The polynomial method for random matrices

N. Raj Rao* Alan Edelman†

October 31, 2007

Abstract

We define a class of "algebraic" random matrices. These are random matrices for which the Stieltjes transform of the limiting eigenvalue distribution function is algebraic, *i.e.*, it satisfies a (bivariate) polynomial equation. The Wigner and Wishart matrices whose limiting eigenvalue distributions are given by the semi-circle law and the Marčenko-Pastur law are special cases.

Algebraicity of a random matrix sequence is shown to act as a certificate of the computability of the limiting eigenvalue density function. The limiting moments of algebraic random matrix sequences, when they exist, are shown to satisfy a finite depth linear recursion so that they may often be efficiently enumerated in closed form.

In this article, we develop the mathematics of the *polynomial method* which allows us to describe the class of algebraic matrices by its generators and map the constructive approach we employ when proving algebraicity into a software implementation that is available for download in the form of the RMTool random matrix "calculator" package. Our characterization of the closure of algebraic probability distributions under free additive and multiplicative convolution operations allows us to simultaneously establish a framework for computational (non-commutative) "free probability" theory. We hope that the tools developed allow researchers to finally harness the power of the infinite random matrix theory.

Key words Random matrices, stochastic eigen-analysis, free probability, algebraic functions, resultants, D-finite series.

1. Introduction

We propose a powerful method that allows us to calculate the limiting eigenvalue distribution of a large class of random matrices. We see this method as allowing us to expand our reach beyond the well known special random matrices whose limiting eigenvalue distributions have the semi-circle density [38], the Marčenko-Pastur density [18], the McKay density [19] or their close cousins [8, 25]. In particular, we encode transforms of the limiting eigenvalue distribution function as solutions of bivariate polynomial equations. Then canonical operations on the random matrices become operations on the bivariate polynomials. We illustrate this with a simple example. Suppose we take the Wigner matrix, sampled in Matlab as:

```
G = sign(randn(N))/sqrt(N); A = (G+G')/sqrt(2);
```

whose eigenvalues in the $N \to \infty$ limit follow the semicircle law, and the Wishart matrix which may be sampled in Matlab as:

```
G = randn(N,2*N)/sqrt(2*N); B = G*G';
```

*MIT Department of Electrical Engineering and Computer Science, Cambridge, MA 02139, raj@mit.edu
†MIT Department of Mathematics, Cambridge, MA 02139, edelman@math.mit.edu

whose eigenvalues in the limit follow the Marčenko-Pastur law. The associated limiting eigenvalue distribution functions have Stieltjes transforms $m_A(z)$ and $m_B(z)$ that are solutions of the equations $L^A_{\rm mz}(m,z)=0$ and $L^B_{\rm mz}(m,z)=0$, respectively, where

$$L^A_{\rm mz}(m,z) = m^2 + z\,m + 1, \qquad L^B_{\rm mz}(m,z) = m^2 z - (-2\,z+1)\,m + 2.$$

The sum and product of independent samples of these random matrices have limiting eigenvalue distribution functions whose Stieltjes transform is a solution of the bivariate polynomial equations $L^{A+B}_{\rm mz}(m,z)=0$ and $L^{AB}_{\rm mz}(m,z)=0$, respectively, which can be calculated from $L^A_{\rm mz}$ and $L^B_{\rm mz}$ alone. To obtain $L^{A+B}_{\rm mz}(m,z)$ we apply the transformation labelled as "Add Atomic Wishart" in Table 7 with $c=2$, $p_1=1$ and $\lambda_1 = 1/c = 0.5$ to obtain the operational law

$$L^{A+B}_{\rm mz}(m,z) = L^A_{\rm mz}\left(m, z - \frac{1}{1+0.5m}\right). \tag{1.1}$$

Substituting $L^A_{\rm mz} = m^2 + z\,m + 1$ in (1.1) and clearing the denominator, yields the bivariate polynomial

$$L^{A+B}_{\rm mz}(m,z) = m^3 + (z+2)\,m^2 - (-2\,z+1)\,m + 2. \tag{1.2}$$

Similarly, to obtain $L^{AB}_{\rm mz}$, we apply the transformation labelled as "Multiply Wishart" in Table 7 with $c=0.5$ to obtain the operational law

$$L^{AB}_{\rm mz}(m,z) = L^A_{\rm mz}\left((0.5-0.5zm)\,m, \frac{z}{0.5-0.5zm}\right). \tag{1.3}$$

Substituting $L^A_{\rm mz} = m^2 + z\,m + 1$ in (1.3) and clearing the denominator, yields the bivariate polynomial

$$L^{AB}_{\rm mz}(m,z) = m^4 z^2 - 2\,m^3 z + m^2 + 4\,mz + 4. \tag{1.4}$$

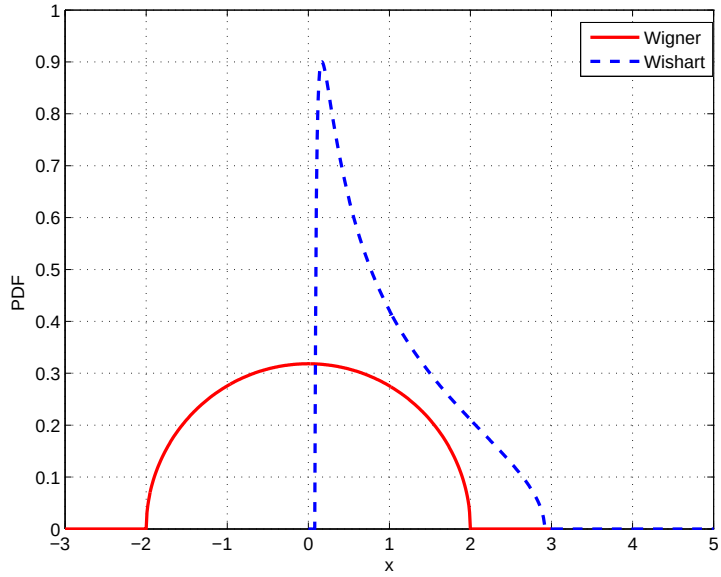
Figure 1 plots the density function associated with the limiting eigenvalue distribution for the Wigner and Wishart matrices as well as their sum and product extracted directly from $L^{A+B}_{\rm mz}(m,z)$ and $L^{AB}_{\rm mz}(m,z)$. In these examples, algebraically extracting the roots of these polynomials using the cubic or quartic formulas is of little use except to determine the limiting density function. As we shall demonstrate in Section 8., the algebraicity of the limiting distribution (in the sense made precise next) is what allows us to readily enumerate the moments efficiently directly from the polynomials $L^{A+B}_{\rm mz}(m,z)$ and $L^{AB}_{\rm mz}(m,z)$.

1.1 Algebraic random matrices: Definition and Utility

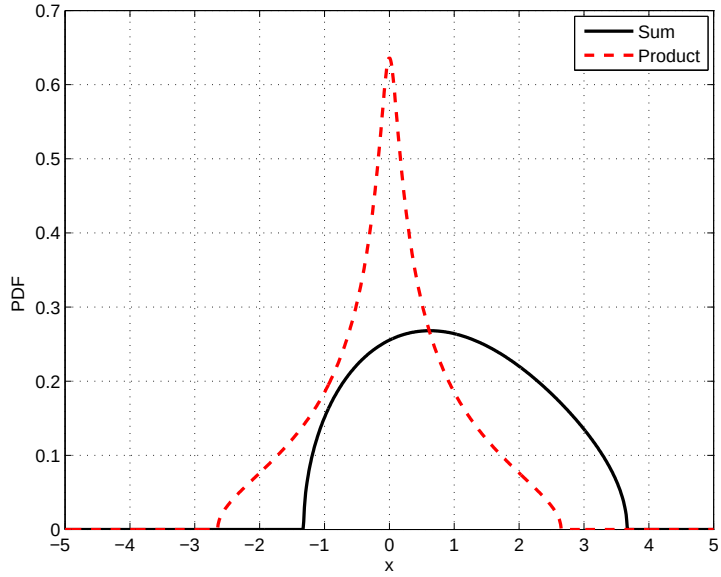
A central object in the study of large random matrices is the empirical distribution function which is defined, for an $N\times N$ matrix $\mathbf{A}_N$ with real eigenvalues, as

$$F^{\mathbf{A}_N}(x) = \frac{\text{Number of eigenvalues of } \mathbf{A}_N \le x}{N}. \tag{1.5}$$

For a large class of random matrices, the empirical distribution function $F^{\mathbf{A}_N}(x)$ converges, for every x, almost surely (or in probability) as $N\to\infty$ to a non-random distribution function $F^A(x)$. The dominant theme of this paper is that "algebraic" random matrices form an important subclass of analytically tractable random matrices and can be effectively studied using combinatorial and analytical techniques that we bring into sharper focus in this paper.



(a) The limiting eigenvalue density function for the GOE and Wishart matrices.



(b) The limiting eigenvalue density function for the sum and product of independent GOE and Wishart matrices.

Figure 1: A representative computation using the random matrix calculator.

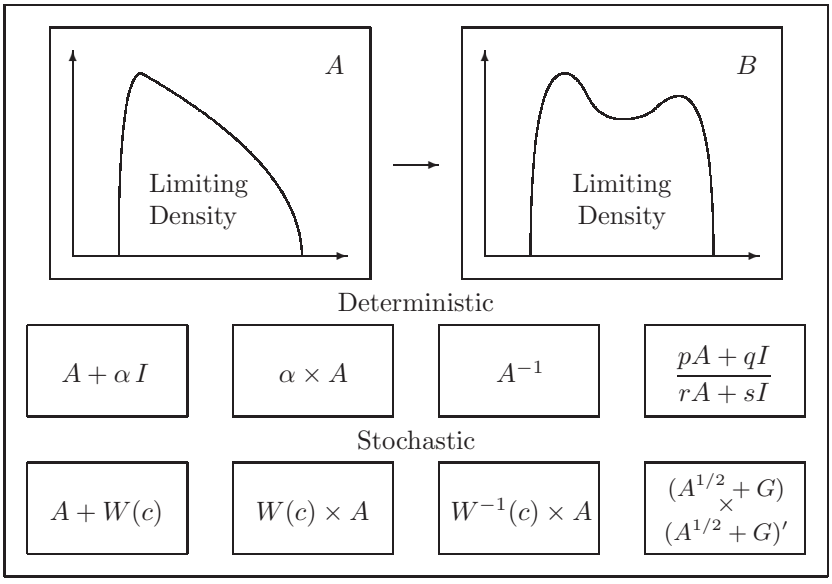


Figure 2: A random matrix calculator where a sequence of deterministic and stochastic operations performed on an algebraic random matrix sequence $\mathbf{A}_N$ produces an algebraic random matrix sequence $\mathbf{B}_N$. The limiting eigenvalue density and moments of a algebraic random matrix can be computed numerically, with the latter often in closed form.

Definition 1 (Algebraic random matrices)**.** *Let $F^A(x)$ denote the limiting eigenvalue distribution function of a sequence of random matrices $\mathbf{A}_N$. If a bivariate polynomial $L_{\mathrm{mz}}(m, z)$ exists such that*

$$m_A(z) = \int \frac{1}{x - z} dF^A(x) \qquad z \in \mathbb{C}^+ \setminus \mathbb{R}$$

is a solution of $L_{\mathrm{mz}}(m_A(z), z) = 0$ then $\mathbf{A}_N$ is said to be an algebraic random matrix. The density function $f_A := dF^A$ (in the distributional sense) is referred to as an algebraic density and we say that $\mathbf{A}_N \in \mathcal{M}_{\mathrm{alg}}$, the class of algebraic random matrices and $f_A \in \mathcal{P}_{alg}$, the class of algebraic distributions.

The utility of this, admittedly technical, definition comes from the fact that we are able to concretely specify the generators of this class. We illustrate this with a simple example. Let $\mathbf{G}$ be an $n \times m$ random matrix with i.i.d. standard normal entries with variance $1/m$. The matrix $\mathbf{W}(c) = \mathbf{G}\mathbf{G}'$ is the Wishart matrix parameterized by $c = n/m$. Let $\mathbf{A}$ be an arbitrary algebraic random matrix independent of $\mathbf{W}(c)$. Figure 2 identifies deterministic and stochastic operations that can be performed on $\mathbf{A}$ so that the resulting matrix is algebraic as well. The calculator analogy is apt because once we start with an algebraic random matrix, if we keep pushing away at the buttons we still get an algebraic random matrix whose limiting eigenvalue distribution is concretely computable using the algorithms developed in Section 6..

The algebraicity definition is important because everything we want to know about the limiting eigenvalue distribution of $\mathbf{A}$ is encoded in the bivariate polynomial $L^A_{\mathrm{mz}}(m, z)$. In this paper, we establish the algebraicity of each of the transformations in Figure 2 using the "hard" approach that we label as the *polynomial method* whereby we explicitly determine the operational law for the polynomial transformation $L^A_{\mathrm{mz}}(m, z) \mapsto$

$L^B_{\rm mz}(m,z)$ corresponding to the random matrix transformation $\mathbf{A} \mapsto \mathbf{B}$. This is in contrast to the "soft" approach taken in a recent paper by Anderson and Zeitouni [3, Section 6] where the algebraicity of Stieltjes transforms under hypotheses frequently fulfilled in RMT is proven using dimension theory for noetherian local rings. The catalogue of admissible transformations, the corresponding "hard" operational law and their software realization is found in Section 6.. This then allows us to calculate the eigenvalue distribution functions of a large class of algebraic random matrices that are generated from other algebraic random matrices. In the simple case involving Wigner and Wishart matrices considered earlier, the transformed polynomials were obtained by hand calculation. Along with the theory of algebraic random matrices we also develop a software realization that maps the entire catalog of transformations (see Tables 7 -9) into symbolic MATLAB code. Thus, for the same example, the sequence of commands:

```
>> syms m z
>> LmzA = m^2+z*m+1;
>> LmzB = m^2-(-2*z+1)*m+2;
>> LmzApB = AplusB(LmzA,LmzB);
>> LmzAtB = AtimesB(LmzA,LmzB);
```

could also have been used to obtain $L^{A+B}_{\rm mz}$ and $L^{AB}_{\rm mz}$. We note that the commands `AplusB` and `AtimesB` implicitly use the free convolution machinery (see Section 9.) to perform the said computation. To summarize, by defining the class of algebraic random matrices, we are able to extend the reach of infinite random matrix theory well beyond the special cases of matrices with Gaussian entries. The key idea is that by encoding probability densities as solutions of bivariate polynomial equations, and deriving the correct operational laws on this encoding, we can take advantage of powerful symbolic and numerical techniques to compute these densities and their associated moments.

