have

$$\left|t\frac{q^*(B-zI)^{-1}A((B-zI)^{-1}q}{1+tq^*(B-zI)^{-1}q}\right| \le \|A\||t|\|\frac{\|(B-zI)^{-1}q\|^2}{|1+tq^*(B-zI)^{-1}q|}.$$

Write $B = \sum_i \lambda_i e_i e_i^*$, its spectral decomposition. Then

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

and

$$|1 + tq^*(B - zI)^{-1} q| \geq |t| \Im(q^*(B - zI)^{-1} q) = |t| \Im z \sum_i \frac{|e_i^* q|^2}{|\lambda_i - z|^2}.$$

Lemma 2.3. *For* $X = (X_1, \ldots, X_n)^T$ *i.i.d. standardized entries,* C $n \times n$, *we have for any* $p \geq 2$

$$\mathsf{E}|X^* C X - \mathsf{tr}\, C|^p \leq K_p \Big(\big(\mathsf{E}|X_1|^4 \mathsf{tr}\, CC^* \big)^{p/2} + \mathsf{E}|X_1|^{2p} \mathsf{tr}\, (CC^*)^{p/2} \Big)$$

where the constant K_p *does not depend on* n, C, *nor on the distribution of* X_1. (Proof given in Bai and S. (1998).)

An example of how the last two lemmas are used is given. Suppose $C = C_n = (A - zI)^{-1}$, where for each n, $A = A_n$ is Hermitian, independent of X, and $z = x + iv \in \mathbb{C}^+$. Then the spectral norm $\|C\|$ of C satisfies $\|C\| \leq 1/v$. Let $q = (1/\sqrt{n})X$. Then for any $p \geq 2$ we have by Lemma 2.3

$$\mathsf{E}|q^*Cq - (1/n)\mathsf{tr}\, C|^p \leq \frac{K_p}{(nv)^p}((\mathsf{E}|X_1|^4 n)^{p/2} + \mathsf{E}|X_1|^{2p}) \leq \frac{K}{n^{p/2}},$$

where K depends also on v and the fourth and $2p$ moments of X. If these moments grow slowly, say, like powers of $\ln n$ as $n \to \infty$, then, considering $p > 2$ (keeping v fixed) we would have for any positive ϵ

$$\mathsf{P}(|q^*Cq - (1/n)\mathsf{tr}\, C| > \epsilon) \leq \frac{\mathsf{E}|q^*Cq - (1/n)\mathsf{tr}\, C|^p}{\epsilon^p},$$

which is summable. Therefore, noticing that $(1/n)\mathsf{tr}\, C = m_A(z)$, we have

$$q^*Cq - m_A(z) \xrightarrow{a.s.} 0.$$

Lemma 2.2 gives us for any sequence $\{t_n\}$ of real numbers

$$q^*Cq - m_{A+t_n qq^*}(z) \xrightarrow{a.s.} 0.$$

Moreover, if for each n we have $O(n)$ of these quantities, say, q_i, A_i, $t_{i,n}$, we get from Boole's inequality for any positive ϵ

$$\mathsf{P}(\max_i |q_i^*(A_i - zI)q_i^{-1} - m_{A_i}(z)| > \epsilon) \leq K\frac{n}{n^{p/2}},$$

which is summable when $p > 4$. Therefore

$$\max_i |q_i^*(A_i - zI)^{-1}q_i - m_{A_i+t_n q_i q_i^*}(z)| \xrightarrow{a.s.} 0.$$

It will be clear where these arguments are needed in the proofs, along with the identities in the first lemma. An important step is the truncation and centralization of the elements of X_n, and truncation of the eigenvalues of T_n in Theorem 1.2 (not needed in Theorem 1.1) and $(1/N)R_nR_n^*$ in Theorem 1.3, all at a rate slower than n ($a \ln n$ for some positive a is sufficient). These steps will be outlined later. We are at this stage able to go through algebraic manipulations, keeping in mind the above three lemmas, and intuitively derive the equations appearing in each of the three theorems. At the same time we can see what technical details need to be worked out.

Before continuing, two more basic properties of matrices is included here.

Lemma 2.4 *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*

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

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

Consider first the B_n in Theorem 1.1. Let q_i denote $1/\sqrt{N}$ times the i^{th} column of X_n^*. Then

$$(1/N) X_n^* T_n X_n = \sum_{i=1}^{n} t_i q_i q_i^*.$$

Let $B_{(i)} = B_n - t_i q_i q_i^*$. For any $z \in \mathbb{C}^+$ and $x \in \mathbb{C}$ we write

$$B_n - zI = A_n - (z - x)I + (1/N) X_n^* T_n X_n - xI.$$

Taking inverses we have

$$(A_n - (z-x)I)^{-1}$$
$$= (B_n - zI)^{-1} + (A_n - (z-x)I)^{-1}((1/N)X_n^* T_n X_n - xI)(B_n - zI)^{-1}.$$

Dividing by N, taking traces and using Lemma 2.1 we find

$$m_{F^{A_n}}(z-x) - m_{F^{B_n}}(z) = (1/N)\mathrm{tr}\,(A_n - (z-x)I)^{-1}\left(\sum_{i=1}^{n} t_i q_i q_i^* - xI\right)(B_n - zI)^{-1}$$

$$= (1/N)\sum_{i=1}^{n} \frac{t_i q_i^* (B_{(i)} - zI)^{-1}(A_n - (z-x)I)^{-1} q_i}{1 + t_i q_i^* (B_{(i)} - zI)^{-1} q_i}$$
$$- x(1/N)\mathrm{tr}\,(B_n - zI)^{-1}(A_n - (z-x)I)^{-1}.$$

Notice that Lemmas 2.2,2.3 give us

$$q_i^*(B_{(i)} - zI)^{-1} q_i \approx m_{F^{B_n}}(z),$$

and when x and q_i are independent, we get

$$q_i^*(B_{(i)}-zI)^{-1}(A_n-(z-x)I)^{-1}q_i \approx (1/N)\mathsf{tr}\,(B_n-zI)^{-1}(A_n-(z-x)I)^{-1}.$$

Letting

$$x = x_n = (1/N)\sum_{i=1}^{n} \frac{t_i}{1+t_i m_{F^{B_n}}(z)}$$

we have

$$m_{F^{A_n}}(z-x_n) - m_{F^{B_n}}(z) = (1/N)\sum_{i=1}^{n} \frac{t_i}{1+t_i m_{F^{B_n}}(z)} d_i$$

where

$$d_i = \frac{1+t_i m_{F^{B_n}}(z)}{1+t_i q_i^*(B_{(i)}-zI)^{-1}q_i} q_i^*(B_{(i)}-zI)^{-1}(A_n-(z-x_n)I)^{-1}q_i - (1/N)\mathsf{tr}\,(B_n-zI)^{-1}(A_n-(z-x_n)I)^{-1}.$$

In order to use Lemma 2.3, for each i, we subtract and add expressions where x_n is replaced by

$$x_{(i)} = (1/N)\sum_{j=1}^{n} \frac{t_j}{1 + t_j m_{F^{B_{(i)}}}(z)}.$$

An outline of the remainder of the proof is given. It is easy to argue that if $\mathcal{A}$ is the zero measure on $\mathbb{R}$ (that is, almost surely, only $o(N)$ eigenvalues of A_n remain bounded), then the Stieltjes transforms of F^{A_n} and F^{B_n} converge a.s. to zero, the limits obviously satisfying (1.2) . So we assume $\mathcal{A}$ is not the zero measure. One can then show

$$\delta = \inf_n \Im(m_{F^{B_n}}(z))$$

is positive almost surely.

Using Lemma 2.3 ($p > 4$) and the fact that all matrix inverses encountered are bounded in spectral norm by $1/\Im z$ we have from standard arguments using Boole's and Chebyshev's inequalities, almost surely

$$(2.1) \qquad \max_{i \le n} \max[|\,\|q_i\|^2 - 1|, |q_i^*(B_{(i)} - zI)^{-1} q_i - m_{F^{B_{(i)}}}(z)|,$$

$$|q_i^*(B_{(i)}-zI)^{-1}(A_n-(z-x_{(i)})I)^{-1}q_i-(1/N)\mathsf{tr}\,(B_{(i)}-zI)^{-1}(A_n-(z-x_{(i)})I)^{-1}|]$$

$$\to 0 \quad \text{as } n\to\infty.$$

Consider now a realization for which (2.1) holds, $\delta > 0$, $F^{T_n} \xrightarrow{D} H$, and $F^{A_n} \xrightarrow{v} \mathcal{A}$. From Lemma 2.2 and (2.1) we have

(2.2)

$$\max_{i\le n}\max[|m_{F^{B_n}}(z)-m_{F^{B_{(i)}}}(z)|,|m_{F^{B_n}}(z)-q_i^*(B_{(i)}-zI)^{-1}q_i|]\to 0,$$

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

$$\max_{i\le n} \max \left[\left| \frac{1+t_i m_{F^{B_n}}(z)}{1+t_i q_i^*(B_{(i)}-zI)^{-1} q_i} - 1 \right|, |x - x_{(i)}| \right] \to 0. \tag{2.3}$$

Therefore, from Lemmas 2.2,2.4, and (2.1) -(2.3), we get $\max_{i\le n} d_i \to 0$, and since

$$\left| \frac{t_i}{1+t_i m_{F^{B_n}}(z)} \right| \le \frac{1}{\delta},$$

we conclude that

$$m_{A_n}(z - x_n) - m_{B_n}(z) \to 0.$$

Consider a subsequence $\{n_j\}$ on which $m_{F^{B_{n_j}}}(z)$ converges to a number m. It follows that

$$x_{n_j} \to c \int \frac{t}{1+tm} dH(t).$$

Therefore, m satisfies (1.1). Uniqueness (to be discussed later) gives us, for this realization $m_{F^{B_n}}(z) \to m$. This event occurs with probability one.

3. The other equations. Let us now derive the equation for the matrix $B_n = (1/N)T_n^{1/2}X_nX_n^*T_n^{1/2}$, after the truncation steps have been taken. Let $c_n = n/N$, $q_j = (1/\sqrt{n})X_{\cdot j}$ (the j^{th} column of X_n), $r_j = (1/\sqrt{N})T_n^{1/2}X_{\cdot j}$, and $B_{(j)} = B_n - r_jr_j^*$. Fix $z \in \mathbb{C}^+$ and let $m_n(z) = m_{F^{B_n}}(z)$, $\mathbf{m}_n(z) = m_{F^{(1/N)X_n^*T_nX_n}}(z)$. By (1.6) we have

$$\mathbf{m}_n(z) = -\frac{1-c_n}{z} + c_n m_n. \tag{3.1}$$

We first derive an identity for $\mathbf{m}_n(z)$. Write

$$B_n - zI + zI = \sum_{j=1}^{N} r_jr_j^*.$$

Taking the inverse of $B_n - zI$ on the right on both sides and using Lemma 2.1 we find

$$I + z(B_n - zI)^{-1} = \sum_{j=1}^{N} \frac{1}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j} r_j r_j^* (B_{(j)} - zI)^{-1}.$$

Taking the trace on both sides and dividing by N we have

$$c_n + z c_n m_n = \frac{1}{N} \sum_{j=1}^{N} \frac{r_j^*(B_{(j)} - zI)^{-1} r_j}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j} = 1 - \frac{1}{N} \sum_{j=1}^{N} \frac{1}{1 + r_j^*(B_{(j)} - zI)^{-1} r_j}.$$

Therefore

$$\mathbf{m}_n(z) = -\frac{1}{N} \sum_{j=1}^{N} \frac{1}{z(1 + r_j^*(B_{(j)} - zI)^{-1} r_j)}. \tag{3.2}$$

Write $B_n - zI - \big(-z\mathbf{m}_n(z) T_n - zI\big) = \sum_{j=1}^{N} r_j r_j^* - (-z\mathbf{m}_n(z)) T_n$.

Taking inverses and using Lemma 2.1, (3.2) we have

$$(-z\mathbf{m}_n(z)T_n - zI)^{-1} - (B_n - zI)^{-1}$$
$$= (-z\mathbf{m}_n(z)T_n - zI)^{-1}\left[\sum_{j=1}^{N} r_j r_j^* - (-z\mathbf{m}_n(z))T_n\right](B_n - zI)^{-1}$$

$$= \sum_{j=1}^{N} \frac{-1}{z(1+r_j^*(B_{(j)}-zI)^{-1}r_j)}\Big[(\mathbf{m}_n(z)T_n + I)^{-1} r_j r_j^*(B_{(j)} - zI)^{-1}$$
$$- (1/N)(\mathbf{m}_n(z)T_n + I)^{-1}T_n(B_n - zI)^{-1}\Big].$$

Taking the trace and dividing by n we find

$$(1/n)\mathsf{tr}\,(-z\mathbf{m}_n(z)T_n - zI)^{-1} - m_n(z) = \frac{1}{N}\sum_{j=1}^{N} \frac{-1}{z(1 + r_j^*(B_{(j)} - zI)^{-1}r_j)} d_j$$

where

$$d_j = q_j^* T_n^{1/2}(B_{(j)} - zI)^{-1}(\mathbf{m}_n(z)T_n + I)^{-1}T_n^{1/2}q_j$$
$$- (1/n)\mathsf{tr}\,(\mathbf{m}_n(z)T_n + I)^{-1}T_n(B_n - zI)^{-1}.$$

The derivation for Theorem 1.3 will proceed in a constructive way. Here we let x_j and r_j denote, respectively, the j^{th} columns of X_n and R_n (after truncation). As before $m_n = m_{F^{B_n}}$, and let

$$\mathbf{m}_n(z) = m_{F^{(1/N)(R_n+\sigma X_n)^*(R_n+\sigma X_n)}}(z).$$

We have again the relationship (3.1) . Notice then equation (1.4) can be written

$$m = \int \frac{1}{\frac{t}{1+\sigma^2 cm} - \sigma^2 z\mathbf{m} - z} dH(t) \tag{3.3}$$

where

$$\mathbf{m} = -\frac{1-c}{z} + cm.$$

Let $B_{(j)} = B_n - (1/N)(r_j + \sigma x_j)(r_j + \sigma x_j)^*$. Then, as in (3.2) we have

(3.4)

$$\mathbf{m}_n(z) = -\frac{1}{N}\sum_{j=1}^{N}\frac{1}{z(1+(1/N)(r_j+\sigma x_j)^*(B_{(j)}-zI)^{-1}(r_j+\sigma x_j))}.$$

Pick $z \in \mathbb{C}^+$. For any $n \times n$ Y_n we write

$$B_n - zI - (Y_n - zI) = \frac{1}{N}\sum_{j=1}^{N}(r_j + \sigma x_j)(r_j + \sigma x_j)^* - Y_n.$$

Taking inverses, dividing by n and using Lemma 2.1 we get

$$(1/n)\mathsf{tr}\,(Y_n - zI)^{-1} - m_n(z)$$

$$\frac{1}{N}\sum_{j=1}^{N}\frac{(1/n)(r_j+\sigma x_j)^*(B_{(j)}-zI)^{-1}(Y_n-zI)^{-1}(r_j+\sigma x_j)}{1+(1/N)(r_j+\sigma x_j)^*(B_{(j)}-zI)^{-1}(r_j+\sigma x_j)}$$

$$-\,(1/n)\mathsf{tr}\,(Y_n - zI)^{-1}Y_n(B_n - zI)^{-1}.$$

The goal is to determine Y_n so that this quantity goes to zero. Notice first that

$$(1/n)x_j^*(B_{(j)} - zI)^{-1}(Y_n - zI)^{-1}x_j \approx (1/n)\mathsf{tr}\,(B_n - zI)^{-1}(Y_n - zI)^{-1},$$

so from (3.4) we see that Y_n should have a term

$$-\sigma^2 z\mathbf{m}_n(z)I.$$

Since for any $n \times n$ C bounded in norm

$$|(1/n)x_j^*Cr_j|^2 = (1/n^2)x_j^*Cr_jr_j^*C^*x_j$$

we have from Lemma 2.3

$$|(1/n)x_j^*Cr_j|^2 \approx (1/n^2)\mathsf{tr}\,Cr_jr_j^*C^* = (1/n^2)r_j^*C^*Cr_j = o(1) \tag{3.5}$$

(from truncation $(1/N)\|r_j\|^2 \leq \ln n$), so the cross terms are negligible.

This leaves us $(1/n)r_j^*(B_{(j)}-zI)^{-1}(Y_n-zI)^{-1}r_j$. Recall Corollary 2.1

$$a^*(A+(a+b)(a+b)^*)^{-1}$$

$$= \frac{1+b^*A^{-1}(a+b)}{1+(a+b)^*A^{-1}(a+b)}a^*A^{-1} - \frac{a^*A^{-1}(a+b)}{1+(a+b)^*A^{-1}(a+b)}b^*A^{-1}.$$

Identify a with $(1/\sqrt{N})r_j$, b with $(1/\sqrt{N})\sigma x_j$, and A with $B_{(j)} - zI$. Using Lemmas 2.2, 2.3 and (3.5), we have

$$(1/n)r_j^*(B_n - zI)^{-1}(Y_n - zI)^{-1}r_j$$

$$\approx \frac{1+\sigma^2 c_n m_n(z)}{1+\frac{1}{N}(r_j+\sigma x_j)^*(B_{(j)}-zI)^{-1}(r_j+\sigma x_j)}(1/n)r_j^*(B_{(j)}-zI)^{-1}(Y_n-zI)^{-1}r_j.$$

Therefore

$$\frac{1}{N}\sum_{j=1}^{N}\frac{(1/n)r_j^*(B_{(j)}-zI)^{-1}(Y_n-zI)^{-1}r_j}{1+\frac{1}{N}(r_j+\sigma x_j)^*(B_{(j)}-zI)^{-1}(r_j+\sigma x_j)}$$

$$\approx \frac{1}{N}\sum_{j=1}^{N}\frac{(1/n)r_j^*(B_n-zI)^{-1}(Y_n-zI)^{-1}r_j}{1+\sigma^2 c_n m_n(z)}$$

$$= (1/n)\frac{1}{1+\sigma^2 c_n m_n(z)}\mathsf{tr}\,(1/N)R_nR_n^*(B_n-zI)^{-1}(Y_n-zI)^{-1}.$$

So we should take

$$Y_n=\frac{1}{1+\sigma^2 c_n m_n(z)}(1/N)R_nR_n^*-\sigma^2 z\mathbf{m}_n(z)I.$$

Then $(1/n)\mathsf{tr}\,(Y_n-zI)^{-1}$ will approach the right hand side of (3.3).

4. Proof of uniqueness of (1.2). For $m \in \mathbb{C}^+$ satisfying (1.2) with $z \in \mathbb{C}^+$ we have

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

$$= \int \frac{1}{\tau - \Re\left(z - c\int \frac{t}{1+tm} dH(t)\right) - i\left(\Im z + c\int \frac{t^2 \Im m}{|1+tm|^2} dH(t)\right)} d\mathcal{A}(\tau)$$

Therefore

(4.1)

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

Suppose $\mathbf{m} \in \mathbb{C}^+$ also satisfies (1.2). Then

(4.2)

$$m - \mathbf{m} = c\int \frac{\left[\int \frac{t}{1+t\mathbf{m}} - \frac{t}{1+tm}\right] dH(t)}{\left(\tau - z + c\int \frac{t}{1+tm} dH(t)\right)\left(\tau - z + c\int \frac{t}{1+t\mathbf{m}} dH(t)\right)} d\mathcal{A}(\tau)$$

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

Using Cauchy-Schwarz and (4.1) we have

$$\Bigg|c\int\frac{t^2}{(1+tm)(1+t\mathbf{m})}dH(t)$$
$$\times\int\frac{1}{\left(\tau-z+c\int\frac{t}{1+tm}dH(t)\right)\left(\tau-z+c\int\frac{t}{1+t\mathbf{m}}dH(t)\right)}d\mathcal{A}(\tau)\Bigg|$$

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

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

$$= \left(c\int \frac{t^2}{|1+tm|^2} dH(t) \frac{\Im m}{\left(\Im z + c\int \frac{t^2 \Im m}{|1+tm|^2} dH(t)\right)} \right)^{1/2}$$

$$\times \left(c\int \frac{t^2}{|1+t\mathbf{m}|^2} dH(t) \frac{\Im \mathbf{m}}{\left(\Im z + c\int \frac{t^2 \Im \mathbf{m}}{|1+t\mathbf{m}|^2} dH(t)\right)} \right)^{1/2} < 1.$$

Therefore, from (4.2) we must have $m = \mathbf{m}$.

5. Truncation and Centralization. We outline here the steps taken to enable us to assume in the proof of Theorem 1.1, for each n, the X_{ij}'s are bounded by a multiple of $\ln n$. The following lemmas are needed.