1.2 Outline

This paper is organized as follows. We introduce various transform representations of the distribution function in Section 2.. We define algebraic distributions and the various manners in which they can be implicitly represented in 3. and describe how they may be algebraically manipulated in 4.. The class of algebraic random matrices is described in Section 5. where the theorems are stated and proved by obtaining the operational law on the bivariate polynomials summarized in Section 6.. Techniques for determining the density function of the limiting eigenvalue distribution function and the associated moments are discussed in Sections 7. and 8., respectively. We discuss the relevance of the polynomial method to computational free probability in Section 9., provide some applications in Section 10. and conclude with some open problems in Section 11..

2. Transform representations

We now describe the various ways in which transforms of the empirical distribution function can be encoded and manipulated.

2.1 The Stieltjes transform and some minor variations

The Stieltjes transform of the distribution function $F^A(x)$ is given by

$$m_A(z) = \int \frac{1}{x-z} dF^A(x) \qquad \text{for } z \in \mathbb{C}^+ \setminus \mathbb{R}. \tag{2.1}$$

The Stieltjes transform may be interpreted as the expectation

$$m_A(z) = E_x\left[\frac{1}{x-z}\right],$$

with respect to the random variable x with distribution function $F^A(x)$. Consequently, for any invertible function $h(x)$ continuous over the support of $dF^A(x)$, the Stieltjes transform $m_A(z)$ can also be written in terms of the distribution of the random variable $y = h(x)$ as

$$m_A(z) = E_x\left[\frac{1}{x-z}\right] = E_y\left[\frac{1}{h^{\langle -1\rangle}(y)-z}\right], \tag{2.2}$$

where $h^{\langle -1\rangle}(\cdot)$ is the inverse of $h(\cdot)$ with respect to composition i.e. $h(h^{\langle -1\rangle}(x)) = x$. Equivalently, for $y = h(x)$, we obtain the relationship

$$E_y\left[\frac{1}{y-z}\right] = E_x\left[\frac{1}{h(x)-z}\right]. \tag{2.3}$$

The well-known Stieltjes-Perron inversion formula [1]

$$f_A(x) \equiv dF^A(x) = \frac{1}{\pi}\lim_{\xi\to 0^+} \operatorname{Im} m_A(x+i\xi). \tag{2.4}$$

can be used to recover the probability density function $f_A(x)$ from the Stieltjes transform. Here and for the remainder of this thesis, the density function is assumed to be distributional derivative of the distribution function. In a portion of the literature on random matrices, the Cauchy transform is defined as

$$g_A(z) = \int \frac{1}{z-x} dF^A(x) \qquad for z \in \mathbb{C}^{-1} \setminus \mathbb{R}.$$

The Cauchy transform is related to the Stieltjes transform, as defined in (2.1), by

$$g_A(z) = -m_A(z). \tag{2.5}$$

2.2 The moment transform

When the probability distribution is compactly supported, the Stieltjes transform can also be expressed as the series expansion

$$m_A(z) = -\frac{1}{z} - \sum_{j=1}^{\infty} \frac{M_j^A}{z^{j+1}}, \tag{2.6}$$

about $z = \infty$, where $M_j^A := \int x^j dF^A(x)$ is the j-th moment. The ordinary moment generating function, $\mu_A(z)$, is the power series

$$\mu_A(z) = \sum_{j=0}^{\infty} M_j^A\, z^j, \tag{2.7}$$

with $M_0^A = 1$. The moment generating function, referred to as the moment transform, is related to the Stieltjes transform by

$$\mu_A(z) = -\frac{1}{z} m_A\left(\frac{1}{z}\right). \tag{2.8}$$

The Stieltjes transform can be expressed in terms of the moment transform as

$$m_A(z) = -\frac{1}{z} \mu_A\left(\frac{1}{z}\right). \tag{2.9}$$

The eta transform, introduced by Tulino and Verdù in [32], is a minor variation of the moment transform. It can be expressed in terms of the Stieltjes transform as

$$\eta_A(z) = \frac{1}{z} m_A\left(-\frac{1}{z}\right), \tag{2.10}$$

while the Stieltjes transform can be expressed in terms of the eta transform as

$$m_A(z) = -\frac{1}{z}\eta_A\left(-\frac{1}{z}\right). \tag{2.11}$$

2.3 The R transform

The R transform is defined in terms of the Cauchy transform as

$$r_A(z) = g_A^{\langle -1\rangle}(z) - \frac{1}{z}, \tag{2.12}$$

where $g_A^{\langle -1\rangle}(z)$ is the functional inverse of $g_A(z)$ with respect to composition. It will often be more convenient to use the expression for the R transform in terms of the Cauchy transform given by

$$r_A(g) = z(g) - \frac{1}{g}. \tag{2.13}$$

The R transform can be written as a power series whose coefficients K_j^A are known as the "free cumulants." For a combinatorial interpretation of free cumulants, see [28]. Thus the R transform is the (ordinary) free cumulant generating function

$$r_A(g) = \sum_{j=0}^{\infty} K_{j+1}^A\, g^j. \tag{2.14}$$

2.4 The S transform

The S transform is relatively more complicated. It is defined as

$$s_A(z) = \frac{1+z}{z}\,\Upsilon_A^{\langle -1\rangle}(z) \tag{2.15}$$

where $\Upsilon_A(z)$ can be written in terms of the Stieltjes transform $m_A(z)$ as

$$\Upsilon_A(z) = -\frac{1}{z}m_A(1/z) - 1. \tag{2.16}$$

This definition is quite cumbersome to work with because of the functional inverse in (2.15). It also places a technical restriction (to enable series inversion) that $M_1^A \neq 0$. We can, however, avoid this by expressing the S transform algebraically in terms of the Stieltjes transform as shown next. We first plug in $\Upsilon_A(z)$ into the left-hand side of (2.15) to obtain

$$s_A(\Upsilon_A(z)) = \frac{1+\Upsilon_A(z)}{\Upsilon_A(z)}\,z.$$

This can be rewritten in terms of $m_A(z)$ using the relationship in (2.16) to obtain

$$s_A(-\frac{1}{z}m(1/z) - 1) = \frac{z\,m(1/z)}{m(1/z)+z}$$

or, equivalently:

$$s_A(-z\,m(z) - 1) = \frac{m(z)}{z\,m(z)+1}. \tag{2.17}$$

We now define $y(z)$ in terms of the Stieltjes transform as $y(z) = -z\,m(z) - 1$. It is clear that $y(z)$ is an invertible function of $m(z)$. The right hand side of (2.17) can be rewritten in terms of $y(z)$ as

$$s_A(y(z)) = -\frac{m(z)}{y(z)} = \frac{m(z)}{z\,m(z)+1}. \tag{2.18}$$

Equation (2.18) can be rewritten to obtain a simple relationship between the Stieltjes transform and the S transform

$$m_A(z) = -y\, s_A(y). \tag{2.19}$$

Noting that $y = -z\, m(z) - 1$ and $m(z) = -y\, s_A(y)$ we obtain the relationship

$$y = z\, y\, s_A(y) - 1$$

or, equivalently

$$z = \frac{y+1}{y\, s_A(y)}. \tag{2.20}$$

3. Algebraic distributions

Notation 3.1 (Bivariate polynomial)**.** *Let* L_{uv} *denote a bivariate polynomial of degree* $\mathrm{D_u}$ *in* u *and* $\mathrm{D_v}$ *in* v *defined as*

$$L_{\mathrm{uv}} \equiv L_{\mathrm{uv}}(\cdot,\cdot) = \sum_{j=0}^{\mathrm{D_u}} \sum_{k=0}^{\mathrm{D_v}} c_{jk}\, u^j\, v^k = \sum_{j=0}^{\mathrm{D_u}} l_j(v)\, u^j. \tag{3.1}$$

The scalar coefficients c_{jk} *are real valued.*

The two letter subscripts for the bivariate polynomial L_{uv} provide us with a convention of which dummy variables we will use. We will generically use the first letter in the subscript to represent a transform of the density with the second letter acting as a mnemonic for the dummy variable associated with the transform. By consistently using the same pair of letters to denote the bivariate polynomial that encodes the transform and the associated dummy variable, this abuse of notation allows us to readily identify the encoding of the distribution that is being manipulated.

Remark 3.2 (Irreducibility)**.** *Unless otherwise stated it will be understood that* $L_{\mathrm{uv}}(u,v)$ *is "irreducible" in the sense that the conditions:*

- $l_0(v), \ldots, l_{\mathrm{D_u}}(v)$ *have no common factor involving* v*,*
- $l_{\mathrm{D_u}}(v) \neq 0$*,*
- $\mathrm{disc_L}(v) \neq 0$*,*

are satisfied, where $\mathrm{disc}_L(v)$ *is the discriminant of* $L_{\mathrm{uv}}(u,v)$ *thought of as a polynomial in* v*.*

We are particularly focused on the solution "curves," $u_1(v), \ldots, u_{\mathrm{D_u}}(v)$, *i.e.*,

$$L_{\mathrm{uv}}(u,v) = l_{\mathrm{D_u}}(v) \prod_{i=1}^{D_{\mathrm{u}}} (u - u_i(v))\,.$$

Informally speaking, when we refer to the bivariate polynomial equation $L_{\mathrm{uv}}(u,v) = 0$ with solutions $u_i(v)$ we are actually considering the equivalence class of rational functions with this set of solution curves.

Remark 3.3 (Equivalence class)**.** *The equivalence class of* $L_{\mathrm{uv}}(u,v)$ *may be characterized as functions of the form* $L_{\mathrm{uv}}(u,v)g(v)/h(u,v)$ *where* h *is relatively prime to* $L_{\mathrm{uv}}(u,v)$ *and* $g(v)$ *is not identically* 0*.*

A few technicalities (such as poles and singular points) that will be catalogued later in Section 6. remain, but this is sufficient for allowing us to introduce rational transformations of the arguments and continue to use the language of polynomials.

Definition 3.4 (Algebraic distributions)**.** *Let $F(x)$ be a probability distribution function and $f(x)$ be its distributional derivative (here and henceforth). Consider the Stieltjes transform $m(z)$ of the distribution function, defined as*

$$m(z) = \int \frac{1}{x - z} dF(x) \qquad \text{for } z \in \mathbb{C}^+ \setminus \mathbb{R}. \tag{3.2}$$

If there exists a bivariate polynomial L_{mz} such that $L_{mz}(m(z), z) = 0$ then we refer to $F(x)$ as algebraic (probability) distribution function, $f(x)$ as an algebraic (probability) density function and say the $f \in \mathcal{P}_{alg}$. Here $\mathcal{P}_{alg}$ denotes the class of algebraic (probability) distributions.

Definition 3.5 (Atomic distribution)**.** *Let $F(x)$ be a probability distribution function of the form*

$$F(x) = \sum_{i=1}^{K} p_i \, \mathbb{I}_{[\lambda_i, \infty)},$$

where the K atoms at $\lambda_i \in \mathbb{R}$ have (non-negative) weights p_i subject to $\sum_i p_i = 1$ and $\mathbb{I}_{[x,\infty)}$ is the indicator (or characteristic) function of the set $[x, \infty)$. We refer to $F(x)$ as an atomic (probability) distribution function. Denoting its distributional derivative by $f(x)$, we say that $f(x) \in \mathcal{P}_{atom}$. Here $\mathcal{P}_{atom}$ denotes the class of atomic distributions.

Example 3.6. *An atomic probability distribution, as in Definition 3.5, has a Stieltjes transform*

$$m(z) = \sum_{i=1}^{K} \frac{p_i}{\lambda_i - z}$$

which is the solution of the equation $L_{mz}(m, z) = 0$ where

$$L_{mz}(m, z) \equiv \prod_{i=1}^{K} (\lambda_i - z)\, m - \sum_{i=1}^{K} \prod_{\substack{j \neq i \\ j=1}}^{K} p_i (\lambda_j - z).$$

Hence it is an algebraic distribution; consequently $\mathcal{P}_{atom} \subset \mathcal{P}_{alg}$.

Example 3.7. *The Cauchy distribution whose density*

$$f(x) = \frac{1}{\pi(x^2 + 1)},$$

has a Stieltjes transform $m(z)$ which is the solution of the equation $L_{mz}(m, z) = 0$ where

$$L_{mz}(m, z) \equiv \left(z^2 + 1\right) m^2 + 2\, z\, m + 1.$$

Hence it is an algebraic distribution.

It is often the case that the probability density functions of algebraic distributions, according to our definition, will also be algebraic functions themselves. We conjecture that this is a necessary but not sufficient condition. We show that it is not sufficient by providing the counter-example below.

Counter-example 3.8. *Consider the quarter-circle distribution with density function*

$$f(x) = \frac{\sqrt{4-x^2}}{\pi} \qquad \text{for } x \in [0,2].$$

Its Stieltjes transform :

$$m(z) = -\frac{4 - 2\sqrt{-z^2+4}\ln\left(-\frac{2+\sqrt{-z^2+4}}{z}\right) + z\pi}{2\pi},$$

is clearly not an algebraic function. Thus $f(x) \notin \mathcal{P}_{alg}$.