LEMMA 5.1. *Let* $X_1, \ldots, X_n$ *be i.i.d. Bernoulli with* $p = \mathsf{P}(X_1 = 1) < 1/2$. *Then for any* $\epsilon > 0$ *such that* $p + \epsilon \leq 1/2$ *we have*

$$\mathsf{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - p \geq \epsilon\right) \leq e^{-\frac{n\epsilon^2}{2(p+\epsilon)}}.$$

LEMMA 5.2. *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)
$$\|F^{A+Q^*TQ} - F^{A+\overline{Q}^*T\overline{Q}}\| \leq \frac{2}{N}\text{rank}(Q - \overline{Q})$$

and

b)
$$\|F^{A+Q^*TQ} - F^{A+Q^*\overline{T}Q}\| \leq \frac{1}{N}\text{rank}(T - \overline{T}).$$

LEMMA 5.3. *For rectangular A, rank$(A) \leq$ the number of nonzero entries of A.*

LEMMA 5.4 *For Hermitian $N \times N$ matrices A, B*

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

LEMMA 5.5 *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 $[a-\frac{1}{m}, a]$, $[b, b+\frac{1}{m}]$. Then*

a) *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.

b) *For* $F_N, G_N \in \mathcal{M}(\mathbb{R})$

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

c) *For empirical distribution functions* F, G *on the (respective) sets* $\{x_1, \ldots, x_N\}, \{y_1, \ldots, y_N\}$

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

Let $p_n = \mathsf{P}(|X_{11}| \geq \sqrt{n})$. Since the second moment of X_{11} is finite we have

$$np_n = o(1). \tag{5.1}$$

Let $\widehat{X}_{ij} = X_{ij} I_{(|X_{ij}| < \sqrt{n})}$ and $\widehat{B}_n = A_n + (1/N)\widehat{X}_n^* T_n \widehat{X}_n$, where $\widehat{X} =$

(X_{ij}). Then from Lemmas 5.2 a) 5.3, for any positive ϵ

$$\mathsf{P}(\|F^{B_n} - F^{\widehat{B}_n}\| \ge \epsilon) \le \mathsf{P}\left(\frac{2}{N}\sum_{ij} I_{(|X_{ij}|\ge\sqrt{n})} \ge \epsilon\right)$$
$$= \mathsf{P}\left(\frac{1}{Nn}\sum_{ij} I_{(|X_{ij}|\ge\sqrt{n})} - p_n \ge \frac{\epsilon}{2n} - p_n\right).$$

Then by Lemma 5.1, for all n large

$$\mathsf{P}(\|F^{B_n} - F^{\widehat{B}_n}\| \ge \epsilon) \le e^{-\frac{N\epsilon}{16}},$$

which is summable. Therefore

$$\|F^{B_n} - F^{\widehat{B}_n}\| \xrightarrow{a.s.} 0.$$

Let $\widetilde{B}_n = A_n + (1/N)\widetilde{X}_n T_n \widetilde{X}_n^*$ where $\widetilde{X}_n = \widehat{X}_n - \mathsf{E}\widehat{X}_n$. Since $\text{rank}(\mathsf{E}\widehat{X}_n) \le 1$, we have from Lemma 5.2 a)

$$\|F^{\widehat{B}_n} - F^{\widetilde{B}_n}\| \longrightarrow 0.$$

For $\alpha > 0$ define $T_\alpha = \mathrm{diag}(t_1^n I_{(|t_1^n|\le\alpha)}, \dots, t_n^n I_{(|t_n^n|\le\alpha)})$, and let Q be any $n \times N$ matrix. If α and $-\alpha$ are continuity points of H, we have by Lemma 5.2 b)

$$\|F^{A_n+Q^*T_nQ} - F^{A_n+Q^*T_\alpha Q}\| \\ \le \frac{1}{N}\mathrm{rank}(T_n - T_\alpha) = \frac{1}{N}\sum_{i=1}^n I_{(|t_i^n|>\alpha)} \xrightarrow{a.s.} cH\{[-\alpha,\alpha]^c\}.$$

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

$$\|F^{A_n+Q^*T_nQ} - F^{A_n+Q^*T_\alpha Q}\| \xrightarrow{a.s.} 0.$$

Let $\overline{X}_{ij} = \widetilde{X}_{ij} I_{(|X_{ij}|<\ln n)} - \mathsf{E}\widetilde{X}_{ij} I_{(|X_{ij}|<\ln n)}$, $\overline{X}_n = ((1/\sqrt{N})\overline{X}_{ij})$, $\overline{\overline{X}}_{ij} = \widetilde{X}_{ij} - \overline{X}_{ij}$, and $\overline{\overline{X}}_n = ((1/\sqrt{N})\overline{\overline{X}}_{ij})$. Then, from Lemmas 5.5 c) and 5.4 and simple applications of Cauchy-Schwarz we have

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

$$\leq \frac{1}{N}[\mathsf{tr}\,(\overline{\overline{X}}_n T_\alpha \overline{\overline{X}}_n^*)^2 + 4\mathsf{tr}\,(\overline{\overline{X}}_n T_\alpha \overline{X}_n^* \overline{X}_n T_\alpha \overline{\overline{X}}_n^*)$$

$$+ 4(\mathsf{tr}\,(\overline{\overline{X}}_n T_\alpha \overline{X}_n^* \overline{X}_n T_\alpha \overline{\overline{X}}_n^*)\mathsf{tr}\,(\overline{\overline{X}}_n T_\alpha \overline{\overline{X}}_n^*)^2)^{1/2}].$$

We have

$$\mathsf{tr}\,(\overline{\overline{X}}_n T_\alpha \overline{\overline{X}}_n^*)^2 \leq \alpha^2 \mathsf{tr}\,(\overline{\overline{X}}\,\overline{\overline{X}}^*)^2$$

and

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

Therefore, to verify

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

it is sufficient to find a sequence $\{\alpha_n\}$ increasing to ∞ so that

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

The details are omitted.

Notice the matrix $\text{diag}(\mathsf{E}|\overline{X}_{1\,1}|^2 t_1^n, \ldots, \mathsf{E}|\overline{X}_{1\,1}|^2 t_n^n)$ also satisfies assumption a) of Theorem 1.1. Just substitute this matrix for T_n, and replace $\overline{X}_n$ by $(1/\sqrt{\mathsf{E}|\overline{X}_{1\,1}|^2})\overline{X}_n$. Therefore we may assume

1) X_{ij} are i.i.d. for fixed n,
2) $|X_{11}| \leq a \ln n$ for some positive a,
3) $\mathsf{E}X_{11} = 0$, $\mathsf{E}|X_{11}|^2 = 1$.

6. The limiting distributions. The Stieltjes transform provides a great deal of information to the nature of the limiting distribution $\hat{F}$ when $A_n = 0$ in Theorem 1.1, and F in Theorems 1.2, 1.3. For the first two

$$z = -\frac{1}{\mathbf{m}} + c\int \frac{t}{1+t\mathbf{m}} dH(t)$$

is the inverse of $\mathbf{m} = m_{\hat{F}}(z)$, the limiting Stieltjes transform of $F^{(1/N)X_n^* T_n X_n}$. Recall, when T_n is nonnegative definite, the relationships between F, the limit of $F^{(1/N)T_n^{1/2} X_n X_n^* T_n^{1/2}}$ and $\hat{F}$

$$\hat{F}(x) = 1 - cI_{[0,\infty)}(x) + cF(x),$$

and m_F and $m_{\hat{F}}$

$$m_{\hat{F}}(z) = -\frac{1-c}{z} + cm_F(z).$$

Based solely on the inverse of $m_{\hat{F}}$ the following is shown in S. and Choi (1995):

1. For all $x \in \mathbb{R}$, $x \neq 0$

$$\lim_{z \in \mathbb{C}^+ \to x} m_{\hat{F}}(z) \equiv m_0(x)$$

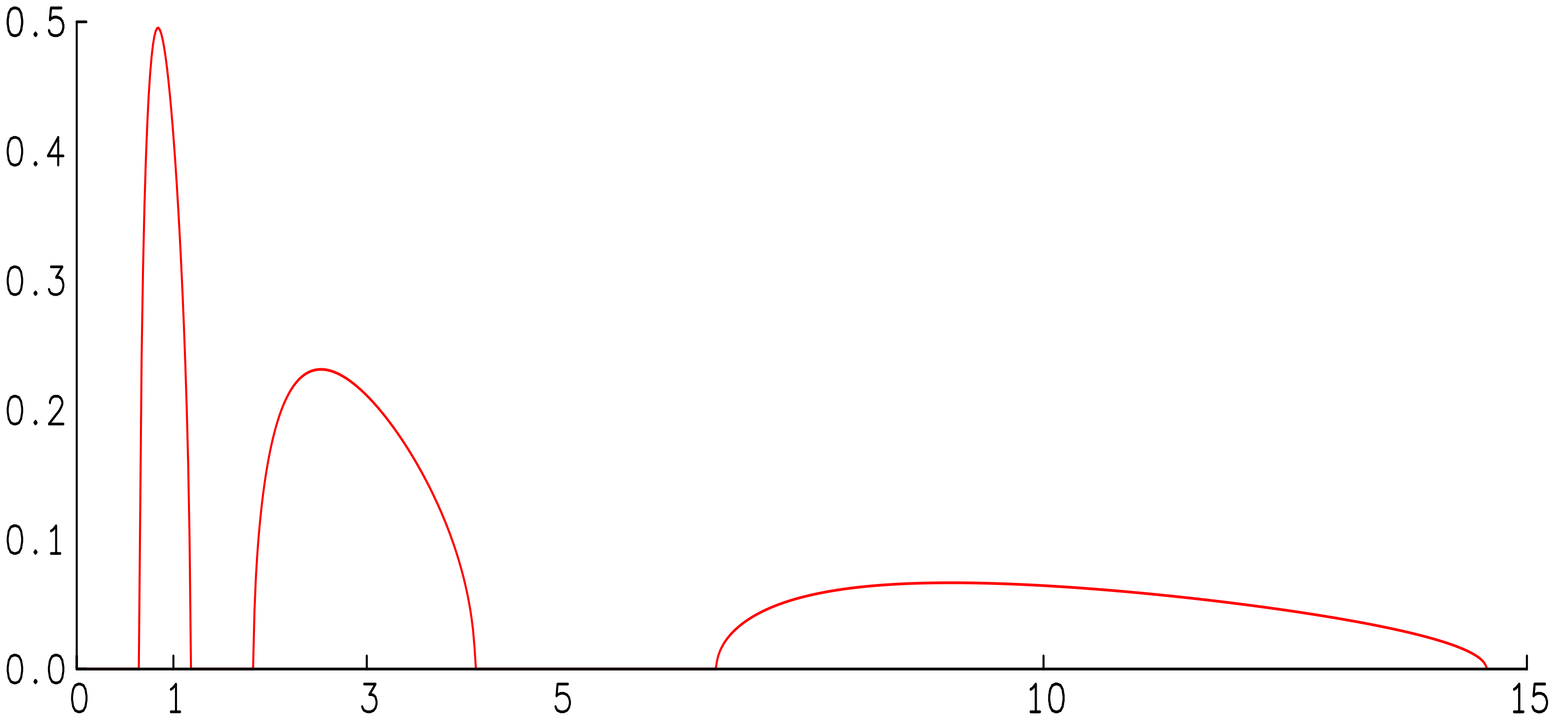
exists. The function m_0 is continuous on $\mathbb{R} - \{0\}$. Consequently, by property 5. of Stieltjes transforms, $\hat{F}$ has a continuous derivative $\hat{f}$ on $\mathbb{R} - \{0\}$ given by $\hat{f}(x) = \frac{1}{\pi} \Im m_0(x)$ (F subsequently has derivative $f = \frac{1}{c}\hat{f}$). The density $\hat{f}$ is analytic (possess a power series expansion) for every $x \neq 0$ for which $\hat{f}(x) > 0$. Moreover, for these x, $\pi \hat{f}(x)$ is the imaginary part of the unique $\mathbf{m} \in \mathbb{C}^+$ satisfying

$$x = -\frac{1}{\mathbf{m}} + c \int \frac{t}{1 + t\mathbf{m}} dH(t).$$

2. Let $x_{\hat{F}}$ denote the above function of $\mathbf{m}$. It is defined and analytic on $B \equiv \{\mathbf{m} \in \mathbb{R} : \mathbf{m} \neq 0, -\mathbf{m}^{-1} \in S_H^c\}$ (S_G^c denoting the complement of the support of distribution G). Then if $x \in S_{\hat{F}}^c$ we have $\mathbf{m} = m_0(x) \in B$ and $x'_{\hat{F}}(\mathbf{m}) > 0$. Conversely, if $\mathbf{m} \in B$ and $x = x'_{\hat{F}}(\mathbf{m}) > 0$, then $x \in S_{\hat{F}}^c$.