3.1 Implicit representations of algebraic distributions

We now define six interconnected bivariate polynomials denoted by $L_{\rm mz}$, $L_{\rm gz}$, $L_{\rm rg}$, $L_{\rm sy}$, $L_{\mu z}$, and $L_{\eta z}$. We assume that $L_{\rm uv}(u,v)$ is an irreducible bivariate polynomial of the form in (3.1). The main protagonist of the transformations we consider is the bivariate polynomial $L_{\rm mz}$ which implicitly defines the Stieltjes transform $m(z)$ via the equation $L_{\rm mz}(m,z) = 0$. Starting off with this polynomial we can obtain the polynomial $L_{\rm gz}$ using the relationship in (2.5) as

$$L_{\rm gz}(g,z) = L_{\rm mz}(-g,z). \tag{3.3}$$

Perhaps we should explain our abuse of notation once again, for the sake of clarity. Given any one polynomial, all the other polynomials can be obtained. The two letter subscripts not only tell us which of the six polynomials we are focusing on, it provides a convention of which dummy variables we will use. The first letter in the subscript represents the transform; the second letter is a mnemonic for the variable associated with the transform that we use consistently in the software based on this framework. With this notation in mind, we can obtain the polynomial $L_{\rm rg}$ from $L_{\rm gz}$ using (2.13) as

$$L_{\rm rg}(r,g) = L_{\rm gz}\left(g, r + \frac{1}{g}\right). \tag{3.4}$$

Similarly, we can obtain the bivariate polynomial $L_{\rm sy}$ from $L_{\rm mz}$ using the expressions in (2.19) and (2.20) to obtain the relationship

$$L_{\rm sy} = L_{\rm mz}\left(-y\,s, \frac{y+1}{sy}\right). \tag{3.5}$$

Based on the transforms discussed in Sectin 2., we can derive transformations between additional pairs of bivariate polynomials represented by the bidirectional arrows in Figure 3 and listed in the third column of Table 3. Specifically, the expressions in (2.8) and (2.11) can be used to derive the transformations between $L_{\rm mz}$ and $L_{\mu z}$ and $L_{\rm mz}$ and $L_{\eta z}$ respectively. The fourth column of Table 3 lists the MATLAB function, implemented using its MAPLE based Symbolic Toolbox, corresponding to the bivariate polynomial transformations represented in Figure 3. In the MATLAB functions, the function `irreducLuv(u,v)` listed in Table 1 ensures that the resulting bivariate polynomial is irreducible by clearing the denominator and making the resulting polynomial square free.

Example: Consider an atomic probability distribution with

$$F(x) = 0.5\,\mathbb{I}_{[0,\infty)} + 0.5\,\mathbb{I}_{[1,\infty)}, \tag{3.6}$$

whose Stieltjes transform

$$m(z) = \frac{0.5}{0-z} + \frac{0.5}{1-z},$$

is the solution of the equation

$$m(0-z)(1-z) - 0.5(1-2z) = 0,$$

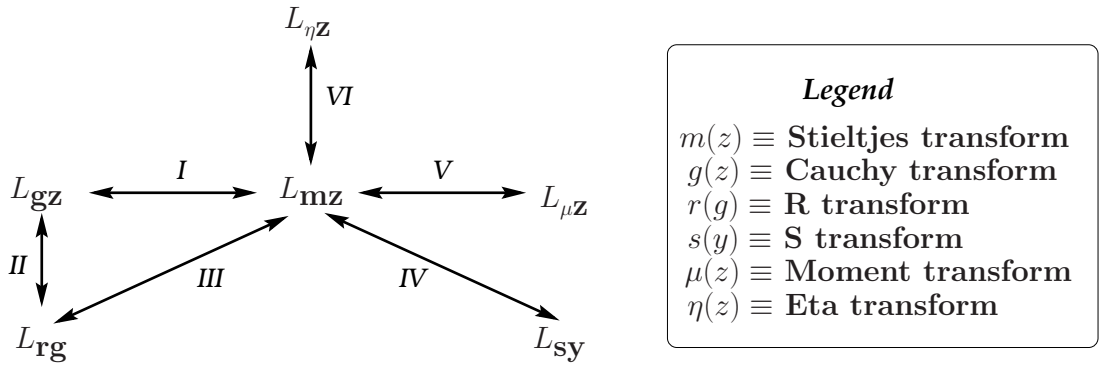


Figure 3: The six interconnected bivariate polynomials; transformations between the polynomials, indicated by the labelled arrows, are given in Table 3.

or equivalently, the solution of the equation $L_{\mathrm{mz}}(m, z) = 0$ where

$$L_{\mathrm{mz}}(m, z) \equiv m(2\,z^2 - 2\,z) - (1 - 2z). \tag{3.7}$$

We can obtain the bivariate polynomial $L_{\mathrm{gz}}(g, z)$ by applying the transformation in (3.3) to the bivariate polynomial L_{mz} given by (3.7) so that

$$L_{\mathrm{gz}}(g, z) = -g(2\,z^2 - 2\,z) - (1 - 2z). \tag{3.8}$$

Similarly, by applying the transformation in (3.4) we obtain

$$L_{\mathrm{rg}}(r, g) = -g\left(2\left(r + \frac{1}{g}\right) - 2\left(r + \frac{1}{g}\right)^2\right) - \left(1 - 2\left(r + \frac{1}{g}\right)\right). \tag{3.9}$$

which, on clearing the denominator and invoking the equivalence class representation of our polynomials (see Remark 3.3), gives us the irreducible bivariate polynomial

$$L_{\mathrm{rg}}(r, g) = -1 + 2\,gr^2 + (2 - 2\,g)\,r. \tag{3.10}$$

By applying the transformation in (3.5) to the bivariate polynomial L_{mz}, we obtain

$$L_{\mathrm{sy}} \equiv (-s\,y)\left(2\,\frac{y+1}{sy} - 2\,\left(\frac{y+1}{sy}\right)^2\right) - \left(1 - 2\frac{y+1}{sy}\right)$$

which on clearing the denominator gives us the irreducible bivariate polynomial

$$L^A_{\mathrm{sy}}(s, y) = (1 + 2\,y)\,s - 2 - 2\,y. \tag{3.11}$$

Table 2 tabulates the six bivariate polynomial encodings in Figure 3 for the distribution in (3.6), the semi-circle distribution for Wigner matrices and the Marčenko-Pastur distribution for Wishart matrices.

Procedure	MATLAB Code
	`function Luv = irreducLuv(Luv,u,v)`
Simplify and clear the denominator	`L = numden(simplify(expand(Luv)));`
	`L = Luv / maple('gcd',L,diff(L,u));`
Make square free	`L = simplify(expand(L));`
	`L = Luv / maple('gcd',L,diff(L,v));`
Simplify	`Luv = simplify(expand(L));`

Table 1: Making L_{uv} irreducible.

(a) The atomic distribution in (3.6).

L	Bivariate Polynomials
L_{mz}	$m(2\,z^2-2\,z)-(1-2z)$
L_{gz}	$-g(2\,z^2-2\,z)-(1-2z)$
L_{rg}	$-1+2\,gr^2+(2-2\,g)\,r$
L_{sy}	$(1+2\,y)\,s-2-2\,y$
$L_{\mu\mathrm{z}}$	$(-2+2\,z)\,\mu+2-z$
$L_{\eta\mathrm{z}}$	$(2\,z+2)\,\eta-2-z$

(b) The Marčenko-Pastur distribution.

L	Bivariate Polynomials
L_{mz}	$czm^2-(1-c-z)\,m+1$
L_{gz}	$czg^2+(1-c-z)\,g+1$
L_{rg}	$(cg-1)\,r+1$
L_{sy}	$(cy+1)\,s-1$
$L_{\mu\mathrm{z}}$	$\mu^2zc-(zc+1-z)\,\mu+1$
$L_{\eta\mathrm{z}}$	$\eta^2zc+(-zc+1-z)\,\eta-1$

(c) The semi-circle distribution.

L	Bivariate polynomials
L_{mz}	$m^2+m\,z+1$
L_{gz}	$g^2-g\,z+1$
L_{rg}	$r-g$
L_{sy}	$s^2\,y-1$
$L_{\mu\mathrm{z}}$	$\mu^2z^2-\mu+1$
$L_{\eta\mathrm{z}}$	$z^2\eta^2-\eta+1$

Table 2: Bivariate polynomial representations of some algebraic distributions.

Label	***Conversion***	***Transformation***	***MATLAB Code***
I	$L_{\mathbf{mz}} \rightleftharpoons L_{\mathbf{gz}}$	$L_{\mathbf{mz}} = L_{\mathbf{gz}}(-m, z)$	function Lmz = Lgz2Lmz(Lgz) syms m g z Lmz = subs(Lgz,g,-m);
		$L_{\mathbf{gz}} = L_{\mathbf{mz}}(-g, z)$	function Lgz = Lmz2Lgz(Lmz) syms m g z Lgz = subs(Lmz,m,-g);
II	$L_{\mathbf{gz}} \rightleftharpoons L_{\mathbf{rg}}$	$L_{\mathbf{gz}} = L_{\mathbf{rg}}(z - \frac{1}{g}, z)$	function Lgz = Lrg2Lgz(Lrg) syms r g z Lgz = subs(Lrg,r,z-1/g); Lgz = irreducLuv(Lgz,g,z);
		$L_{\mathbf{rg}} = L_{\mathbf{gz}}(g, r + \frac{1}{g})$	function Lrg = Lgz2Lrg(Lgz) syms r g z Lrg = subs(Lgz,g,r+1/g); Lrg = irreducLuv(Lrg,r,g);
III	$L_{\mathbf{mz}} \rightleftharpoons L_{\mathbf{rg}}$	$L_{\mathbf{mz}} \rightleftharpoons L_{\mathbf{gz}} \rightleftharpoons L_{\mathbf{rg}}$	function Lmz = Lrg2Lmz(Lrg) syms m z r g Lgz = Lrg2Lgz(Lrg); Lmz = Lgz2Lmz(Lgz);
			function Lrg = Lmz2Lrg(Lmz) syms m z r g Lgz = Lmz2Lgz(Lmz); Lrg = Lgz2Lrg(Lgz);
IV	$L_{\mathbf{mz}} \rightleftharpoons L_{\mathbf{sy}}$	$L_{\mathbf{mz}} = L_{\mathbf{sy}}(\frac{m}{z\,m+1}, -z\,m - 1)$	function Lmz = Lsy2Lmz(Lsy) syms m z s y Lmz = subs(Lsy,s,m/(z*m+1)); Lmz = subs(Lmz,y,-z*m-1); Lmz = irreducLuv(Lmz,m,z);
		$L_{\mathbf{sy}} = L_{\mathbf{mz}}(-y\,s, \frac{y+1}{s\,y})$	function Lsy = Lmz2Lsy(Lmz) syms m z s y Lsy = subs(Lmz,m,-y*s); Lsy = subs(Lsy,z,(y+1)/y/s); Lsy = irreducLuv(Lsy,s,y);
V	$L_{\mathbf{mz}} \rightleftharpoons L_{\mu\mathbf{z}}$	$L_{\mathbf{mz}} = L_{\mu\mathbf{z}}(-m\,z, \frac{1}{z})$	function Lmz = Lmyuz2Lmz(Lmyuz) syms m myu z Lmz = subs(Lmyuz,z,1/z); Lmz = subs(Lmz,myu,-m*z); Lmz = irreducLuv(Lmz,m,z);
		$L_{\mu\mathbf{z}} = L_{\mathbf{mz}}(-\mu\,z, \frac{1}{z})$	function Lmyuz = Lmz2Lmyuz(Lmz) syms m myu z Lmyuz = subs(Lmz,z,1/z); Lmyuz = subs(Lmyuz,m,-myu*z); Lmyuz = irreducLuv(Lmyuz,myu,z);
VI	$L_{\mathbf{mz}} \rightleftharpoons L_{\eta\mathbf{z}}$	$L_{\mathbf{mz}} = L_{\eta\mathbf{z}}(-z\,m, -\frac{1}{z})$	function Lmz = Letaz2Lmz(Letaz) syms m eta z Lmz = subs(Letaz,z,-1/z); Lmz = subs(Lmz,eta,-z*m); Lmz = irreducLuv(Lmz,m,z);
		$L_{\eta\mathbf{z}} = L_{\mathbf{mz}}(z\,\eta, -\frac{1}{z})$	function Letaz = Lmz2Letaz(Lmz) syms m eta z Letaz = subs(Lmz,z,-1/z); Letaz = subs(Letaz,m,z*eta); Letaz = irreducLuv(Letaz,eta,z);

Table 3: Transformations between the different bivariate polynomials. As a guide to MATLAB notation, the command `syms` declares a variable to be symbolic while the command `subs` symbolically substitutes every occurrence of the second argument in the first argument with the third argument. Thus, for example, the command `y=subs(x-a,a,10)` will yield the output `y=x-10` if we have previously declared x and a to be symbolic using the command `syms x a`.

4. Algebraic operations on algebraic functions

Algebraic functions are closed under addition and multiplication. Hence we can add (or multiply) two algebraic functions and obtain another algebraic function. We show, using purely matrix theoretic arguments, how to obtain the polynomial equation whose solution is the sum (or product) of two algebraic functions without ever actually computing the individual functions. In Section 4.2, we interpret this computation using the concept of resultants [31] from elimination theory. These tools will feature prominently in Section 5. when we encode the transformations of the random matrices as algebraic operations on the appropriate form of the bivariate polynomial that encodes their limiting eigenvalue distributions.

4.1 Companion matrix based computation

Definition 4.1 (Companion Matrix)**.** *The companion matrix* $\mathbf{C}_{a(x)}$ *to a monic polynomial*

$$a(x) \equiv a_0 + a_1\, x + \ldots + a_{n-1}\, x^{n-1} + x^n$$

is the $n \times n$ *square matrix*

$$\mathbf{C}_{a(x)} = \begin{bmatrix} 0 & \ldots & \ldots & \ldots & -a_0 \\ 1 & \ldots & \ldots & \ldots & -a_1 \\ 0 & \ddots & & & -a_2 \\ \vdots & & \ddots & & \vdots \\ 0 & \ldots & \ldots & 1 & -a_{n-1} \end{bmatrix}$$

with ones on the sub-diagonal and the last column given by the negative coefficients of $a(x)$*.*

Remark 4.2. *The eigenvalues of the companion matrix are the solutions of the equation* $a(x) = 0$*. This is intimately related to the observation that the characteristic polynomial of the companion matrix equals* $a(x)$*, i.e.,*

$$a(x) = \det(x\, \mathbf{I}_n - \mathbf{C}_{a(x)}).$$

Consider the bivariate polynomial L_{uv} as in (3.1). By treating it as a polynomial in u whose coefficients are polynomials in v, *i.e.*, by rewriting it as

$$L_{\mathrm{uv}}(u, v) \equiv \sum_{j=0}^{\mathrm{D_u}} l_j(v)\, u^j, \tag{4.1}$$

we can create a companion matrix $\mathbf{C}^{\mathrm{u}}_{\mathrm{uv}}$ whose characteristic polynomial as a function of u is the bivariate polynomial L_{uv}. The companion matrix $\mathbf{C}^{\mathrm{u}}_{\mathrm{uv}}$ is the $\mathrm{D_u} \times \mathrm{D_u}$ matrix in Table 4.