We see then a systematic way of determining the support of $\hat{F}$: Plot $x_{\hat{F}}(\mathbf{m})$ for $\mathbf{m} \in B$. Remove all intervals on the vertical axis corresponding to places where $x_{\hat{F}}$ is increasing. What remains is $S_{\hat{F}}$, the support of $\hat{F}$.

Let us look at an example where H places mass at 1, 3, and 10, with respective probabilities .2, .4, and .4, and $c = .1$.



(a)

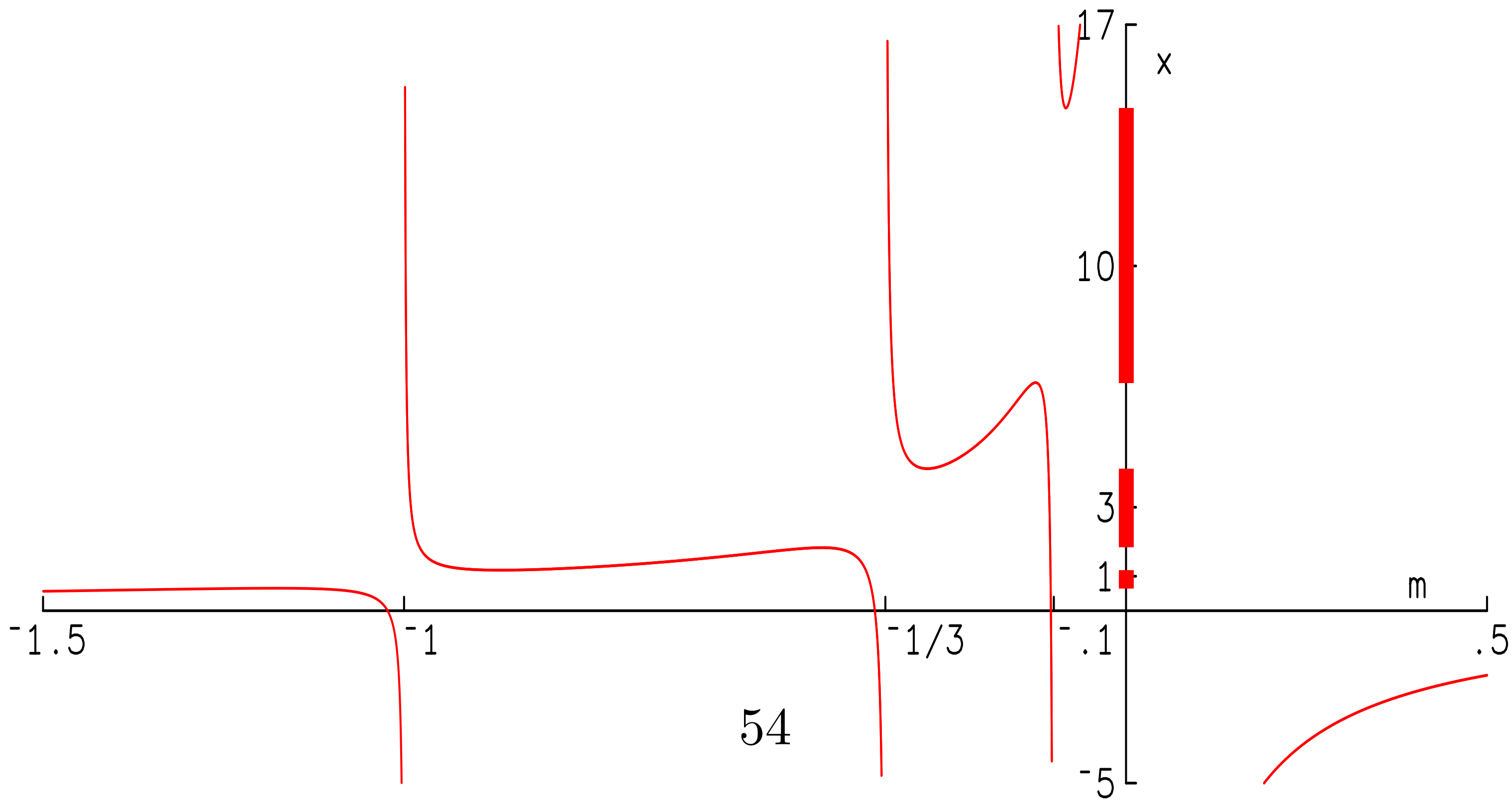


Figure (b) is the graph of

$$x_{\hat{F}}(\mathbf{m}) = -\frac{1}{\mathbf{m}} + .1\left(.2\frac{1}{1+\mathbf{m}} + .4\frac{3}{1+3\mathbf{m}} + .4\frac{10}{1+10\mathbf{m}}\right).$$

We see the support boundaries occur at relative extreme values. These values were estimated and for values of $x \in S_{\hat{F}}$, $f(x) = \frac{1}{c\pi}\Im m_0(x)$ was computed using Newton's method on $x = x_{\hat{F}}(\mathbf{m})$, resulting in figure (a).

It is possible for a support boundary to occur at a boundary of the support of B, which would only happen for a nondiscrete H. However, we have

3. Suppose support boundary a is such that $m_{\hat{F}}(a) \in B$, and is a left-endpoint in the support of $\hat{F}$. Then for $x > a$ and near a

$$f(x) = \left(\int_a^x g(t)dt\right)^{1/2}.$$

where $g(a) > 0$ (analogous statement holds for a a right-endpoint in the support of $\hat{F}$). Thus, near support boundaries, f and the square root function share common features, as can be seen in figure (a).

It is remarked here that similar results have been obtained for the matrices in Theorem 1.3. See Dozier and S. b).

Explicit solutions can be derived in a few cases. Consider the Mařcenko-Pastur distribution, where $T_n = I$. Then $\mathbf{m} = m_0(x)$ solves

$$x = -\frac{1}{\mathbf{m}} + c\frac{1}{1+\mathbf{m}},$$

resulting in the quadratic equation

$$x\mathbf{m}^2 + \mathbf{m}(x+1-c) + 1 = 0$$

with solution

$$m = \frac{-(x+1-c) \pm \sqrt{(x+1-c)^2 - 4x}}{2x}$$
$$= \frac{-(x+1-c) \pm \sqrt{x^2 - 2x(1+c) + (1-c)^2}}{2x}$$

$$= \frac{-(x+1-c) \pm \sqrt{(x-(1-\sqrt{c})^2)(x-(1+\sqrt{c})^2)}}{2x}$$

We see the imaginary part of m is zero when x lies outside the interval $[(1-\sqrt{c})^2, (1+\sqrt{c})^2]$, and we conclude that

$$f(x) = \begin{cases} \frac{\sqrt{(x-(1-\sqrt{c})^2)((1+\sqrt{c})^2-x)}}{2\pi c x} & x \in ((1-\sqrt{c})^2, (1+\sqrt{c})^2) \\ 0 & \text{otherwise} \end{cases}.$$

The Stieltjes transform in the multivariate F matrix case, that is, when $T_n = ((1/N')\underline{X}_n\underline{X}_n^*)^{-1}$, $\underline{X}_n$ $n \times N'$ containing i.i.d. standardized entries, $n/N' \to c' \in (0,1)$, also satisfies a quadratic equation. Indeed, H now is the distribution of the reciprocal of a Marčenko-Pastur distributed random variable which we'll denote by $X_{c'}$, the Stieltjes transform of its distribution denoted by $m_{X_{c'}}$. We have

$$x = -\frac{1}{\mathbf{m}} + c\mathsf{E}\left(\frac{\frac{1}{X_{c'}}}{1+\frac{1}{X_{c'}}\mathbf{m}}\right) = -\frac{1}{\mathbf{m}} + c\mathsf{E}\left(\frac{1}{X_{c'}+\mathbf{m}}\right)$$

$$= -\frac{1}{\mathbf{m}} + cm_{X_{c'}}(-\mathbf{m}).$$

From above we have

$$m_{X_{c'}}(z) = \frac{1-c'}{c'z} + \frac{-(z+1-c) + \sqrt{(z+1-c)^2 - 4z}}{2zc'}$$

$$= \frac{-z+1-c' + \sqrt{(z+1-c')^2 - 4z}}{2zc'}$$

(the square root defined so that the expression is a Stieltjes transform) so that $\mathbf{m} = m_0(x)$ satisfies

$$x = -\frac{1}{\mathbf{m}} + c\left(\frac{\mathbf{m}+1-c+\sqrt{(-\mathbf{m}+1-c)^2+4\mathbf{m}}}{-2\mathbf{m}c'}\right).$$

It follows that $\mathbf{m}$ satisfies

$$\mathbf{m}^2(c'x^2 + cx) + \mathbf{m}(2c'x - c^2 + c + cx(1-c')) + c' + c(1-c') = 0.$$

Solving for **m** we conclude that, with

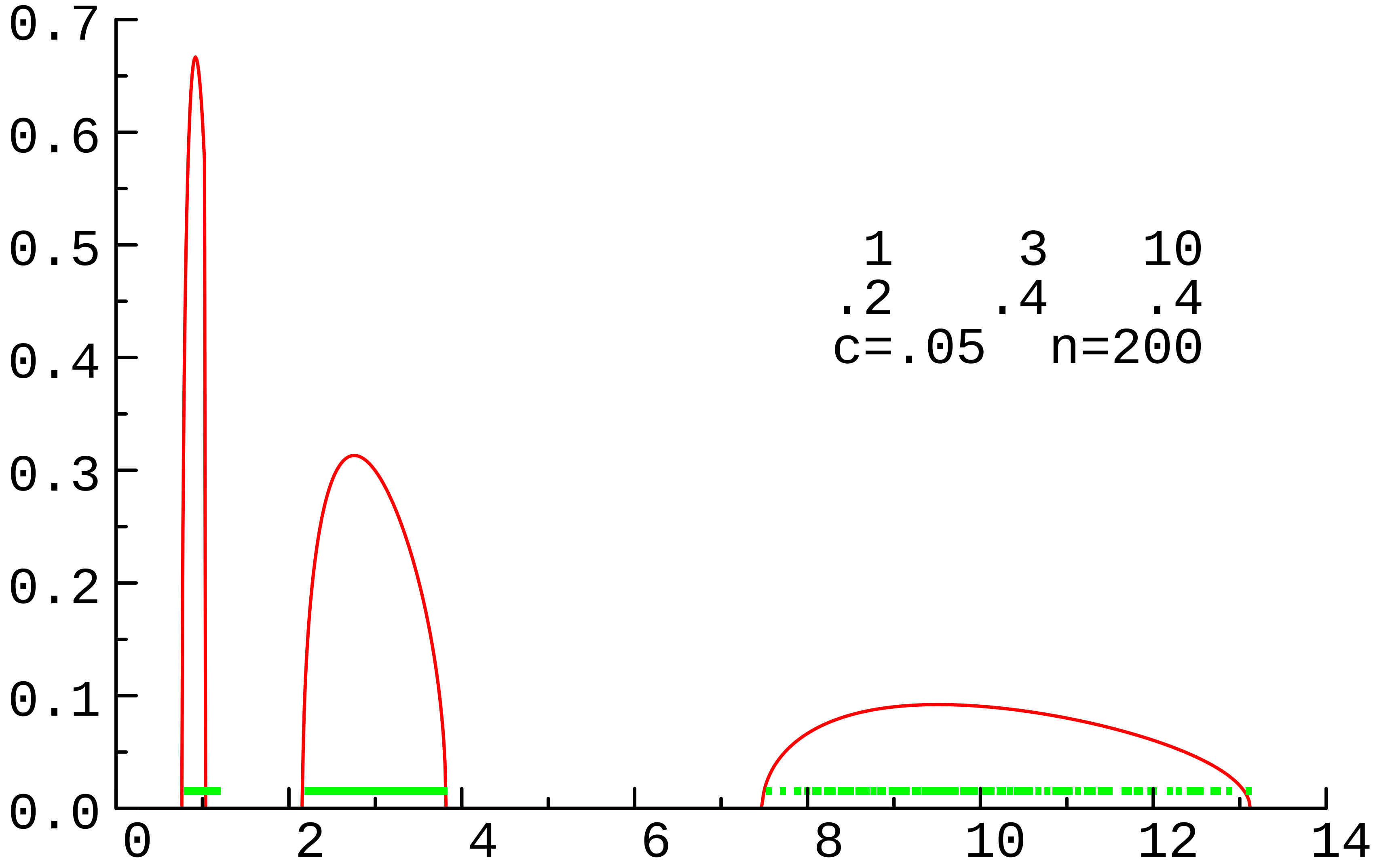
$$b_1 = \left(\frac{1-\sqrt{1-(1-c)(1-c')}}{1-c'}\right)^2 \quad b_2 = \left(\frac{1+\sqrt{1-(1-c)(1-c')}}{1-c'}\right)^2$$