$\mathbf{C}^{\mathrm{u}}_{\mathrm{uv}}$	MATLAB code
$\begin{bmatrix} 0 & \ldots & \ldots & \ldots & -l_0(v)/l_{\mathrm{D_u}}(v) \\ 1 & \ldots & \ldots & \ldots & -l_1(v)/l_{\mathrm{D_u}}(v) \\ 0 & \ddots & & & -l_2(v)/l_{\mathrm{D_u}}(v) \\ \vdots & & \ddots & & \vdots \\ 0 & \ldots & \ldots & 1 & -l_{\mathrm{D_u}-1}(v)/l_{\mathrm{D_u}}(v) \end{bmatrix}$	function Cu = Luv2Cu(Luv,u) Du = double(maple('degree',Luv,u)); LDu = maple('coeff',Luv,u,Du); Cu = sym(zeros(Du))+ .. +diag(ones(Du-1,1),-1)); for Di = 0:Du-1 LtuDi = maple('coeff',Lt,u,Di); Cu(Di+1,Du) = -LtuDi/LDu; end

Table 4: The companion matrix $\mathbf{C}^{\mathrm{u}}_{\mathrm{uv}}$, with respect to u, of the bivariate polynomial L_{uv} given by (4.1).

Remark 4.3. *Analogous to the univariate case, the characteristic polynomial of* $\mathbf{C}^{\mathrm{u}}_{\mathrm{uv}}$ *is* $\det(u\,\mathbf{I} - \mathbf{C}^{\mathrm{u}}_{\mathrm{uv}}) = L_{\mathrm{uv}}(u,v)/l_{\mathrm{D_u}}(v)^{\mathrm{D_u}}$. *Since* $l_{\mathrm{D_u}}(v)$ *is not identically zero, we say that* $\det(u\,\mathbf{I} - \mathbf{C}^{\mathrm{u}}_{\mathrm{uv}}) = L_{\mathrm{uv}}(u,v)$ *where the equality is understood to be with respect to the equivalence class of* L_{uv} *as in Remark 3.3. The eigenvalues of* $\mathbf{C}^{\mathrm{u}}_{\mathrm{uv}}$ *are the solutions of the algebraic equation* $L_{\mathrm{uv}}(u,v) = 0$*; specifically, we obtain the algebraic function* $u(v)$.

Definition 4.4 (Kronecker product)**.** *If* $\mathbf{A}_m$ *(with entries* a_{ij}*) is an* $m \times m$ *matrix and* $\mathbf{B}_n$ *is an* $n \times n$ *matrix then the Kronecker (or tensor) product of* $\mathbf{A}_m$ *and* $\mathbf{B}_n$*, denoted by* $\mathbf{A}_m \otimes \mathbf{B}_n$*, is the* $mn \times mn$ *matrix defined as:*

$$\mathbf{A}_m \otimes \mathbf{B}_n = \begin{bmatrix} a_{11}\mathbf{B}_n & \dots & a_{1n}\mathbf{B}_n \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B}_n & \dots & a_{mn}\mathbf{B}_n \end{bmatrix}$$

Lemma 4.5. *If* α_i *and* β_j *are the eigenvalues of* $\mathbf{A}_m$ *and* $\mathbf{B}_n$ *respectively, then*

1. $\alpha_i + \beta_j$ *is an eigenvalue of* $(\mathbf{A}_m \otimes \mathbf{I}_n) + (\mathbf{I}_m \otimes \mathbf{B}_n)$,
2. $\alpha_i\,\beta_j$ *is an eigenvalue of* $\mathbf{A}_m \otimes \mathbf{B}_n$,

for $i = 1, \dots, m$, $j = 1, \dots, n$.

Proof. The statements are proved in [16, Theorem 4.4.5] and [16, Theorem 4.2.12]. □

Proposition 4.6. *Let* $u_1(v)$ *be a solution of the algebraic equation* $L^1_{\mathrm{uv}}(u,v) = 0$*, or equivalently an eigenvalue of the* $D^1_{\mathrm{u}} \times D^1_{\mathrm{u}}$ *companion matrix* $\mathbf{C}^{\mathrm{u}_1}_{\mathrm{uv}}$*. Let* $u_2(v)$ *be a solution of the algebraic equation* $L^2_{\mathrm{uv}}(u,v) = 0$*, or equivalently an eigenvalue of the* $D^2_{\mathrm{u}} \times D^2_{\mathrm{u}}$ *companion matrix* $\mathbf{C}^{\mathrm{u}_2}_{\mathrm{uv}}$*. Then*

1. $u_3(v) = u_1(v) + u_2(v)$ *is an eigenvalue of the matrix* $\mathbf{C}^{\mathrm{u}_3}_{\mathrm{uv}} = \left(\mathbf{C}^{\mathrm{u}_1}_{\mathrm{uv}} \otimes \mathbf{I}_{D^2_{\mathrm{u}}}\right) + \left(\mathbf{I}_{D^1_{\mathrm{u}}} \otimes \mathbf{C}^{\mathrm{u}_2}_{\mathrm{uv}}\right)$,
2. $u_3(v) = u_1(v)u_2(v)$ *is an eigenvalue of the matrix* $\mathbf{C}^{\mathrm{u}_3}_{\mathrm{uv}} = \mathbf{C}^{\mathrm{u}_1}_{\mathrm{uv}} \otimes \mathbf{C}^{\mathrm{u}_2}_{\mathrm{uv}}$.

Equivalently $u_3(v)$ *is a solution of the algebraic equation* $L^3_{\mathrm{uv}} = 0$ *where* $L^3_{\mathrm{uv}} = \det(u\,\mathbf{I} - \mathbf{C}^{\mathrm{u}_3}_{\mathrm{uv}})$.

Proof. This follows directly from Lemma 4.5. □ We represent the binary addition and multiplication operators on the space of algebraic functions by the symbols $\boxplus_{\mathrm{u}}$ and $\boxtimes_{\mathrm{u}}$ respectively. We define addition and multiplication as in Table 5 by applying Proposition 4.6. Note that the subscript 'u' in $\boxplus_{\mathrm{u}}$ and $\boxtimes_{\mathrm{u}}$ provides us with an indispensable convention of which dummy variable we are using. Table 6 illustrates the $\boxplus$ and $\boxtimes$ operations on a pair of bivariate polynomials and underscores the importance of the symbolic software developed. The $(\mathrm{D_u}+1) \times (\mathrm{D_v}+1)$ matrix $\mathbf{T}_{\mathrm{uv}}$ lists only the coefficients c_{ij} for the term $u^i\,v^j$ in the polynomial $L_{\mathrm{uv}}(u,v)$. Note that the indexing for i and j starts with zero.

4.2 Resultants based computation

Addition (and multiplication) of algebraic functions produces another algebraic function. We now demonstrate how the concept of resultants from elimination theory can be used to obtain the polynomial whose zero set is the required algebraic function.

Definition 4.7 (Resultant)**.** *Given a polynomial*

$$a(x) \equiv a_0 + a_1\,x + \ldots + a_{n-1}\,x^{n-1} + a_n x^n$$

of degree n *with roots* α_i*, for* $i = 1, \dots, n$ *and a polynomial*

$$b(x) \equiv b_0 + b_1\,x + \ldots + b_{m-1}\,x^{m-1} + b_m x^m$$

Operation: $L^1_{\mathrm{uv}}, L^2_{\mathrm{uv}} \longmapsto L^3_{\mathrm{uv}}$	MATLAB Code

$L^3_{\mathrm{uv}} = L^1_{\mathrm{uv}} \boxplus_{\mathrm{u}} L^2_{\mathrm{uv}} \equiv \det(u\,\mathbf{I} - \mathbf{C}^{\mathrm{u}_3}_{\mathrm{uv}})$, where $\mathbf{C}^{\mathrm{u}_3}_{\mathrm{uv}} = \begin{cases} 2\,\mathbf{C}^{\mathrm{u}_1}_{\mathrm{uv}} & \text{if } L^1_{\mathrm{uv}} = L^2_{\mathrm{uv}}, \\ (\mathbf{C}^{\mathrm{u}_1}_{\mathrm{uv}} \otimes \mathbf{I}_{D^1_{\mathrm{u}}}) + (\mathbf{I}_{D^1_{\mathrm{u}}} \otimes \mathbf{C}^{\mathrm{u}_2}_{\mathrm{uv}}) & \text{otherwise.} \end{cases}$	function Luv3 = L1plusL2(Luv1,Luv2,u)
	Cu1 = Luv2Cu(Luv1,u);
	if (Luv1 == Luv2)
	  Cu3 = 2*Cu1;
	else
	  Cu2 = Luv2Cu(Luv2,u);
	  Cu3 = kron(Cu1,eye(length(Cu2))) + ..
	  +kron(eye(length(Cu1)),Cu2);
	end
	Luv3 = det(u*eye(length(Cu3))-Cu3);

$L^3_{\mathrm{uv}} = L^1_{\mathrm{uv}} \boxtimes_{\mathrm{u}} L^2_{\mathrm{uv}} \equiv \det(u\,\mathbf{I} - \mathbf{C}^{\mathrm{u}_3}_{\mathrm{uv}})$, where $\mathbf{C}^{\mathrm{u}_3}_{\mathrm{uv}} = \begin{cases} \mathbf{C}^{\mathrm{u}_3}_{\mathrm{uv}} = (\mathbf{C}^{\mathrm{u}_1}_{\mathrm{uv}})^2 & \text{if } L^1_{\mathrm{uv}} = L^2_{\mathrm{uv}}, \\ \mathbf{C}^{\mathrm{u}_3}_{\mathrm{uv}} = \mathbf{C}^{\mathrm{u}_1}_{\mathrm{uv}} \otimes \mathbf{C}^{\mathrm{u}_2}_{\mathrm{uv}} & \text{otherwise.} \end{cases}$	function Luv3 = L1timesL2(Luv1,Luv2,u)
	Cu1 = Luv2Cu(Luv1,u);
	if (Luv1 == Luv2)
	  Cu3 = Cû2;
	else
	  Cu2 = Luv2Cu(Luv2,u);
	  Cu3 = kron(Cu1,Cu2);
	end
	Luv3 = det(u*eye(length(Cu3))-Cu3);

Table 5: Formal and computational description of the $\boxplus_{\mathrm{u}}$ and $\boxtimes_{\mathrm{u}}$ operators acting on the bivariate polynomials $L^1_{\mathrm{uv}}(u,v)$ and $L^2_{\mathrm{uv}}(u,v)$ where $\mathbf{C}^{\mathrm{u}_1}_{\mathrm{uv}}$ and $\mathbf{C}^{\mathrm{u}_2}_{\mathrm{uv}}$ are their corresponding companion matrices constructed as in Table 4 and $\otimes$ is the matrix Kronecker product.

of degree m with roots β_j, for $j = 1, \ldots, m$, the resultant is defined as

$$\mathrm{Res}_{\mathrm{x}}\left(a(x)\,,\, b(x)\right) = a_n^m\, b_m^n \prod_{i=1}^{n} \prod_{j=1}^{m} (\beta_j - \alpha_i).$$

From a computational standpoint, the resultant can be directly computed from the coefficients of the polynomials itself. The computation involves the formation of the Sylvester matrix and exploiting an identity that relates the determinant of the Sylvester matrix to the resultant.

Definition 4.8 (Sylvester matrix)**.** *Given polynomials $a(x)$ and $b(x)$ with degree n and m respectively and coefficients as in Definition 4.7, the Sylvester matrix is the $(n+m) \times (n+m)$ matrix*

$$\mathbf{S}(a,b) = \begin{bmatrix} a_n & 0 & \cdots & 0 & 0 & b_m & 0 & \cdots & 0 & 0 \\ a_{n-1} & a_n & \cdots & 0 & 0 & b_{m-1} & b_m & \cdots & 0 & 0 \\ \ldots & \ldots & \cdots & \ldots & \ldots & \ldots & \ldots & \cdots & \ldots & \ldots \\ 0 & 0 & \cdots & a_0 & a_1 & 0 & 0 & \cdots & b_0 & b_1 \\ 0 & 0 & \cdots & 0 & a_0 & 0 & 0 & \cdots & 0 & b_0 \end{bmatrix}$$

Proposition 4.9. *The resultant of two polynomials $a(x)$ and $b(x)$ is related to the determinant of the Sylvester matrix by*

$$\det(\mathbf{S}(a,b)) = \mathrm{Res}_{\mathrm{x}}\left(a(x)\,,\, b(x)\right)$$

PROOF. This identity can be proved using standard linear algebra arguments. A proof may be found in [2]. □ For our purpose, the utility of this definition is that the $\boxplus_{\mathrm{u}}$ and $\boxtimes_{\mathrm{u}}$ operations can be expressed in terms of resultants. Suppose we are given two bivariate polynomials L^1_{uv} and L^2_{uv}. By using the definition of the resultant and treating the bivariate polynomials as polynomials in u whose coefficients are polynomials in v, we obtain the identities

$$\boxed{L^3_{\mathrm{uv}}(t,v) = L^1_{\mathrm{uv}} \boxplus_{\mathrm{u}} L^2_{\mathrm{uv}} \equiv \mathrm{Res}_{\mathrm{u}}\left(L^1_{\mathrm{uv}}(t-u,v)\,,\, L^2_{\mathrm{uv}}(u,v)\right),} \tag{4.2}$$

$L_{\rm uv}$	$\mathbf{T}_{\rm uv}$	$\mathbf{C}^{\rm u}_{\rm uv}$	$\mathbf{C}^{\rm v}_{\rm uv}$
$L^1_{\rm uv} \equiv u^2 v + u\,(1-v) + v^2$	$\begin{array}{c} \\ 1 \\ u \\ u^2 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & v & v^2 \end{array} \\ \begin{bmatrix} \cdot & \cdot & 1 \\ 1 & -1 & \cdot \\ \cdot & 1 & \cdot \end{bmatrix} \end{array}$	$\begin{bmatrix} 0 & -v \\ 1 & \dfrac{-1+v}{v} \end{bmatrix}$	$\begin{bmatrix} 0 & -u \\ 1 & -u^2+u \end{bmatrix}$
$L^2_{\rm uv} \equiv u^2\left(v^2 - 3\,v + 1\right) + u\,(1+v) + v^2$	$\begin{array}{c} \\ 1 \\ u \\ u^2 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & v & v^2 \end{array} \\ \begin{bmatrix} \cdot & \cdot & 1 \\ 1 & 1 & \cdot \\ 1 & -3 & 1 \end{bmatrix} \end{array}$	$\begin{bmatrix} 0 & \dfrac{-v^2}{v^2-3\,v+1} \\ 1 & \dfrac{-1-v}{v^2-3\,v+1} \end{bmatrix}$	$\begin{bmatrix} 0 & \dfrac{-u^2-u}{u^2+1} \\ 1 & \dfrac{3\,u^2-u}{u^2+1} \end{bmatrix}$