$$f(x) = \begin{cases} \frac{(1-c')\sqrt{(x-b_1)(b_2-x)}}{2\pi x(xc'+c)} & b_1 < x < b_2 \\ 0 & \text{otherwise.} \end{cases}$$

(S. (1985)).

7. Other uses of the Stieltjes transform. We conclude these lectures with two results requiring Stieltjes transforms.

The first concerns the eigenvalues of matrices in Theorem 1.2 outside the support of the limiting distribution. The results mentioned so far clearly say nothing about the possibility of some eigenvalues lingering in this region. Consider this example with T_n given earlier, but now $c = .05$. Below is a scatterplot of the eigenvalues from a simulation with $n = 200$ ($N = 4000$), superimposed on the limiting density.



0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
2
4
6
8
10
12
14
1 3 10
.2 .4 .4
c=.05 n=200

Here the entries of X_n are $N(0,1)$. All the eigenvalues appear to stay close to the limiting support. Such simulations were the prime motivation to prove

THEOREM 7.1 (Bai and S. (1998)). *Let, for any $d > 0$ and d.f. G, $\hat{F}^{d,G}$ denote the limiting e.d.f. of $(1/N)X_n^* T_n X_n$ corresponding to limiting ratio d and limiting F^{T_n} G.*

Assume in addition to the previous assumptions:

a) $\mathsf{E}X_{11} = 0$, $\mathsf{E}|X_{11}|^2 = 1$, *and* $\mathsf{E}|X_{11}|^4 < \infty$.

b) T_n *is nonrandom and* $\|T_n\|$ *is bounded in* n.

c) *The interval* $[a, b]$ *with* $a > 0$ *lies in an open interval outside the support of* $\hat{F}^{c_n, H_n}$ *for all large* n, *where* $H_n = F^{T_n}$.

Then

$$\mathsf{P}(\text{ no eigenvalue of } B_n \text{ appears in } [a, b] \text{ for all large } n\ ) = 1.$$

Steps in proof:

1. Let $\underline{B}_n = (1/N)X_n^* T_n X_n$ $\underline{m}_n = m_{F^{\underline{B}_n}}$ and $\underline{m}_n^0 = m_{\hat{F}^{c_n, H_n}}$. Then for $z = x + iv_n$

$$\sup_{x \in [a,b]} |\underline{m}_n(z) - \underline{m}_n^0(z)| = o(1/Nv_n) \quad a.s.$$

when $v_n = N^{-1/68}$.

2. The proof of 1. allows 1. to hold for $Im(z) = \sqrt{2}v_n, \sqrt{3}v_n, \ldots, \sqrt{34}v_n$. Then almost surely

$$\max_{k \in \{1,\ldots,34\}} \sup_{x \in [a,b]} |\underline{m}_n(x + i\sqrt{k}v_n) - \underline{m}_n^0(x + i\sqrt{k}v_n)| = o(v_n^{67}).$$

We take the imaginary part of these Stieltjes transforms and get

$$\max_{k \in \{1,2\ldots,34\}} \sup_{x \in [a,b]} \left| \int \frac{d(F^{\underline{B}_n}(\lambda) - \hat{F}^{c_n,H_n}(\lambda))}{(x-\lambda)^2 + kv_n^2} \right| = o(v_n^{66}) \quad a.s.$$

Upon taking differences we find with probability one

$$\max_{k_1 \neq k_2} \sup_{x \in [a,b]} \left| \int \frac{v_n^2 \, d(F^{\underline{B}_n}(\lambda) - \hat{F}^{c_n, H_n}(\lambda))}{((x-\lambda)^2 + k_1 v_n^2)((x-\lambda)^2 + k_2 v_n^2)} \right| = o(v_n^{66})$$

$$\max_{\substack{k_1, k_2, k_3 \\ \text{distinct}}} \sup_{x \in [a,b]} \left| \int \frac{(v_n^2)^2 \, d(F^{\underline{B}_n}(\lambda) - \hat{F}^{c_n, H_n}(\lambda))}{((x-\lambda)^2 + k_1 v_n^2)((x-\lambda)^2 + k_2 v_n^2)((x-\lambda)^2 + k_3 v_n^2)} \right| = o(v_n^{66})$$

$$\vdots$$

$$\sup_{x \in [a,b]} \left| \int \frac{(v_n^2)^{33} \, d(F^{\underline{B}_n}(\lambda) - \hat{F}^{c_n, H_n}(\lambda))}{((x-\lambda)^2 + v_n^2)((x-\lambda)^2 + 2v_n^2) \cdots ((x-\lambda)^2 + 34 v_n^2)} \right| = o(v_n^{66}).$$

Thus with probability one

$$\sup_{x \in [a,b]} \left| \int \frac{d(F^{\underline{B}_n}(\lambda) - \hat{F}^{c_n, H_n}(\lambda))}{((x-\lambda)^2 + v_n^2)((x-\lambda)^2 + 2v_n^2) \cdots ((x-\lambda)^2 + 34 v_n^2)} \right| = o(1)$$

Let $0 < a' < a$, $b' > b$ be such that $[a', b']$ is also in the open interval outside the support of $\hat{F}^{c_n, H_n}$ for all large n. We split up the integral and get with probability one

$$\sup_{x\in[a,b]}\left|\int \frac{I_{[a',b']^c}(\lambda)\,d(F^{\underline{B}_n}(\lambda)-\hat{F}^{c_n,H_n}(\lambda))}{((x-\lambda)^2+v_n^2)((x-\lambda)^2+2v_n^2)\cdots((x-\lambda)^2+34v_n^2)}\right.$$

$$\left.+\sum_{\lambda_j\in[a',b']}\frac{v_n^{68}}{((x-\lambda_j)^2+v_n^2)((x-\lambda_j)^2+2v_n^2)\cdots((x-\lambda_j)^2+34v_n^2)}\right|=o(1).$$

Now if, for each term in a subsequence satisfying the above, there is at least one eigenvalue contained in $[a,b]$, then the sum, with x evaluated at these eigenvalues, will be uniformly bounded away from 0. Thus, at these same x values, the integral must also stay uniformly bounded away from 0. But the integral MUST converge to zero $a.s.$ since the integrand is bounded and with probability one, both $F^{\underline{B}_n}$ and $\hat{F}^{c_n,H_n}$ converge weakly to the same limit having no mass on $\{a',b'\}$. Contradiction!

The last result is on the rate of convergence of linear statistics of the eigenvalues of B_n, that is, quantities of the form

$$\int f(x)dF^{B_n}(x) = \frac{1}{n}\sum_{i=1}^{n} f(\lambda_i)$$

where f is a function defined on $[0, \infty)$, and the λ_i's are the eigenvalues of B_n. The result establishes the rate to be $1/n$ for analytic f. It considers integrals of functions with respect to

$$G_n(x) = n[F^{B_n}(x) - F^{c_n,H_n}(x)]$$

where for any $d > 0$ and d.f. G, $F^{d,G}$ is the limiting e.d.f. of $B_n = (1/N)T_n^{1/2}X_nX_n^*T_n^{1/2}$ corresponding to limiting ratio d and limiting F^{T_n} G.

THEOREM 7.2 (Bai and S. (2004)). *Under the assumptions in Theorem 7.1, let $f_1, \ldots, f_r$ be C^1 functions on $\mathbb{R}$ with bounded derivatives, and analytic on an open interval containing*

$$[\liminf_n \lambda_{\min}^{T_n} I_{(0,1)}(c)(1-\sqrt{c})^2, \limsup_n \lambda_{\max}^{T_n}(1+\sqrt{c})^2].$$

Let $\underline{m} = m_{\hat{F}}$. Then

(1) *the random vector*

$$\left(\int f_1(x)\, dG_n(x), \dots, \int f_r(x)\, dG_n(x)\right) \tag{7.1}$$

forms a tight sequence in n.

(2) *If X_{11} and T_n are real and $\mathsf{E}(X_{11}^4) = 3$, then (7.1) converges weakly to a Gaussian vector $(X_{f_1}, \dots, X_{f_r})$, with means*

$$\mathsf{E}X_f = -\frac{1}{2\pi i}\int f(z) \frac{c\int \frac{\underline{m}(z)^3 t^2\, dH(t)}{(1+t\underline{m}(z))^3}}{\left(1 - c\int \frac{\underline{m}(z)^2 t^2\, dH(t)}{(1+t\underline{m}(z))^2}\right)^2} dz \tag{7.2}$$

and covariance function

(7.3)

$$\mathrm{Cov}(X_f, X_g) = -\frac{1}{2\pi^2}\iint \frac{f(z_1)g(z_2)}{(\underline{m}(z_1) - \underline{m}(z_2))^2} \frac{d}{dz_1}\underline{m}(z_1) \frac{d}{dz_2}\underline{m}(z_2) dz_1 dz_2$$

$(f, g \in \{f_1, \ldots, f_r\})$. *The contours in (7.2) and (7.3)* (*two in (7.3), which we may assume to be non-overlapping*) *are closed and are taken in the positive direction in the complex plane, each enclosing the support of* $F^{c,H}$.

(3) *If* X_{11} *is complex with* $\mathsf{E}(X_{11}^2) = 0$ *and* $\mathsf{E}(|X_{11}|^4) = 2$, *then* (2) *also holds, except the means are zero and the covariance function is 1/2 the function given in (7.3)* .