$L^1_{\rm uv} \boxplus_{\mathbf{u}} L^2_{\rm uv}$

$$\begin{array}{c} \\ 1 \\ u \\ u^2 \\ u^3 \\ u^4 \end{array} \begin{array}{c} \begin{array}{ccccccccc} 1 & v & v^2 & v^3 & v^4 & v^5 & v^6 & v^7 & v^8 \end{array} \\ \begin{bmatrix} \cdot & \cdot & 2 & -6 & 11 & -10 & 18 & -8 & 1 \\ 2 & \cdot & 2 & -8 & 4 & \cdot & \cdot & \cdot & \cdot \\ 5 & \cdot & 1 & -4 & 2 & \cdot & \cdot & \cdot & \cdot \\ 4 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \end{array}$$

$L^1_{\rm uv} \boxtimes_{\mathbf{u}} L^2_{\rm uv}$

$$\begin{array}{c} \\ 1 \\ u \\ u^2 \\ u^3 \\ u^4 \end{array} \begin{array}{c} \begin{array}{ccccccccccccccc} 1 & v & v^2 & v^3 & v^4 & v^5 & v^6 & v^7 & v^8 & v^9 & v^{10} & v^{11} & v^{12} & v^{13} & v^{14} \end{array} \\ \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & -6 & 11 & -6 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & -1 & 3 & \cdot & -3 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & -4 & 10 & -6 & 7 & -2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \end{array}$$

$L^1_{\rm uv} \boxplus_{\mathbf{v}} L^2_{\rm uv}$

$$\begin{array}{c} \\ 1 \\ u \\ u^2 \\ u^3 \\ u^4 \\ u^5 \\ u^6 \\ u^7 \\ u^8 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & v & v^2 & v^3 & v^4 \end{array} \\ \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & 4 & \cdot & \cdot \\ \cdot & \cdot & 1 & -4 & \cdot \\ \cdot & -8 & 6 & \cdot & \cdot \\ 1 & -2 & 3 & \cdot & \cdot \\ 8 & -12 & \cdot & \cdot & \cdot \\ 3 & 2 & \cdot & \cdot & \cdot \\ 2 & \cdot & \cdot & \cdot & \cdot \\ -1 & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \end{array}$$

$L^2_{\rm uv} \boxtimes_{\mathbf{v}} L^2_{\rm uv}$

$$\begin{array}{c} \\ 1 \\ u \\ u^2 \\ u^3 \\ u^4 \\ u^5 \\ u^6 \\ u^7 \\ u^8 \\ u^9 \\ u^{10} \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & v & v^2 & v^3 & v^4 \end{array} \\ \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -2 & 1 & \cdot \\ \cdot & \cdot & \cdot & -4 & \cdot \\ 1 & 1 & -9 & 3 & \cdot \\ 2 & -3 & 7 & \cdot & \cdot \\ 3 & \cdot & \cdot & \cdot & \cdot \\ 4 & \cdot & -1 & \cdot & \cdot \\ 3 & -1 & 1 & \cdot & \cdot \\ 2 & 3 & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \end{array}$$

Table 6: Examples of $\boxplus$ and $\boxtimes$ operations on a pair of bivariate polynomials, $L^1_{\rm uv}$ and $L^2_{\rm uv}$.

and

$$\boxed{L^3_{\rm uv}(t,v) = L^1_{\rm uv} \boxtimes_{\rm u} L^2_{\rm uv} \equiv \mathrm{Res}_{\rm u}\left(u^{D^1_{\rm u}} L^1_{\rm uv}(t/u,v), L^2_{\rm uv}(u,v)\right),} \tag{4.3}$$

where $D^1_{\rm u}$ is the degree of $L^1_{\rm uv}$ with respect to u. By Proposition 4.9, evaluating the $\boxplus_{\rm u}$ and $\boxtimes_{\rm u}$ operations via the resultant formulation involves computing the determinant of the $(D^1_{\rm u}+D^2_{\rm u})\times(D^1_{\rm u}+D^2_{\rm u})$ Sylvester matrix. When $L^1_{\rm uv} \neq L^2_{\rm uv}$, this results in a steep computational saving relative to the companion matrix based formulation in Table 5 which involves computing the determinant of a $(D^1_{\rm u}D^2_{\rm u})\times(D^1_{\rm u}D^2_{\rm u})$ matrix. Fast algorithms for computing the resultant exploit this and other properties of the Sylvester matrix formulation. In MAPLE , the computation $L^3_{\rm uv} = L^1_{\rm uv} \boxplus_{\rm u} L^2_{\rm uv}$ may be performed using the command:

```
Luv3 = subs(t=u,resultant(subs(u=t-u,Luv1),Luv2,u));
```

The computation $L^3_{\rm uv} = L^1_{\rm uv} \boxtimes_{\rm u} L^2_{\rm uv}$ can be performed via the sequence of commands:

```
Du1 = degree(Luv1,u);
Luv3 =  subs(t=u,resultant(simplify(u^Du1*subs(u=t/u,Luv1)),Luv2,u));
```

When $L^1_{\rm uv} = L^2_{\rm uv}$, however, the $\boxplus_{\rm u}$ and $\boxtimes_{\rm u}$ operations are best performed using the companion matrix formulation in Table 5. The software implementation of the operations in Table 5 in [22] uses the companion matrix formulation when $L^1_{\rm uv} = L^2_{\rm uv}$ and the resultant formulation otherwise.
Thus far we have established our ability to encode algebraic distribution as solutions of bivariate polynomial equations and to manipulate the solutions. This sets the stage for defining the class of "algebraic" random matrices next.

5. Class of algebraic random matrices

We are interested in identifying canonical random matrix operations for which the limiting eigenvalue distribution of the resulting matrix is an algebraic distribution. This is equivalent to identifying operations for which the transformations in the random matrices can be mapped into transformations of the bivariate polynomial that encodes the limiting eigenvalue distribution function. This motivates the construction of the class of "algebraic" random matrices which we shall define next.

The practical utility of this definition, which will become apparent in Section 6. and 10. can be succinctly summarized: if a random matrix is shown to be algebraic then its limiting eigenvalue density function can be computed using a simple root-finding algorithm. Furthermore, if the moments exist, they will satisfy a finite depth linear recursion (see Theorem 8.6) with polynomial coefficients so that we will often be able to enumerate them efficiently in closed form. Algebraicity of a random matrix thus acts as a certificate of the computability of its limiting eigenvalue density function and the associated moments. In this chapter our objective is to specify the class of algebraic random matrices by its generators.

5.1 Preliminaries

Let $\mathbf{A}_N$, for $N = 1, 2, \ldots$ be a sequence of $N \times N$ random matrices with real eigenvalues. Let $F^{\mathbf{A}_N}$ denote the e.d.f., as in (1.5). Suppose $F^{\mathbf{A}_N}(x)$ converges almost surely (or in probability), for every x, to $F^A(x)$ as $N \to \infty$, then we say that $\mathbf{A}_N \mapsto A$. We denote the associated (non-random) limiting probability density function by $f_A(x)$.

Notation 5.1 (Mode of convergence of the empirical distribution function)**.** *When necessary we highlight the mode of convergence of the underlying distribution function thus: if* $\mathbf{A}_N \overset{a.s.}{\longmapsto} A$ *then it is shorthand for the statement that the empirical distribution function of* $\mathbf{A}_N$ *converges almost surely to the distribution function* F^A*; likewise* $\mathbf{A}_N \overset{p}{\longmapsto} A$ *is shorthand for the statement that the empirical distribution function of* $\mathbf{A}_N$ *converges in probability to the distribution function* F^A*. When the distinction is not made then almost sure convergence is assumed.*

Remark 5.2. *The element A above is not to be interpreted as a matrix. There is no convergence in the sense of an $\infty \times \infty$ matrix. The notation $\mathbf{A}_N \overset{a.s}{\longmapsto} A$ is shorthand for describing the convergence of the associated distribution functions and not of the matrix itself. We think of A as being an (abstract) element of a probability space with distribution function F^A and associated density function f_A.*

Definition 5.3 (Atomic random matrix)**.** *If $f_A \in \mathcal{P}_{atom}$ then we say that $\mathbf{A}_N$ is an atomic random matrix. We represent this as $\mathbf{A}_N \mapsto A \in \mathcal{M}_{atom}$ where $\mathcal{M}_{atom}$ denotes the class of atomic random matrices.*

Definition 5.4 (Algebraic random matrix)**.** *If $f_A \in \mathcal{P}_{alg}$ then we say that $\mathbf{A}_N$ is an algebraically characterizable random matrix (often suppressing the word characterizable for brevity). We represent this as $\mathbf{A}_N \longmapsto A \in \mathcal{M}_{alg}$ where $\mathcal{M}_{alg}$ denotes the class of algebraic random matrices. Note that, by definition, $\mathcal{M}_{atom} \subset \mathcal{M}_{alg}$.*

5.2 Key idea used in proving algebraicity preserving nature of a random matrix transformation

The ability to describe the class of algebraic random matrices and the technique needed to compute the associated bivariate polynomial is at the crux our investigation. In the theorems that follow, we accomplish the former by cataloguing random matrix operations that preserve algebraicity of the limiting distribution.

Our proofs shall rely on exploiting the fact that some random matrix transformations, say $\mathbf{A}_N \longmapsto \mathbf{B}_N$, can be most naturally expressed as transformations of $L^A_{\mathrm{mz}} \longmapsto L^B_{\mathrm{mz}}$; others as $L^A_{\mathrm{rg}} \longmapsto L^B_{\mathrm{rg}}$ while some as $L^A_{\mathrm{sy}} \longmapsto L^B_{\mathrm{sy}}$. Hence, we manipulate the bivariate polynomials (using the transformations depicted in Figure 3) to the form needed to apply the appropriate operational law, which we derive as part of the proof, and then reverse the transformations to obtain the bivariate polynomial L^B_{mz}. Once we have derived the operational law for computing L^B_{mz} from L^A_{mz}, we have established the algebraicity of the limiting eigenvalue distribution of $\mathbf{B}_N$ and we are done. Readers interested in the operational law may skip directly to Section 6..

The following property of the convergence of distributions will be invaluable in the proofs that follow .

Proposition 5.5 (Continuous mapping theorem)**.** *Let $\mathbf{A}_N \longmapsto A$. Let f_A and $\mathcal{S}^\delta_A$ denote the corresponding limiting density function and the atomic component of the support, respectively. Consider the mapping $y = h(x)$ continuous everywhere on the real line except on the set of its discontinuities denoted by $\mathcal{D}_h$. If $D_h \cap \mathcal{S}^\delta_A = \emptyset$ then $\mathbf{B}_N = h(\mathbf{A}_N) \longmapsto B$. The associated non-random distribution function, F^B is given by $F^B(y) = F^A\left(h^{\langle -1\rangle}(y)\right)$. The associated probability density function is its distributional derivative.*

Proof. This is a restatement of continuous mapping theorem which follows from well-known facts about the convergence of distributions [7]. □

5.3 Deterministic operations

We first consider some simple deterministic transformations on an algebraic random matrix $\mathbf{A}_N$ that produce an algebraic random matrix $\mathbf{B}_N$.

Theorem 5.6. *Let $\mathbf{A}_N \mapsto A \in \mathcal{M}_{alg}$ and p, q, r, and s be real-valued scalars. Then,*

$$\mathbf{B}_N = (p\,\mathbf{A}_N + q\,\mathbf{I}_N)/(r\,\mathbf{A}_N + s\,\mathbf{I}_N) \mapsto B \in \mathcal{M}_{alg},$$

provided f_A does not contain an atom at $-s/r$ and r, s are not zero simultaneously.

Proof. Here we have $h(x) = (p\,x + r)/(q\,x + s)$ which is continuous everywhere except at $x = -s/r$ for s and r not simultaneously zero. From Proposition 5.5, unless $f_A(x)$ has an atomic component at $-s/r$,

$\mathbf{B}_N \mapsto B$. The Stieltjes transform of F^B can be expressed as

$$m_B(z) = E_y\left[\frac{1}{y-z}\right] = E_x\left[\frac{r\,x+s}{p\,x+q-z(r\,x+s)}\right]. \tag{5.1}$$

Equation (5.1) can be rewritten as

$$m_B(z) = \int \frac{rx+s}{(p-rz)x+(q-sz)} dF^A(x) = \frac{1}{p-rz}\int \frac{rx+s}{x+\frac{q-sz}{p-rz}} dF^A(x). \tag{5.2}$$

With some algebraic manipulations, we can rewrite (5.2) as

$$\begin{aligned} m_B(z) &= \beta_z \int \frac{rx+s}{x+\alpha_z} dF^A(x) = \beta_z\left(r\int \frac{x}{x+\alpha_z} dF^A(x) + s\int \frac{1}{x+\alpha_z} dF^A(x)\right) \\ &= \beta_z\left(r\int dF^A(x) - r\,\alpha_z \int \frac{1}{x+\alpha_z} dF^A(x) + s\int \frac{1}{x+\alpha_z} dF^A(x)\right). \end{aligned} \tag{5.3}$$

where $\beta_z = 1/(p-r\,z)$ and $\alpha_z = (q-s\,z)/(p-r\,z)$. Using the definition of the Stieltjes transform and the identity $\int dF^A(x) = 1$, we can express $m_B(z)$ in (5.3) in terms of $m_A(z)$ as

$$m_B(z) = \beta_z\, r + (\beta_z\, s - \beta\, r\,\alpha_z)\, m_A(-\alpha_z). \tag{5.4}$$

Equation (5.4) can, equivalently, be rewritten as

$$m_A(-\alpha_z) = \frac{m_B(z) - \beta_z\, r}{\beta_z\, s - \beta_z\, r\,\alpha_z}. \tag{5.5}$$

Equation (5.5) can be expressed as an operational law on $L^A_{\rm mz}$ as

$$\boxed{L^B_{\rm mz}(m,z) = L^A_{\rm mz}((m-\beta_z\, r)/(\beta_z\, s - \beta_z\, r\,\alpha_z), -\alpha_z).} \tag{5.6}$$

Since $L^A_{\rm mz}$ exists, we can obtain $L^B_{\rm mz}$ by applying the transformation in (5.6), and clearing the denominator to obtain the irreducible bivariate polynomial consistent with Remark 3.3. Since $L^B_{\rm mz}$ exists, this proves that $f_B \in \mathcal{P}_{\rm alg}$ and $\mathbf{B}_N \mapsto B \in \mathcal{M}_{\rm alg}$. □

Appropriate substitutions for the scalars p, q, r and s in Theorem 5.6 leads to the following Corollary.

Corollary 5.7. *Let $\mathbf{A}_N \mapsto A \in \mathcal{M}_{alg}$ and let α be a real-valued scalar. Then,*

1. $\mathbf{B}_N = \mathbf{A}_N^{-1} \mapsto B \in \mathcal{M}_{alg}$, *provided f_A does not contain at atom at 0,*
2. $\mathbf{B}_N = \alpha\,\mathbf{A}_N \mapsto B \in \mathcal{M}_{alg}$,
3. $\mathbf{B}_N = \mathbf{A}_N + \alpha\,\mathbf{I}_N \mapsto B \in \mathcal{M}_{alg}$.

Theorem 5.8. *Let $\mathbf{X}_{n,N}$ be an $n \times N$ matrix. If $\mathbf{A}_N = \mathbf{X}_{n,N}\mathbf{X}_{n,N}^{'} \mapsto A \in \mathcal{M}_{alg}$ then*

$$\mathbf{B}_N = \mathbf{X}_{n,N}^{'}\mathbf{X}_{n,N} \mapsto B \in \mathcal{M}_{alg}\,.$$

PROOF. Here $\mathbf{X}_{n,N}$ is an $n \times N$ matrix, so that $\mathbf{A}_n$ and $\mathbf{B}_N$ are $n \times n$ and $N \times N$ sized matrices respectively. Let $c_N = n/N$. When $c_N < 1$, $\mathbf{B}_N$ will have $N-n$ eigenvalues of magnitude zero while the remaining n

eigenvalues will be identically equal to the eigenvalues of $\mathbf{A}_n$. Thus, the e.d.f. of $\mathbf{B}_N$ is related to the e.d.f. of $\mathbf{A}_n$ as

$$\begin{aligned} F^{\mathbf{B}_N}(x) &= \frac{N-n}{N}\,\mathbb{I}_{[0,\infty)} + \frac{n}{N}F^{\mathbf{A}_n}(x) \\ &= (1-c_N)\,\mathbb{I}_{[0,\infty)} + c_N\,F^{\mathbf{A}_n}(x). \end{aligned} \tag{5.7}$$

where $\mathbb{I}_{[0,\infty)}$ is the indicator function that is equal to 1 when $x \geq 0$ and is equal to zero otherwise.

Similarly, when $c_N > 1$, $\mathbf{A}_n$ will have $n - N$ eigenvalues of magnitude zero while the remaining N eigenvalues will be identically equal to the eigenvalues of $\mathbf{B}_N$. Thus the e.d.f. of $\mathbf{A}_n$ is related to the e.d.f. of $\mathbf{B}_N$ as

$$\begin{aligned} F^{\mathbf{A}_n}(x) &= \frac{n-N}{n}\,\mathbb{I}_{[0,\infty)} + \frac{N}{n}F^{\mathbf{B}_N}(x) \\ &= \left(1-\frac{1}{c_N}\right)\mathbb{I}_{[0,\infty)} + \frac{1}{c_N}\,F^{\mathbf{B}_N}(x). \end{aligned} \tag{5.8}$$

Equation (5.8) is (5.7) rearranged; so we do not need to differentiate between the case when $c_N < 1$ and $c_N > 1$.

Thus, as $n, N \to \infty$ with $c_N = n/N \to c$, if $F^{\mathbf{A}_n}$ converges to a non-random d.f. F^A, then $F^{\mathbf{B}_N}$ will also converge to a non-random d.f. F^B related to F^A by

$$F^B(x) = (1-c)\mathbb{I}_{[0,\infty)} + c\,F^A(x). \tag{5.9}$$

From (5.9), it is evident that the Stieltjes transform of the limiting distribution functions F^A and F^B are related as

$$m_A(z) = -\left(1-\frac{1}{c}\right)\frac{1}{z} + \frac{1}{c}m_B(z). \tag{5.10}$$

Rearranging the terms on either side of (5.10) allows us to express $m_B(z)$ in terms of $m_A(z)$ as

$$m_B(z) = -\frac{1-c}{z} + c\,m_A(z). \tag{5.11}$$

Equation (5.11) can be expressed as an operational law on L^A_{mz} as

$$\boxed{L^B_{\mathrm{mz}}(m,z) = L^A_{\mathrm{mz}}\left(-\left(1-\frac{1}{c}\right)\frac{1}{z} + \frac{1}{c}m, z\right).} \tag{5.12}$$

Given L^A_{mz}, we can obtain L^B_{mz} by using (5.12). Hence $\mathbf{B}_N \mapsto B \in \mathcal{M}_{\mathrm{alg}}$. □

Theorem 5.9. *Let* $\mathbf{A}_N \mapsto A \in \mathcal{M}_{alg}$. *Then*

$$\mathbf{B}_N = (\mathbf{A}_N)^2 \mapsto B \in \mathcal{M}_{alg}\,.$$

Proof. Here we have $h(x) = x^2$ which is continuous everywhere. From Proposition 5.5, $\mathbf{B}_N \mapsto B$. The Stieltjes transform of F^B can be expressed as

$$m_B(z) = E_Y\left[\frac{1}{y-z}\right] = E_X\left[\frac{1}{x^2-z}\right]. \tag{5.13}$$

Equation (5.13) can be rewritten as

$$m_B(z) = \frac{1}{2\sqrt{z}} \int \frac{1}{x - \sqrt{z}} dF^A(x) - \frac{1}{2\sqrt{z}} \int \frac{1}{x + \sqrt{z}} dF^A(x) \tag{5.14}$$

$$= \frac{1}{2\sqrt{z}} m_A(\sqrt{z}) - \frac{1}{2\sqrt{z}} m_A(-\sqrt{z}). \tag{5.15}$$

Equation (5.14) leads to the operational law

$$\boxed{L^B_{\rm mz}(m, z) = L^A_{\rm mz}(2m\sqrt{z}, \sqrt{z}) \boxplus_{\rm m} L^A_{\rm mz}(-2m\sqrt{z}, \sqrt{z}).} \tag{5.16}$$

Given $L^A_{\rm mz}$, we can obtain $L^B_{\rm mz}$ by using (5.16). This proves that $\mathbf{B}_N \mapsto B \in \mathcal{M}_{\rm alg}$. □

Theorem 5.10. *Let $\mathbf{A}_n \mapsto A \in \mathcal{M}_{alg}$ and $\mathbf{B}_N \mapsto B \in \mathcal{M}_{alg}$. Then,*

$$\mathbf{C}_M = \mathrm{diag}(\mathbf{A}_n, \mathbf{B}_N) \mapsto C \in \mathcal{M}_{alg},$$

where $M = n + N$ and $n/N \to c > 0$ as $n, N \to \infty$.

Proof. Let $\mathbf{C}_N$ be an $N \times N$ block diagonal matrix formed from the $n \times n$ matrix $\mathbf{A}_n$ and the $M \times M$ matrix $\mathbf{B}_M$. Let $c_N = n/N$. The e.d.f. of $\mathbf{C}_N$ is given by

$$F^{\mathbf{C}_N} = c_N F^{\mathbf{A}_n} + (1 - c_N) F^{\mathbf{B}_M}.$$

Let $n, N \to \infty$ and $c_N = n/N \to c$. If $F^{\mathbf{A}_n}$ and $F^{\mathbf{B}_M}$ converge in distribution almost surely (or in probability) to non-random d.f.'s F^A and F^B respectively, then $F^{\mathbf{C}_N}$ will also converge in distribution almost surely (or in probability) to a non-random distribution function F^C given by

$$F^C(x) = c F^A(x) + (1 - c) F^B(x). \tag{5.17}$$

The Stieltjes transform of the distribution function F^C can hence be written in terms of the Stieltjes transforms of the distribution functions F^A and F^B as

$$m_C(z) = c m_A(z) + (1 - c) m_B(z) \tag{5.18}$$

Equation (5.18) can be expressed as an operational law on the bivariate polynomial $L^A_{\rm mz}(m, z)$ as

$$\boxed{L^C_{\rm mz} = L^A_{\rm mz}\left(\frac{m}{c}, z\right) \boxplus_{\rm m} L^B_{\rm mz}\left(\frac{m}{1 - c}, z\right).} \tag{5.19}$$

Given $L^A_{\rm mz}$ and $L^B_{\rm mz}$, and the definition of the $\boxplus_{\rm m}$ operator in Section 4., $L^C_{\rm mz}$ is a polynomial which can be constructed explicitly. This proves that $\mathbf{C}_N \mapsto C \in \mathcal{M}_{\rm alg}$. □

Theorem 5.11. *If $\mathbf{A}_n = \mathrm{diag}(\mathbf{B}_N, \alpha \mathbf{I}_{n-N})$ and α is a real valued scalar. Then,*

$$\mathbf{B}_N \mapsto B \in \mathcal{M}_{alg},$$

as $n, N \to \infty$ with $c_N = n/N \to c$,

Proof. Assume that as $n, N \to \infty$, $c_N = n/N \to c$. As we did in the proof of Theorem 5.10, we can show that the Stieltjes transform $m_A(z)$ can be expressed in terms of $m_B(z)$ as

$$m_A(z) = \left(\frac{1}{c} - 1\right) \frac{1}{\alpha - z} + \frac{1}{c} m_B(z). \tag{5.20}$$

This allows us to express $L^B_{\rm mz}(m,z)$ in terms of $L^A_{\rm mz}(m,z)$ using the relationship in (5.20) as

$$\boxed{L^B_{\rm mz}(m,z) = L^A_{\rm mz}\left(-\left(\frac{1}{c}-1\right)\frac{1}{\alpha-z}+\frac{1}{c}\,m, z\right).} \tag{5.21}$$

We can hence obtain $L^B_{\rm mz}$ from $L^A_{\rm mz}$ using (5.21). This proves that $\mathbf{B}_N \mapsto B \in \mathcal{M}_{\rm alg}$. □

Corollary 5.12. *Let $\mathbf{A}_N \mapsto A \in \mathcal{M}_{alg}$. Then*

$$\mathbf{B}_N = \mathrm{diag}(\mathbf{A}_n, \alpha\,\mathbf{I}_{N-n}) \mapsto B \in \mathcal{M}_{alg},$$

for $n/N \to c > 0$ as $n, N \to \infty$.

Proof. This follows directly from Theorem 5.10. □

5.4 Gaussian-like operations

We now consider some simple stochastic transformations that "blur" the eigenvalues of $\mathbf{A}_N$ by injecting additional randomness. We show that canonical operations involving an algebraic random matrix $\mathbf{A}_N$ and Gaussian-like and Wishart-like random matrices (defined next) produce an algebraic random matrix $\mathbf{B}_N$.

Definition 5.13 (Gaussian-like random matrix)**.** *Let $\mathbf{Y}_{N,L}$ be an $N \times L$ matrix with independent, identically distributed (i.i.d.) elements having zero mean, unit variance and bounded higher order moments. We label the matrix $\mathbf{G}_{N,L} = \frac{1}{\sqrt{L}}\mathbf{Y}_{N,L}$ as a Gaussian-like random matrix.*

We can sample a Gaussian-like random matrix in Matlab as

```
G = sign(randn(N,L))/sqrt(L);
```

Gaussian-like matrices are labelled thus because they exhibit the same limiting behavior in the $N \to \infty$ limit as "pure" Gaussian matrices which may be sampled in Matlab as

```
G = randn(N,L)/sqrt(L);
```

Definition 5.14 (Wishart-like random matrix)**.** *Let $\mathbf{G}_{N,L}$ be a Gaussian-like random matrix. We label the matrix $W_N = \mathbf{G}_{N,L} \times \mathbf{G}'_{N,L}$ as a Wishart-like random matrix. Let $c_N = N/L$. We denote a Wishart-like random matrix thus formed by $\mathbf{W}_N(c_N)$.*

Remark 5.15 (Algebraicity of Wishart-like random matrices)**.** *The limiting eigenvalue distribution of the Wishart-like random matrix has the Marčenko-Pastur density which is an algebraic density since L^W_{mz} exists (see Table 1(b)).*

Proposition 5.16. *Assume that $\mathbf{G}_{N,L}$ is an $N \times L$ Gaussian-like random matrix. Let $\mathbf{A}_N \overset{\text{a.s.}}{\longmapsto} A$ be an $N \times N$ symmetric/Hermitian random matrix and $\mathbf{T}_L \overset{\text{a.s.}}{\longmapsto} T$ be an $L \times L$ diagonal atomic random matrix respectively. If $\mathbf{G}_{N,L}$, $\mathbf{A}_N$ and $\mathbf{T}_L$ are independent then $\mathbf{B}_N = \mathbf{A}_N + \mathbf{G}'_{N,L}\mathbf{T}_L\mathbf{G}_{N,L} \overset{\text{a.s.}}{\longmapsto} B$, as $c_L = N/L \to c$ for $N, L \to \infty$,. The Stieltjes transform $m_B(z)$ of the unique distribution function F^B is satisfies the equation*

$$\boxed{m_B(z) = m_A\left(z - c\int \frac{x\,dF^T(x)}{1 + x\,m_B(z)}\right).} \tag{5.22}$$

PROOF. This result may be found in Marčenko-Pastur [18] and Silverstein [26]. □

We can reformulate Proposition 5.16 to obtain the following result on algebraic random matrices.

Theorem 5.17. *Let* $\mathbf{A}_N$, $\mathbf{G}_{N,L}$ *and* $\mathbf{T}_L$ *be defined as in Proposition 5.16. Then*

$$\mathbf{B}_N = \mathbf{A}_N + \mathbf{G}_{L,N}^{'} \mathbf{T}_L \mathbf{G}_{L,N} \xrightarrow{\text{a.s.}} B \in \mathcal{M}_{alg},$$

as $c_L = N/L \to c$ *for* $N, L \to \infty$.