(4) *If the assumptions in (2) or (3) were to hold, then* G_n, *considered as a random element in* $D[0, \infty)$ *(the space of functions on* $[0, \infty)$ *right-continuous with left-hand limits, together with the Skorohod metric) cannot form a tight sequence in* $D[0, \infty)$.

The proof relies on the identity

$$\int f(x)dG(x) = -\frac{1}{2\pi i}\int f(z)m_G(z)dz$$

(f analytic on the support of G, contour positively oriented around the support), and establishes the following results on

$$M_n(z) = n[m_{F^{B_n}}(z) - m_{F^{c_n,H_n}}(z)].$$

a) $\{M_n(z)\}$ forms a tight sequence for z in a sufficiently large contour about the origin.
b) If X_{11} is complex with $\mathsf{E}(X_{11}^2) = 0$ and $\mathsf{E}(X_{11}^4) = 2$, then for $z_1, \ldots, z_r$ with nonzero imaginary parts,

$$(ReM_n(z_1), ImM_n(z_1), \ldots, ReM_n(z_r), ImM_n(z_r))$$

converges weakly to a mean zero Gaussian vector. It follows that M_n, viewed as a random element in the metric space of continuous

$\mathbb{R}^2$-valued functions with domain restricted to a contour in the complex plane, converges weakly to a (2 dimensional) Gaussian process M. The limiting covariance function can be derived from the formula

$$\mathsf{E}(M(z_1)M(z_2)) = \frac{\underline{m}'(z_1)\underline{m}'(z_2)}{(\underline{m}(z_1) - \underline{m}(z_2))^2} - \frac{1}{(z_1 - z_2)^2}.$$

c) If X_{11} is real and $\mathsf{E}(X_{11}^4) = 3$ then b) still holds, except the limiting mean can be derived from

$$\mathsf{E}M(z) = \frac{c\int \frac{\underline{m}^3 t^2 dH(t)}{(1+t\underline{m})^3}}{\left(1 - c\int \frac{\underline{m}^2 t^2 dH(t)}{(1+t\underline{m})^2}\right)^2}$$

and "covariance function" is twice that of the above function.

The difference between (2) and (3), and the difficulty in extending beyond these two cases, arise from

$$\mathsf{E}(X_{\cdot 1}^* A X_{\cdot 1} - \mathsf{tr}\, A)(X_{\cdot 1}^* B X_{\cdot 1} - \mathsf{tr}\, B)$$
$$= (\mathsf{E}(|X_{11}|^4) - |\mathsf{E}(X_{11}^2)|^2 - 2)\sum_i a_{ii}b_{ii} + |\mathsf{E}(X_{11}^2)|^2 \mathsf{tr}\, AB^T + \mathsf{tr}\, AB,$$

valid for square matrices A and B.

Can show

$$(7.2) = \frac{1}{2\pi}\int f'(x) \arg\left(1 - c\int \frac{t^2\underline{m}^2(x)}{(1+t\underline{m}(x))^2} dH(t)\right) dx$$

and

$$(7.3) = \frac{1}{\pi^2}\iint f'(x)g'(y) \ln\left|\frac{\underline{m}(x) - \overline{\underline{m}}(y)}{\underline{m}(x) - \underline{m}(y)}\right| dxdy$$

$$= \frac{1}{2\pi^2} \iint f'(x)g'(y) \ln \left(1 + 4 \frac{\underline{m}_i(x)\underline{m}_i(y)}{|\underline{m}(x) - \underline{m}(y)|^2} \right) dxdy$$

where $\underline{m}_i = \Im \underline{m}$.

For case (2) with $H = I_{[1,\infty)}$ we have for $f(x) = \ln x$ and $c \in (0,1)$

$$\mathsf{E}X_{\ln} = \frac{1}{2}\ln(1-c) \qquad \text{and} \qquad \operatorname{Var} X_{\ln} = -2\ln(1-c).$$

Also, for $c > 0$

$$\mathsf{E}X_{x^r} = \frac{1}{4}((1-\sqrt{c})^{2r} + (1+\sqrt{c})^{2r}) - \frac{1}{2}\sum_{j=0}^{r}\binom{r}{j}^2 c^j$$

and

$$\operatorname{Cov}(X_{x^{r_1}}, X_{x^{r_2}}) = 2c^{r_1+r_2}\sum_{k_1=0}^{r_1-1}\sum_{k_2=0}^{r_2}\binom{r_1}{k_1}\binom{r_2}{k_2}\left(\frac{1-c}{c}\right)^{k_1+k_2} \times \sum_{\ell=1}^{r_1-k_1}\ell\binom{2r_1-1-(k_1+\ell)}{r_1-1}\binom{2r_2-1-k_2+\ell}{r_2-1}$$

(see Jonsson (1982)).

References

Bai, Z.D. and S. "No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices" *Annals of Probability* **26**(1) (1998), pp. 316-345.

Bai, Z.D. and S. "CLT for linear spectral statistics of large dimensional sample covariance matritces" *Annals of Probability* **32** (1A) (2004), pp. 553-605.

Chuah, C.N., Tse, D.N.C., Kahn, J.M., Valenzuela, R.A. "Capacity scaling in MIMO wireless systems under correlated fading" *IEEE Transactions on Information Theory* **48**(3) (2002), pp. 637-650.

Dozier, R.B. and S. a). "On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices " *Journal of Multivariate Analysis* **98** (2007), pp. 678-694.

Dozier, R.B. and S. b). “Analysis of the Limiting Spectral Distribution of Large Dimensional Information-Plus-Noise Type Matrices” to appear in *Journal of Multivariate Analysis*.

Jonsson, D. “Some limit theorems for the eigenvalues of a sample covariance matrix” *Journal of Multivariate Analysis* **12**(1) (1982), pp. 1-38.

V. A. Marčenko and L. A. Pastur, “Distribution of eigenvalues for some sets of random matrices” USSR-Sb., 1 (1967), pp. 457-483.

S. “The limiting eigenvalue distribution of a multivariate F-matrix” *SIAM Journal on Mathematical Analysis* **16**(3) (1985), pp. 641-646.

S. “Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices” *Journal of Multivariate Analysis* **55**(2) (1995), pp. 331-339.

S. and Bai, Z.D. “On the empirical distribution of eigenvalues of a class of large dimensional random matrices” *Journal of Multivariate Analysis* **54**(2) (1995), pp. 175-192.

S. and Choi, Sang-Il. “Analysis of the limiting spectral distribution function of large dimensional random matrices” *Journal of Multivariate Analysis* **54**(2) (1995), pp. 295-309.

Silverstein, J.W. and Combettes, P.L. “Signal detection via spectral theory of large dimensional random matrices” *IEEE Transactions on Signal Processing* **40**(8) (1992), pp. 2100-2105.

Y. Q. Yin, “Limiting spectral distribution for a class of random matrices” *Journal of Multivariate Analysis* 20 (1986), pp. 50-68.

On the Signal-to-Interference-Ratio of CDMA Systems in Wireless Communications

By Z.D. Bai and Jack W. Silverstein

I. Results. *Direct-sequence code-division multiple-access* (or DS-CDMA) network, used in wireless communications:

Each user has assigned to it a vector of high dimension, called a *signature sequence*, used to transmit data.

Suppose:

1. K users and L receive antennas.
2. N is the dimension of the signature sequences, $s_k \in \mathbb{C}^N$ is the signature sequence assigned to user k.
3. At a particular instant of time $X_k \in \mathbb{R}$ is the value transmitted by user k.
4. $T_k \in \mathbb{R}^+$ is user k's transmit power.
5. $\gamma_k(\ell)$ fading channel gain from user k to antenna ℓ.
6. X_k's are independent standardized random variables.

7. $W(\ell) \in \mathbb{C}^N$ noise associated with transmission to antenna ℓ, entries $W_i(\ell)$ i.i.d. across i and ℓ, mean zero and

$$\mathsf{E}|W_i(\ell)|^2 = \sigma^2.$$

Then the data recorded at antenna ℓ is modeled by

$$Y(\ell) = \sum_{k=1}^{K} X_k \sqrt{T_k}\gamma_k(\ell) s_k + W(\ell).$$

Let $Y = [Y(1)^T, \ldots, Y(L)^T]^T \in \mathbb{C}^{NL}$.

Goal: Capture the transmitted X_k for each user in a linear fashion, that is, by taking the inner product of Y with an appropriate vector $c_k \in \mathbb{C}^{NL}$ (the linear receiver for user k). For user 1, $\hat{X}_1 = c_1^* Y$ is the estimate of transmitted X_1.

The output *signal-to-interference ratio*

$$\frac{|c_1^* \hat{s}_1|^2}{\sigma^2 \|c_1\|^2 + \sum_{k=2}^{K} |c_1^* \hat{s}_k|^2}$$

associated with user 1, is typically used as a measure for evaluating the performance of the linear receiver, where

$$\hat{s}_k = \sqrt{T_k}[\gamma_k(1) s_k^T, \ldots, \gamma_k(\ell) s_k^T]^T.$$

Fact: The choice of c_1 which minimizes $\mathsf{E}(\hat{X} - X)^2$ (the *minimum mean-square error*) also maximizes user 1's signal-to-interference ratio, the latter taking the value

$$\mathbf{SIR}_1 = \hat{s}_1^* \left(\sum_{k=2}^{K} \hat{s}_k \hat{s}_k^* + \sigma^2 I \right)^{-1} \hat{s}_1,$$

where I is the $NL \times NL$ identity matrix.

Properties of $\mathbf{SIR_1}$ and their dependency on the $\gamma_k(\ell)$'s, T_k's, σ^2, L, N, and K, when the latter two values are large, are explored by proving limiting results, as N and K approach infinity with their ratio approaching a positive constant, under the assumption that the s_k's are randomly generated (which is usually done in practice). They are independent i.i.d. random vectors containing i.i.d. mean zero, variance $1/N$ entries, independent of the $\gamma_k(\ell)$'s and T_k's.

THEOREM. *Let* $\{s_{ij} : i, j = 1, 2, \ldots\}$ *be a doubly infinite array of i.i.d. complex random variables with* $\mathsf{E}s_{11} = 0$, $\mathsf{E}|s_{11}|^2 = 1$. *Define for* $k = 1, 2, \ldots, K$ $s_k = s_k(N) = (s_{1\,k}, s_{2\,k}, \ldots, s_{N\,k})^T$. *We assume* $K = K(N)$ *and* $K/N \to c > 0$ *as* $N \to \infty$. *For each* N *let* $\gamma_k(\ell) = \gamma_k^N(\ell) \in \mathbb{C}$, $T_k = T_k^N \in \mathbb{R}^+$, $k = 1, \ldots, K$, $\ell = 1, \ldots, L$ *be random variables, independent of* $s_1, \ldots, s_K$. *Let for each* N *and* k

$$\alpha_k = \alpha_k^N = \sqrt{T_k}(\gamma_k(1), \ldots, \gamma_k(L))^T.$$

Assume almost surely, the empirical distribution of $\alpha_1, \ldots, \alpha_K$ *weakly converges to a probability distribution* H *in* $\mathbb{C}^L$.

Let $\beta_k = \beta_k(N) = \sqrt{T_k}(\gamma_k(1)s_k^T, \ldots, \gamma_k(L)s_k^T)^T$, *and*

$$C = C(N) = \frac{1}{N}\sum_{k=2}^{K} \beta_k\beta_k^*.$$

Define

$$\mathrm{SIR}_1 = \frac{1}{N}\beta_1^*(C+\sigma^2 I)^{-1}\beta_1.$$

then, with probability one

$$\lim_{N\to\infty} \mathrm{SIR}_1 = T_1 \sum_{\ell,\ell'=1}^{L} \bar{\gamma}_1(\ell)\gamma_1(\ell')a_{\ell,\ell'}$$

where the $L \times L$ matrix $A = (a_{\ell,\ell'})$ is nonrandom, Hermitian positive definite, and is the unique Hermitian positive definite matrix satisfying

$$A = \left(c\,\mathsf{E}\frac{\alpha\alpha^*}{1+\alpha^* A\alpha} + \sigma^2 I_L\right)^{-1} \tag{1.1}$$

where $\alpha \in \mathbb{C}^L$ has distribution H and I_L is the $L \times L$ identity matrix.

Clearly $\mathbf{SIR}_1$ defined in this theorem is the same as the one initially introduced, the only difference in notation being the removal of the scaling by $1/\sqrt{N}$ in the definition of the s_k's.

Importance: Arbitrary scenarios can be analyzed and compared. In applications the empirical distribution of the α_k's would typically be used for H, the matrix A thereby satisfying

$$A = \left(\frac{1}{N} \sum_{k=2}^{K} \frac{\alpha_k \alpha_k^*}{1 + \alpha_k^* A \alpha_k} + \sigma^2 I_L \right)^{-1}.$$

Although there appears to be no explicit solution to (1.1), the paper shows that A can be computed numerically by iteration of the right side of (1.1), provided the eigenvalues of the initial choice in the iteration lie in a certain closed interal in $(0, \infty)$.

Fact: Let α_ℓ denote the ℓ^{th} entry of the random vector α having distribution H. If the $\gamma_k(\ell)$'s are independent and circularly symmetric, with angles independent of the T_k's, then $\sqrt{T_k}\gamma_k(\ell)$ and $\sqrt{T_k}\gamma_k(\ell')$ are uncorrelated for $\ell \neq \ell'$. It follows that for $\ell \neq \ell'$ and positive $a_1, \dots, a_L$

$$\mathsf{E}\frac{\alpha_\ell \bar{\alpha}_{\ell'}}{1 + \sum_{\underline{\ell}} a_{\underline{\ell}} |\alpha_{\underline{\ell}}|^2} = 0 \tag{1.2}$$

COROLLARY 1. *Under the conditions in the Theorem and (1.2), the limiting $A = diag(a_1, \dots, a_L)$ where the a_ℓ's are positive satisfying*

$$a_\ell = \frac{1}{c\,\mathsf{E}\frac{|\alpha_\ell|^2}{1+\sum_{\underline{\ell}} a_{\underline{\ell}}|\alpha_{\underline{\ell}}|^2} + \sigma^2} \tag{1.3}$$

COROLLARY 2. *Suppose the conditions in the Theorem are met except, for the limiting behavior of the α_k's, it is only known that:*

1. *the empirical distribution of*

$$T_k(|\gamma_k(1)|^2, \ldots, |\gamma_k(L)|^2)^T \quad 2 \le k \le K \tag{1.4}$$

converges almost surely in distribution to a probability distribution G *in* $\mathbb{R}^L$, *and*

2. *for* $\ell \neq \ell'$ *and positive* $a_1, \ldots, a_L$

$$\frac{1}{K-1} \sum_{k=2}^{K} \frac{T_k \gamma_k(\ell) \bar{\gamma}_k(\ell')}{1 + \sum_{\underline{\ell}} a_{\underline{\ell}} T_k |\gamma_k(\underline{\ell})|^2} \to 0$$

almost surely, as $N \to \infty$.

Let $(\delta_1, \ldots, \delta_L)^T \in \mathbb{R}^L$ *denote a random vector having distribution* G. *Then the conclusions of the Theorem and Corollary 1 hold, with each* $|\alpha_\ell|^2$ *in (1.3) replaced by* δ_ℓ.

Remarks:

1. The assumption of the s_{ij} coming from a doubly infinite array can be replaced with $s_{ij} = s_{ij}(N)$, $1 \le i \le N$, $1 \le j \le K$, with no dependency assumptions for different N, provided $\mathsf{E}|s_{11}|^4 < \infty$.

2. The Theorem only provides limiting properties of the signal-to-interference ratio with respect to one user.

Let

$$C_k = \frac{1}{N}\left(\sum_{j=1}^{K} \beta_j \beta_j^* - \beta_k \beta_k^*\right).$$

Then

$$\mathbf{SIR}_k \equiv \frac{1}{N}\beta_k^*(C_k + \sigma^2 I)^{-1}\beta_k = \frac{1}{N}\sum_{\ell,\ell'} \bar{\alpha}_k(\ell)\alpha_k(\ell')s_k^*(C_k + \sigma^2 I)^{-1}_{\ell,\ell'} s_k$$

represents user k's best signal-to-interference ratio.

If $\mathsf{E}|s_{11}|^4 < \infty$, or if the doubly infinite array assumption is dropped, $\mathsf{E}|s_{11}|^6 < \infty$, then

$$\max_{k \le K} |N^{-1} s_k^* (C_k + \sigma^2 I)^{-1}_{\ell,\ell'} s_k - a_{\ell,\ell'}| \to 0 \quad a.s.$$

as $N \to \infty$.

2. Basic tools. For any rectangular matrix X, **vec**X denotes the column vector consisting of stacking the columns of X on top of each other, first column on top, last on bottom.

Lemma 2.1. *Let $\sigma^2 > 0$, B, A $n \times n$ matrices with B Hermitian non-negative definite, and $x \in \mathbb{C}^n$. Then*

$$|\mathrm{tr}\,((B+xx^*+\sigma^2 I)^{-1}-(B+\sigma^2 I)^{-1})A| = \left|\frac{x^*(B+\sigma^2 I)^{-1}A(B+\sigma^2 I)^{-1}x}{1+x^*(B+\sigma^2 I)^{-1}x}\right|$$

$$\leq \frac{\|A\|}{\sigma^2}.$$

Lemma 2.2. *For any matrix A $N \times K$ and $\sigma^2 > 0$*

$$(AA^* + \sigma^2 I_N)^{-1} = \sigma^{-2}(I_N - A(A^*A + \sigma^2 I_K)^{-1}A^*).$$

Lemma 2.3. *Suppose $A_1, ..., A_L$ are $N \times K$, and $\sigma^2 > 0$. Define the ℓ, ℓ' block of the $NL \times NL$ matrix $\mathcal{A}$ by $\mathcal{A}_{\ell,\ell'} = A_\ell A_{\ell'}^*$, and,* *splitting* $(A+\sigma^2 I)^{-1}$ *into* L^2 $N \times N$ *matrices, let* $(A+\sigma^2 I)^{-1}_{\ell,\ell'}$ *denote its* ℓ, ℓ' *block. Then*

$$(A+\sigma^2 I)^{-1}_{\ell,\ell'} = \sigma^{-2}\left(\delta_{\ell,\ell'} I_N - A_\ell \left(\sum_{\underline{\ell}} A^*_{\underline{\ell}} A_{\underline{\ell}} + \sigma^2 I_K\right)^{-1} A^*_{\ell'}\right).$$

LEMMA 2.4 *Given* $A_1, ..., A_L$ *are* $N \times K$ *and* $z_1, \ldots, z_\ell \in \mathbb{C}$ *with* $\sum_\ell |z_\ell|^2 = 1$

$$\left(\sum_\ell A_\ell z_\ell\right)\left(\sum_\ell A^*_\ell \hat{z}_\ell\right) \preceq \sum_\ell A_\ell A^*_\ell,$$

where "$\preceq$*" represents the partial ordering on Hermitian nonnegative definite matrices.*

Lemma 2.5 *For* $A_1, ..., A_L$, A, σ^2 *in Lemma 2.3, the* $L \times L$ *matrix* $(\mathrm{tr}\,(A+\sigma^2 I)^{-1}_{\ell,\ell'})$ *is positive definite with smallest eigenvalue bounded below by*

$$\mathrm{tr}\left(\sum_\ell A_\ell A^*_\ell + \sigma^2 I_N\right)^{-1}.$$

For $A = (a_{ij})$ $m \times n$ and B $p \times q$, the Kronecker product of A and B, denoted by $A \otimes B$, is the $mp \times nq$ matrix, expressed in blocks of $p \times q$ matrices, the i, j block being $a_{ij}B$. We will need the following, which is Lemma 4.2.10 of Horn and Johnson (1991)

LEMMA 2.6. (Lemma 4.2.10 of Horn and Johnson (1991)) *For A $m \times n$, B $p \times q$, C $n \times k$ and D $q \times r$ we have*

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD).$$

LEMMA 2.7, SCHAUDER FIXED POINT THEOREM (Nirenberg (1961)). *If $\mathcal{A}$ is a convex, compact subset of a Banach space $\mathcal{X}$ and $g : \mathcal{A} \to \mathcal{A}$ is continuous, then g has a fixed point in $\mathcal{A}$.*

LEMMA 2.8. *Let $A = (a_{ij}) = (a_1, \ldots, a_n)$ $(m \times n)$ and B $(h \times g)$ be two random matrices, the entries having bounded second moments.*

Then

$$\|\mathsf{E}A \otimes B\| \leq \min\left(\sqrt{\|\mathsf{E}AA^*\|\,\|\mathsf{E}B^*B\|}, \sqrt{\|\mathsf{E}A^*A\|\,\|\mathsf{E}BB^*\|}\right).$$

LEMMA 2.9. *For $X = (X_1, \ldots, X_n)^T$ i.i.d. standardized entries, C $n \times n$, we have for any $p \geq 2$*

$$\mathsf{E}|X^*CX - \mathrm{tr}\, C|^p \leq K_p\left(\left(\mathsf{E}|X_1|^4 \mathrm{tr}\, CC^*\right)^{p/2} + \mathsf{E}|X_1|^{2p} \mathrm{tr}\,(CC^*)^{p/2}\right)$$

where the constant K_p does not depend on n, C, nor on the distribution of X_1.

3. Sketch of proof of Theorem 1.1. Write $\sqrt{T_k}\gamma_k(\ell)$ as $\alpha_k(\ell)$. It is first shown that, after truncating and centralizing the entries of s_1, from Lemma 2.9, with probability one

$$|\mathrm{SIR}_1 - N^{-1}\sum_{\ell,\ell'} \bar{\alpha}_1(\ell)\alpha_1(\ell')\mathrm{tr}\,(C+\sigma^2 I)^{-1}_{\ell,\ell'}| \to 0.$$

After additional truncation and centralization steps can assume for each N:

1. $s_{nk} = s_{nk}(N)$, $1 \leq n \leq N$, $2 \leq k \leq K$, i.i.d. standardized random variables with $|s_{nk}| \leq \log N$.
2. $\max_{2\leq k\leq K,\ell} |\alpha_k(\ell)|^2 \leq \log N$.

Define $C_{(k)} = C - (1/N)\beta_k\beta_k^*$.