PROOF. Let $\mathbf{T}_L$ be an atomic matrix with d atomic masses of weight p_i and magnitude λ_i for $i = 1, 2, \ldots, d$. From Proposition 5.16, $m_B(z)$ can be written in terms of $m_A(z)$ as

$$m_B(z) = m_A\left(z - c\sum_{i=1}^{d} \frac{p_i\,\lambda_i}{1 + \lambda_i\, m_B(z)}\right). \tag{5.23}$$

where we have substituted $F^T(x) = \sum_{i=1}^{d} p_i\, \mathbb{I}_{[\lambda_i, \infty)}$ into (5.22) with $\sum_i p_i = 1$.

Equation (5.23) can be expressed as an operational law on the bivariate polynomial L_{mz}^A as

$$\boxed{L_{\text{mz}}^B(m, z) = L_{\text{mz}}^A(m, z - \alpha_m).} \tag{5.24}$$

where $\alpha_m = c\sum_{i=1}^{d} p_i\,\lambda_i/(1 + \lambda_i\, m)$. This proves that $\mathbf{B}_N \xrightarrow{\text{a.s.}} B \in \mathcal{M}_{\text{alg}}$. □

Proposition 5.18. *Assume that* $\mathbf{W}_N(c_N)$ *is an* $N \times N$ *Wishart-like random matrix. Let* $\mathbf{A}_N \xrightarrow{\text{a.s.}} A$ *be an* $N \times N$ *random Hermitian non-negative definite matrix. If* $\mathbf{W}_N(c_N)$ *and* $\mathbf{A}_N$ *are independent, then* $\mathbf{B}_N = \mathbf{A}_N \times \mathbf{W}_N(c_N) \xrightarrow{\text{a.s.}} B$ *as* $c_N \to c$. *The Stieltjes transform* $m_B(z)$ *of the unique distribution function* F^B *satisfies*

$$m_B(z) = \int \frac{dF^A(x)}{\{1 - c - c\,z\,m_B(z)\}x - z}. \tag{5.25}$$

PROOF. This result may be found in Bai and Silverstein [4, 26]. □

We can reformulate Proposition 5.18 to obtain the following result on algebraic random matrices.

Theorem 5.19. *Let* $\mathbf{A}_N$ *and* $\mathbf{W}_N(c_N)$ *satisfy the hypothesis of Proposition 5.18. Then,*

$$\mathbf{B}_N = \mathbf{A}_N \times \mathbf{W}_N(c_N) \xrightarrow{\text{a.s.}} B \in \mathcal{M}_{alg},$$

as $c_N \to c$.

PROOF. By rearranging the terms in the numerator and denominator, (5.25) can be rewritten as

$$m_B(z) = \frac{1}{1 - c - c\,z\,m_B(z)} \int \frac{dF^A(x)}{x - \frac{z}{1 - c - c\,z\,m_B(z)}}. \tag{5.26}$$

Let $\alpha_{m,z} = 1 - c - c\,z\,m_B(z)$ so that (5.26) can be rewritten as

$$m_B(z) = \frac{1}{\alpha_{m,z}} \int \frac{dF^A(x)}{x - (z/\alpha_{m,z})}. \tag{5.27}$$

We can express $m_B(z)$ in (5.27) in terms of $m_A(z)$ as

$$m_B(z) = \frac{1}{\alpha_{m,z}}\, m_A(z/\alpha_{m,z}). \tag{5.28}$$

Equation (5.28) can be rewritten as

$$m_A(z/\alpha_{m,z}) = \alpha_{m,z}\, m_B(z). \tag{5.29}$$

Equation (5.29) can be expressed as an operational law on the bivariate polynomial L^A_{mz} as

$$\boxed{L^B_{\text{mz}}(m,z) = L^A_{\text{mz}}(\alpha_{m,z}\, m, z/\alpha_{m,z}).} \tag{5.30}$$

This proves that $\mathbf{B}_N \xmapsto{\text{a.s.}} B \in \mathcal{M}_{\text{alg}}$. □

Proposition 5.20. *Assume that $\mathbf{G}_{N,L}$ is an $N \times L$ Gaussian-like random matrix. Let $\mathbf{A}_N \xmapsto{\text{a.s.}} A$ be an $N \times N$ symmetric/Hermitian random matrix independent of $\mathbf{G}_{N,L}$, $\mathbf{A}_N$. Let $\mathbf{A}_N^{1/2}$ denote an $N \times L$ matrix. If s is a positive real-valued scalar then $\mathbf{B}_N = (\mathbf{A}_N^{1/2} + \sqrt{s}\,\mathbf{G}_{N,L})(\mathbf{A}_N^{1/2} + \sqrt{s}\,\mathbf{G}_{N,L})' \xmapsto{\text{a.s.}} B$, as $c_L = N/L \to c$ for $N, L \to \infty$. The Stieltjes transform, $m_B(z)$ of the unique distribution function F^B satisfies the equation*

$$m_B(z) = -\int \frac{dF^A(x)}{z\,\{1 + s\,c\,m_B(z)\} - \frac{x}{1+s\,c\,m_B(z)} + s\,(c-1)}. \tag{5.31}$$

Proof. This result is found in Dozier and Silverstein [12]. □
We can reformulate Proposition 5.20 to obtain the following result on algebraic random matrices.

Theorem 5.21. *Assume $\mathbf{A}_N$, $\mathbf{G}_{N,L}$ and s satisfy the hypothesis of Proposition 5.20. Then*

$$\mathbf{B}_N = (\mathbf{A}_N^{1/2} + \sqrt{s}\,\mathbf{G}_{N,L})(\mathbf{A}_N^{1/2} + \sqrt{s}\,\mathbf{G}_{N,L})' \xmapsto{\text{a.s.}} B \in \mathcal{M}_{alg},$$

as $c_L = N/L \to c$ for $N, L \to \infty$.

Proof. By rearranging the terms in the numerator and denominator, (5.31) can be rewritten as

$$m_B(z) = \int \frac{\{1 + s\,c\,m_B(z)\}\, dF^A(x)}{x - \{1 + s\,c\,m_B(z)\}(z\,\{1 + s\,c\,m_B(z)\} + (c-1)\,s)}. \tag{5.32}$$

Let $\alpha_m = 1 + s\,c\,m_B(z)$ and $\beta_m = \{1 + s\,c\,m_B(z)\}(z\,\{1 + s\,c\,m_B(z)\} + (c-1)\,s)$, so that $\beta = \alpha_m^2\, z + \alpha_m\, s(c-1)$. Equation (5.32) can hence be rewritten as

$$m_B(z) = \alpha_m \int \frac{dF^A(x)}{x - \beta_m}. \tag{5.33}$$

Using the definition of the Stieltjes transform in (2.1), we can express $m_B(z)$ in (5.33) in terms of $m_A(z)$ as

$$\begin{aligned} m_B(z) &= \alpha_m\, m_A(\beta_m) \\ &= \alpha_m\, m_A(\alpha_m^2\, z + \alpha_m(c-1)s). \end{aligned} \tag{5.34}$$

Equation (5.34) can, equivalently, be rewritten as

$$m_A(\alpha_m^2\, z + \alpha_m(c-1)s) = \frac{1}{\alpha_m} m_B(z). \tag{5.35}$$

Equation (5.35) can be expressed as an operational law on the bivariate polynomial L_{mz} as

$$\boxed{L^B_{\text{mz}}(m,z) = L^A_{\text{mz}}(m/\alpha_m, \alpha^2\, z + \alpha_m\, s(c-1)).} \tag{5.36}$$

This proves that $\mathbf{B}_N \xmapsto{\text{a.s.}} B \in \mathcal{M}_{\text{alg}}$. □

5.5 Sums and products

Proposition 5.22. *Let* $\mathbf{A}_N \overset{\text{p}}{\longmapsto} A$ *and* $\mathbf{B}_N \overset{\text{p}}{\longmapsto} B$ *be* $N \times N$ *symmetric/Hermitian random matrices. Let* $\mathbf{Q}_N$ *be a Haar distributed unitary/orthogonal matrix independent of* $\mathbf{A}_N$ *and* $\mathbf{B}_N$*. Then* $\mathbf{C}_N = \mathbf{A}_N + \mathbf{Q}_N\mathbf{B}_N\mathbf{Q}_N' \overset{\text{p}}{\longmapsto} C$*. The associated distribution function* F^C *is the unique distribution function whose* R *transform satisfies*

$$r_C(g) = r_A(g) + r_B(g). \tag{5.37}$$

Proof. This result was obtained by Voiculescu in [34]. □

We can reformulate Proposition 5.22 to obtain the following result on algebraic random matrices.

Theorem 5.23. *Assume that* $\mathbf{A}_N$, $\mathbf{B}_N$ *and* $\mathbf{Q}_N$ *satisfy the hypothesis of Proposition 5.22. Then,*

$$\mathbf{C}_N = \mathbf{A}_N + \mathbf{Q}_N\mathbf{B}_N\mathbf{Q}_N^{'} \overset{\text{p}}{\longmapsto} C \in \mathcal{M}_{alg}$$

Proof. Equation (5.37) can be expressed as an operational law on the bivariate polynomials L^A_{rg} and L^B_{rg} as

$$\boxed{L^C_{\text{rg}} = L^A_{\text{rg}} \boxplus_{\text{r}} L^B_{\text{rg}}} \tag{5.38}$$

If L_{mz} exists then so does L_{rg} and vice-versa. This proves that $\mathbf{C}_N \overset{\text{p}}{\longmapsto} C \in \mathcal{M}_{\text{alg}}$. □

Proposition 5.24. *Let* $\mathbf{A}_N \overset{\text{p}}{\longmapsto} A$ *and* $\mathbf{B}_N \overset{\text{p}}{\longmapsto} B$ *be* $N \times N$ *symmetric/Hermitian random matrices. Let* $\mathbf{Q}_N$ *be a Haar distributed unitary/orthogonal matrix independent of* $\mathbf{A}_N$ *and* $\mathbf{B}_N$*. Then* $\mathbf{C}_N = \mathbf{A}_N \times \mathbf{Q}_N\mathbf{B}_N\mathbf{Q}_N^{'} \overset{\text{p}}{\longmapsto} C$ *where* $\mathbf{C}_N$ *is defined only if* $\mathbf{C}_N$ *has real eigenvalues for every sequence* $\mathbf{A}_N$ *and* $\mathbf{B}_N$*. The associated distribution function* F^C *is the unique distribution function whose* S *transform satisfies*

$$s_C(y) = s_A(y)s_B(y). \tag{5.39}$$

Proof. This result was obtained by Voiculescu in [35, 36]. □

We can reformulate Proposition 5.24 to obtain the following result on algebraic random matrices.

Theorem 5.25. *Assume that* $\mathbf{A}_N$, *and* $\mathbf{B}_N$ *satisfy the hypothesis of Proposition 5.24. Then*

$$\mathbf{C}_N = \mathbf{A}_N \times \mathbf{Q}_N\mathbf{B}_N\mathbf{Q}_N^{'} \overset{\text{p}}{\longmapsto} C \in \mathcal{M}_{alg}\,.$$

Proof. Equation (5.39) can be expressed as an operational law on the bivariate polynomials L^A_{sy} and L^B_{sy} as

$$\boxed{L^C_{\text{sy}} = L^A_{\text{sy}} \boxtimes_{\text{s}} L^B_{\text{sy}}} \tag{5.40}$$

If L_{mz} exists then so does L_{sy} and vice versa. This proves that $\mathbf{B}_N \overset{\text{p}}{\longmapsto} B \in \mathcal{M}_{\text{alg}}$. □

Definition 5.26 (Orthogonally/Unitarily invariant random matrix)**.** *If the joint distribution of the elements of a random matrix* $\mathbf{A}_N$ *is invariant under orthogonal/unitary transformations, it is referred to as an orthogonally/unitarily invariant random matrix.*

If $\mathbf{A}_N$ (or $\mathbf{B}_N$) or both are an orthogonally/unitarily invariant sequences of random matrices then Theorems 5.23 and 5.25 can be stated more simply.

Corollary 5.27. *Let $\mathbf{A}_N \xrightarrow{\text{p}} A \in \mathcal{M}_{alg}$ and $\mathbf{B}_N \to B \xrightarrow{\text{p}} \mathcal{M}_{alg}$ be a orthogonally/unitarily invariant random matrix independent of $\mathbf{A}_N$. Then,*

1. $\mathbf{C}_N = \mathbf{A}_N + \mathbf{B}_N \xrightarrow{\text{p}} C \in \mathcal{M}_{alg}$
2. $\mathbf{C}_N = \mathbf{A}_N \times \mathbf{B}_N \xrightarrow{\text{p}} C \in \mathcal{M}_{alg}$

Here multiplication is defined only if $\mathbf{C}_N$ has real eigenvalues for every sequence $\mathbf{A}_N$ and $\mathbf{B}_N$.

When both the limiting eigenvalue distributions of $\mathbf{A}_N$ and $\mathbf{B}_N$ have compact support, it is possible to strengthen the mode of convergence in Theorems 5.23 and 5.25 to almost surely [15]. We suspect that almost sure convergence must hold when the distributions are not compactly supported; this remains an open problem.

6. Operational laws on bivariate polynomials

The key idea behind the definition of algebraic random matrices in Section 5. was that when the limiting eigenvalue distribution of a random matrix can be encoded by a bivariate polynomial, then for the broad class of random matrix operations identified in Section 5., algebraicity of the eigenvalue distribution is preserved under the transformation.

These operational laws, the associated random matrix transformation and the symbolic MATLAB code for the operational law are summarized in Tables 7-9. The remainder of this chapter discusses techniques for extracting the density function from the polynomial and the special structure in the moments that allows them to be efficiently enumerated using symbolic methods.