Define the $L \times L$ matrix $\underline{B} = (\underline{b}_{\ell,\ell'})$ with

$$\underline{b}_{\ell,\ell'} = \frac{1}{N}\sum_{k=2}^{K} \frac{\bar{\alpha}_k(\ell')\alpha_k(\ell)}{1+\frac{1}{N}\beta_k^*(C_{(k)}+\sigma^2 I)^{-1}\beta_k},$$

and define the $NL \times NL$ matrix B in terms of the Kronecker product: $B = \underline{B} \otimes I_N$. We have $(B + \sigma^2 I)^{-1} = (\underline{B} + \sigma^2 I_L)^{-1} \otimes I_N$. Denote the ℓ, ℓ' entry of $(\underline{B} + \sigma^2 I_L)^{-1}$ by $\hat{b}_{\ell,\ell'}$.

Let $I_{\ell',\ell}$ denote the $NL \times NL$ matrix consisting of the $N \times N$ identity matrix in the ℓ', ℓ block, zeros elsewhere. We write

$$C + \sigma^2 I - (B + \sigma^2 I) = \frac{1}{N} \sum_{k=2}^{K} \beta_k \beta_k^* - B.$$

Taking inverses on each side, we have

$$(B + \sigma^2 I)^{-1} - (C + \sigma^2 I)^{-1}$$

$$= \frac{1}{N} \sum_{k=2}^{K} (B + \sigma^2 I)^{-1} \beta_k \beta_k^* (C + \sigma^2 I)^{-1} - (B + \sigma^2 I)^{-1} B (C + \sigma^2 I)^{-1}$$

$$= \frac{1}{N} \sum_{k=2}^{K} \frac{(B + \sigma^2 I)^{-1} \beta_k \beta_k^* (C_{(k)} + \sigma^2 I)^{-1}}{1 + (1/N) \beta_k^* (C_{(k)} + \sigma^2 I)^{-1} \beta_k} - (B + \sigma^2 I)^{-1} B (C + \sigma^2 I)^{-1}.$$

Multiplying on the right by $I_{\ell',\ell}$, taking traces, and dividing by N we get

$$N^{-1}\mathrm{tr}\,(B+\sigma^2 I)^{-1}_{\ell,\ell'} - N^{-1}\mathrm{tr}\,(C+\sigma^2 I)^{-1}_{\ell,\ell'}$$

$$= \frac{1}{N}\sum_{k=2}^{K} \frac{\frac{1}{N}\beta_k^*(C_{(k)}+\sigma^2 I)^{-1} I_{\ell',\ell}(B+\sigma^2 I)^{-1}\beta_k}{1+\frac{1}{N}\beta_k^*(C_{(k)}+\sigma^2 I)^{-1}\beta_k}$$

$$-N^{-1}\mathrm{tr}\,B(C+\sigma^2 I)^{-1} I_{\ell',\ell}(B+\sigma^2 I)^{-1}$$

$$= \frac{1}{N}\sum_{k=2}^{K} \frac{\sum_{\underline{\ell},\underline{\ell}'} \bar{\alpha}_k(\underline{\ell})\alpha_k(\underline{\ell}')\frac{1}{N} s_k^*[(C_{(k)}+\sigma^2 I)^{-1} I_{\ell',\ell}(B+\sigma^2 I)^{-1}]_{\underline{\ell},\underline{\ell}'} s_k}{1+\frac{1}{N}\beta_k^*(C_{(k)}+\sigma^2 I)^{-1}\beta_k}$$

$$-\frac{1}{N}\mathrm{tr}\sum_{\underline{\ell},\underline{\ell}'} B_{\underline{\ell}',\underline{\ell}}[(C+\sigma^2 I)^{-1} I_{\ell',\ell}(B+\sigma^2 I)^{-1}]_{\underline{\ell},\underline{\ell}'}$$

$$= \sum_{\underline{\ell},\underline{\ell}'} \frac{1}{N} \left[\sum_{k=2}^{K} \frac{1}{N} \frac{s_k^* [(C_{(k)} + \sigma^2 I)^{-1} I_{\ell',\ell} (B + \sigma^2 I)^{-1}]_{\underline{\ell},\underline{\ell}'} s_k \bar{\alpha}_k(\underline{\ell}) \alpha_k(\underline{\ell}')}{1 + \frac{1}{N} \beta_k^* (C_{(k)} + \sigma^2 I)^{-1} \beta_k} \right.$$

$$\left. - \operatorname{tr} B_{\underline{\ell}',\underline{\ell}} [(C + \sigma^2 I)^{-1} I_{\ell',\ell} (B + \sigma^2 I)^{-1}]_{\underline{\ell},\underline{\ell}'} \right]$$

$$= \sum_{\underline{\ell},\underline{\ell}'} \frac{1}{N} \left[\sum_{k=2}^{K} \frac{1}{N} \frac{s_k^* (C_{(k)} + \sigma^2 I)^{-1}_{\underline{\ell},\ell'} (B + \sigma^2 I)^{-1}_{\ell,\underline{\ell}'} s_k \bar{\alpha}_k(\underline{\ell}) \alpha_k(\underline{\ell}')}{1 + \frac{1}{N} \beta_k^* (C_{(k)} + \sigma^2 I)^{-1} \beta_k} \right.$$

$$\left. - \operatorname{tr} B_{\underline{\ell}',\underline{\ell}} [(C + \sigma^2 I)^{-1}_{\underline{\ell},\ell'} (B + \sigma^2 I)^{-1}_{\ell,\underline{\ell}'} \right]$$

$$= \sum_{\underline{\ell},\underline{\ell}'} \frac{1}{N} \sum_{k=2}^{K} \frac{\hat{b}_{\ell,\underline{\ell}'} \bar{\alpha}_k(\underline{\ell}) \alpha_k(\underline{\ell}') N^{-1} (s_k^* (C_{(k)} + \sigma^2 I)^{-1}_{\underline{\ell},\ell'} s_k - \operatorname{tr} (C + \sigma^2 I)^{-1}_{\underline{\ell},\ell'})}{1 + \frac{1}{N} \beta_k^* (C_{(k)} + \sigma^2 I)^{-1} \beta_k}.$$

Noticing that

$$N^{-1}\mathrm{tr}\,(B+\sigma^2 I)^{-1}_{\ell,\ell'} = \hat{b}_{\ell,\ell'},$$

we get from Lemmas 2.1 and 2.9

$$|\hat{b}_{\ell,\ell'} - N^{-1}\mathrm{tr}\,(C+\sigma^2 I)^{-1}_{\ell,\ell'}| \to 0 \quad a.s.$$

Again, from these lemmas

$$\left| \underline{b}_{\ell,\ell'} - \frac{1}{N}\sum_{k=2}^{K} \frac{\bar{\alpha}_k(\ell')\alpha_k(\ell)}{1+\sum_{\underline{\ell},\underline{\ell}'} \bar{\alpha}_k(\underline{\ell})\alpha_k(\underline{\ell}')N^{-1}\mathrm{tr}\,(C+\sigma^2 I)^{-1}_{\underline{\ell},\underline{\ell}'}} \right| \to 0 \quad a.s.$$

4. Proof of uniqueness. Suppose A and $\widetilde{A}$ are two different $L \times L$ Hermitian positive definite matrices satisfying (1.1). Then

$$A - \widetilde{A} = c\,\mathsf{E}\frac{A\alpha\alpha^*\widetilde{A}\alpha^*(A-\widetilde{A})\alpha}{(1+\alpha^*A\alpha)(1+\alpha^*\widetilde{A}\alpha)}.$$

Multiplying $A^{-1/2}$ on the left and $\widetilde{A}^{-1/2}$ on the right we obtain

$$A^{1/2}\widetilde{A}^{-1/2} - A^{-1/2}\widetilde{A}^{1/2} = c\,\mathsf{E}\frac{A^{1/2}\alpha\alpha^*\widetilde{A}^{1/2}\alpha^*(A-\widetilde{A})\alpha}{(1+\alpha^*A\alpha)(1+\alpha^*\widetilde{A}\alpha)}$$

$$= c\,\mathsf{E}\frac{\eta\widetilde{\eta}^*\eta^*(A^{1/2}\widetilde{A}^{-1/2} - A^{-1/2}\widetilde{A}^{1/2})\widetilde{\eta}}{(1+\alpha^*A\alpha)(1+\alpha^*\widetilde{A}\alpha)},$$

where $\eta = A^{1/2}\alpha$ and $\widetilde{\eta} = \widetilde{A}^{1/2}\alpha$. Write

$$\mu = \mathbf{vec}(A^{1/2}\widetilde{A}^{-1/2} - A^{-1/2}\widetilde{A}^{1/2}).$$

With the aid of the Kronecker product we can write the above equation as

$$\mu = c\,\mathsf{E}\frac{(\overline{\widetilde{\eta}} \otimes \eta)(\widetilde{\eta}^T \otimes \eta^*)}{(1+\alpha^* A\alpha)(1+\alpha^* \widetilde{A}\alpha)}\mu. \tag{4.1}$$

Using Lemma 2.6 we have

$$c\,\mathsf{E}\frac{(\overline{\widetilde{\eta}} \otimes \eta)(\widetilde{\eta}^T \otimes \eta^*)}{(1+\alpha^* A\alpha)(1+\alpha^* \widetilde{A}\alpha)} = c\,\mathsf{E}\left[\frac{\overline{\widetilde{\eta}\widetilde{\eta}^*}}{1+\alpha^* \widetilde{A}\alpha} \otimes \frac{\eta\eta^*}{1+\alpha^* \widetilde{A}\alpha}\right]$$

and, since $\mu \neq 0$, this matrix has an eigenvalue equal to 1. By Lemma 2.8 its largest squared eigenvalue cannot be greater than

$$\left\| c\,\mathsf{E}\left(\frac{\widetilde{\eta}\widetilde{\eta}^*}{1+\alpha^* \widetilde{A}\alpha}\right)^2 \right\| \left\| c\,\mathsf{E}\left(\frac{\eta\eta^*}{1+\alpha^* A\alpha}\right)^2 \right\|.$$

We have

$$c\,\mathsf{E}\left(\frac{\eta\eta^*}{1+\alpha^* A\alpha}\right)^2 = c\,\mathsf{E}\frac{A^{1/2}\alpha\alpha^* A^{1/2}\alpha^* A\alpha}{(1+\alpha^* A\alpha)^2},$$

and since

$$\frac{A^{1/2}\alpha\alpha^* A^{1/2}}{1+\alpha^* A\alpha} - \frac{A^{1/2}\alpha\alpha^* A^{1/2}\alpha^* A\alpha}{(1+\alpha^* A\alpha)^2}$$

is non-negative definite we have

$$c\,\mathsf{E}\left(\frac{\eta\eta^*}{1+\alpha^* A\alpha}\right)^2 \preceq c\,\mathsf{E}\frac{A^{1/2}\alpha\alpha^* A^{1/2}}{1+\alpha^* A\alpha} = A^{1/2}(A^{-1} - \sigma^2 I_L)A^{1/2}$$

$$= I_L - \sigma^2 A,$$

the eigenvalues of which must all be less than one. The same result applies for the other matrix involving $\widetilde{A}$. Therefore, the matrix in (4.1) cannot have an eiqenvalue equal to one, a contradiction. So we conclude that there is only one Hermitian positive definite solution to (1.1).