$\mathbf{B}$	Operation	$L^B_{\mathbf{mz}}(m,z)$	MATLAB code
Deterministic Transformations			

$\frac{p\,\mathbf{A}+q\,\mathbf{I}}{r\,\mathbf{A}+s\,\mathbf{I}}$	''Mobius''	$L^A_{\mathbf{mz}}\left(\frac{m-\beta_z\,r}{\beta_z\,s-\beta_z\,r\,\alpha_z},-\alpha_z\right)$, where $\alpha_z=(q-s\,z)/(p-r\,z)$, and $\beta_z=1/(p-r\,z)$.	function LmzB = mobiusA(LmzA,p,q,r,s)
			syms m z
			alpha = ((q-s*z)/(p-r*z);beta=1/(p-r*z);
			temp_pol = subs(LmzA,z,-alpha);
			temp_pol = subs(temp_pol,m,((m/beta)-r)/(s-r*alpha));
			LmzB = irreducLuv(temp_pol,m,z);

$\mathbf{A}^{-1}$	''Invert''	$L^A_{\mathbf{mz}}\left(-z-z^2\,m,\frac{1}{z}\right)$	function LmzB = invA(LmzA) LmzB = mobiusA(LmzA,0,1,1,0);
$\mathbf{A}+\alpha\,\mathbf{I}$	''Translate''	$L^A_{\mathbf{mz}}(m,z-\alpha)$	function LmzB = shiftA(LmzA,alpha) LmzB = mobiusA(LmzA,1,alpha,0,1);
$\alpha\,\mathbf{A}$	''Scale''	$L^A_{\mathbf{mz}}\left(\alpha\,m,\frac{z}{\alpha}\right)$	function LmzB = scaleA(LmzA) LmzB = mobiusA(LmzA,alpha,0,0,1);

$\begin{bmatrix}\mathbf{A} & 0\\ 0 & \alpha\mathbf{I}\end{bmatrix}$	''Projection/ Transpose''	$\frac{\text{Size of }\mathbf{A}}{\text{Size of }\mathbf{B}}\to c>1$
		$L^A_{\mathbf{mz}}\left((1-\frac{1}{c})\frac{1}{\alpha-z}+\frac{m}{c},z\right)$
		$\frac{\text{Size of }\mathbf{A}}{\text{Size of }\mathbf{B}}\to c<1$
			function LmzB = projectA(LmzA,c,alpha)
			syms m z
			mb = (1-(1/c))*(1/(alpha-z))+m/c;
			temp_pol = subs(LmzA,m,mb);
			LmzB = irreducLuv(temp_pol,m,z);

$\mathbf{A}=\begin{bmatrix}\mathbf{B} & 0\\ 0 & \alpha\mathbf{I}\end{bmatrix}$	''Augmentation ''		function LmzB = augmentA(LmzA,c,alpha)
			syms m z
			mb = (1-(1/c))*(1/(alpha-z))+m/c;
			temp_pol = subs(LmzA,m,mb);
			LmzB = irreducLuv(temp_pol,m,z);

		Stochastic Transformations	

$\mathbf{A}+\mathbf{G}^{'}\,\mathbf{T}\,\mathbf{G}$	''Add Atomic Wishart ''	$L^A_{\mathbf{mz}}(m,z-\alpha_m)$, where $\alpha_m=c\sum_{i=1}^{d}\frac{p_i\,\lambda_i}{1+\lambda_i\,m}$, with $\sum_i p_i=1$.	function LmzB = AplusWish(LmzA,c,p,lambda)
			syms m z
			alpha = z-c*sum(p.*(lambda./(1+lambda*m)));
			temp_pol = subs(LmzA,z,z-alpha);
			LmzB = irreducLuv(temp_pol,m,z);

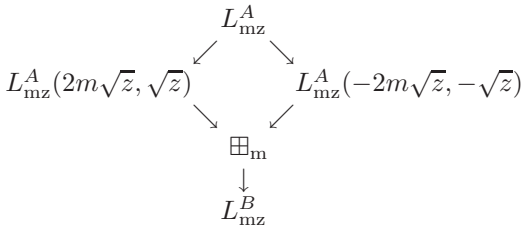
$\mathbf{A}\times\mathbf{W}(c)$	''Multiply Wishart''	$L^A_{\mathbf{mz}}\left(\alpha_{m,z}\,m,\frac{z}{\alpha_{m,z}}\right)$, where $\alpha_{m,z}=(1-c-c\,z\,m)$.	function LmzB = AtimesWish(LmzA,c)
			syms m z z1
			alpha = (1-c-c*z1*m); temp_pol = subs(LmzA,m,m*alpha);
			temp_pol = subs(temp_pol,z,z1/alpha);
			temp_pol = subs(temp_pol,z1,z); % Replace dummy variable
			LmzB = irreducLuv(temp_pol,m,z);

$(\mathbf{A}^{1/2}+\sqrt{s}\,\mathbf{G})$ $\times$ $(\mathbf{A}^{1/2}+\sqrt{s}\,\mathbf{G})'$	''Grammian''	$L^A_{\mathbf{mz}}\left(\frac{m}{\alpha_m},\alpha_m^2\,z+\alpha_m\,s(c-1)\right)$, where $\alpha_m=1+s\,c\,m$.	function LmzB = AgramG(LmzA,c,s)
			syms m z
			alpha = (1+s*c*m); beta = alpha*(z*alpha+s*(c-1));
			temp_pol = subs(subs(LmzA,m,m/alpha),z,beta);
			LmzB = irreducLuv(temp_pol,m,z);

Table 7: Operational laws on the bivariate polynomial encodings (and their computational realization in MATLAB) corresponding to a class of deterministic and stochastic transformations. The Gaussian-like random matrix $\mathbf{G}$ is an $N\times L$, the Wishart-like matrix $\mathbf{W}(c)=\mathbf{G}\,\mathbf{G}^{'}$ where $N/L\to c>0$ as $N,L\to\infty$, and the matrix $\mathbf{T}$ is a diagonal atomic random matrix.

(a) $L^A_{\mathrm{mz}} \longmapsto L^B_{\mathrm{mz}}$ for $\mathbf{A} \longmapsto \mathbf{B} = \mathbf{A}^2$.

Operational Law	MATLAB Code

	`function LmzB = squareA(LmzA)`
	`syms m z`
	
	`Lmz1 = subs(LmzA,z,sqrt(z));`
	`Lmz1 = subs(Lmz1,m,2*m*sqrt(z));`
	`Lmz2 = subs(LmzA,z,-sqrt(z));`
	`Lmz2 = subs(Lmz2,m,-2*m*sqrt(z));`
	
	`LmzB = L1plusL2(Lmz1,Lmz2,m);`
	`LmzB = irreducLuv(LmzB,m,z);`

(b) $L^A_{\mathrm{mz}}, L^A_{\mathrm{mz}} \longmapsto L^C_{\mathrm{mz}}$ for $\mathbf{A}, \mathbf{B} \longmapsto \mathbf{C} = \mathrm{diag}(\mathbf{A}, \mathbf{B})$ where Size of $\mathbf{A}$/ Size of $\mathbf{C} \to c$.

Operational Law	MATLAB Code

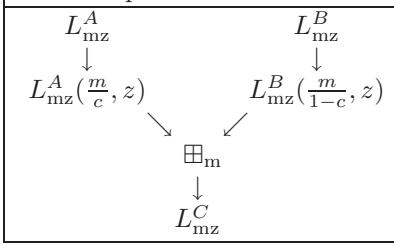
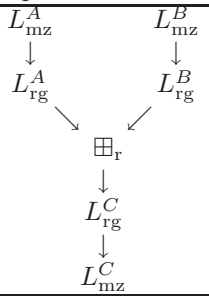
	`function LmzC = AblockB(LmzA,LmzB,c)`
	`syms m z mu`
	
	`LmzA1 = subs(LmzA,m,m/c);`
	`LmzB1 = subs(LmzB,m,m/(1-c));`
	
	`LmzC = L1plusL2(LmzA1,LmzB1,m);`
	`LmzC = irreducLuv(LmzC,m,z);`

Table 8: Operational laws on the bivariate polynomial encodings for some deterministic random matrix transformations. The operations $\boxplus_{\mathrm{u}}$ and $\boxtimes_{\mathrm{u}}$ are defined in Table 5.

(a) $L^A_{\rm mz}, L^B_{\rm mz} \longmapsto L^C_{\rm mz}$ for $\mathbf{A}, \mathbf{B} \longmapsto \mathbf{C} = \mathbf{A} + \mathbf{QBQ}'$.

Operational Law	MATLAB Code

	`function LmzC = AplusB(LmzA,LmzB)`
	`syms m z r g`
	
	`LrgA = Lmz2Lrg(LmzA);`
	`LrgB = Lmz2Lrg(LmzB);`
	
	`LrgC = L1plusL2(LrgA,LrgB,r);`
	
	`LmzC = Lrg2Lmz(LrgC);`

(b) $L^A_{\rm mz}, L^B_{\rm mz} \longmapsto L^C_{\rm mz}$ for $\mathbf{A}, \mathbf{B} \longmapsto \mathbf{C} = \mathbf{A} \times \mathbf{QBQ}'$.

Operational Law	MATLAB Code

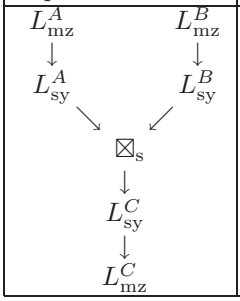
	`function LmzC = AtimesB(LmzA,LmzB)`
	`syms m z s y`
	
	`LsyA = Lmz2Lsy(LmzA);`
	`LsyB = Lmz2Lsy(LmzB);`
	
	`LsyC = L1timesL2(LsyA,LsyB,s);`
	
	`LmzC = Lsy2Lmz(LsyC);`

Table 9: Operational laws on the bivariate polynomial encodings for some canonical random matrix transformations. The operations $\boxplus_{\rm u}$ and $\boxtimes_{\rm u}$ are defined in Table 5.

7. Interpreting the solution curves of polynomial equations

Consider a bivariate polynomial $L_{\rm mz}$. Let $\mathrm{D_m}$ be the degree of $L_{\rm mz}(m,z)$ with respect to m and $l_k(z)$, for $k=0,\ldots,\mathrm{D_m}$, be polynomials in z that are the coefficients of m^k. For every z along the real axis, there are at most $\mathrm{D_m}$ solutions to the polynomial equation $L_{\rm mz}(m,z)=0$. The solutions of the bivariate polynomial equation $L_{\rm mz}=0$ define a locus of points (m,z) in $\mathbb{C}\times\mathbb{C}$ referred to as a complex algebraic curve. Since the limiting density is over $\mathbb{R}$, we may focus on real values of z.

For almost every $z\in\mathbb{R}$, there will be $\mathrm{D_m}$ values of m. The exception consists of the singularities of $L_{\rm mz}(m,z)$. A singularity occurs at $z=z_0$ if:

- There is a reduction in the degree of m at z_0 so that there are less than $\mathrm{D_m}$ roots for $z=z_0$. This occurs when $l_{\mathrm{D_m}}(z_0)=0$. Poles of $L_{\rm mz}(m,z)$ occur if some of the m-solutions blow up to infinity.
- There are multiple roots of $L_{\rm mz}$ at z_0 so that some of the values of m coalesce.

The singularities constitute the so-called exceptional set of $L_{\rm mz}(m,z)$. Singularity analysis, in the context of algebraic functions, is a well studied problem [14] from which we know that the singularities of $L^A_{\rm mz}(m,z)$ are constrained to be *branch points*.

A *branch* of the algebraic curve $L_{\rm mz}(m,z)=0$ is the choice of a locally analytic function $m_j(z)$ defined outside the exceptional set of $L^A_{\rm mz}(m,z)$ together with a connected region of the $\mathbb{C}\times\mathbb{R}$ plane throughout which this particular choice $m_j(z)$ is analytic. These properties of singularities and branches of algebraic curve are helpful in determining the atomic and non-atomic component of the encoded probability density from $L_{\rm mz}$. We note that, as yet, we do not have a fully automated algorithm for extracting the limiting density function from the bivariate polynomial. Development of efficient computational algorithms that exploit the algebraic properties of the solution curve would be of great benefit to the community.

7.1 The atomic component

If there are any atomic components in the limiting density function, they will necessarily manifest themselves as poles of $L_{\rm mz}(m,z)$. This follows from the definition of the Stieltjes transform in (2.1). As mentioned in the discussion on the singularities of algebraic curves, the poles are located at the roots of $l_{\mathrm{D_m}}(z)$. These may be computed in MAPLE using the sequence of commands:

```
> Dm := degree(LmzA,m);
> lDmz := coeff(LmzA,m,Dm);
> poles := solve(lDmz=0,z);
```

We can then compute the Puiseux expansion about each of the poles at $z=z_0$. This can be computed in MAPLE using the `algcurves` package as:

```
> with(algcurves):
> puiseux(Lmz,z=pole,m,1);
```

For the pole at $z=z_0$, we inspect the Puiseux expansions for branches with leading term $1/(z_0-z)$. An atomic component in the limiting spectrum occurs if and only if the coefficient of such a branch is non-negative and not greater than one. This constraint ensures that the branch is associated with the Stieltjes transform of a valid probability distribution function.

Of course, as is often the case with algebraic curves, pathological cases can be easily constructed. For example, more than one branch of the Puiseux expansion might correspond to a candidate atomic component, *i.e.*, the coefficients are non-negative and not greater than one. In our experimentation, whenever this has happened it has been possible to eliminate the spurious branch by matrix theoretic arguments. Demonstrating this rigorously using analytical arguments remains an open problem.

Sometimes it is possible to encounter a double pole at $z=z_0$ corresponding to two admissible weights. In such cases, empirical evidence suggests that the branch with the largest coefficient (less than one) is the "right" Puiseux expansion though we have no theoretical justification for this choice.

7.2 The non-atomic component

The probability density function can be recovered from the Stieltjes transform by applying the inversion formula in (2.4). Since the Stieltjes transform is encoded in the bivariate polynomial L_{mz}, we accomplish this by first computing all $\mathrm{D_m}$ roots along $z \in \mathbb{R}$ (except at poles or singularities). There will be $\mathrm{D_m}$ roots of which one solution curve will be the "correct" solution , *i.e.*, the non-atomic component of the desired density function is the imaginary part of the correct solution normalized by π. In MATLAB , the $\mathrm{D_m}$ roots can be computed using the sequence of commands:

```
Lmz_roots = [];
x_range = [x_start:x_step:x_end];
for x = x_range
   Lmz_roots_unnorm = roots(sym2poly(subs(Lmz,z,x)));
   Lmz_roots = [Lmz_roots;
                real(Lmz_roots_unnorm) + i*imag(Lmz_roots_unnorm)/pi];
end
```

The density of the limiting eigenvalue distribution function can be, generically, be expressed in closed form when $\mathrm{D_m} = 2$. When using root-finding algorithms, for $\mathrm{D_m} = 2, 3$, the correct solution can often be easily identified; the imaginary branch will always appear with its complex conjugate. The density is just the scaled (by $1/\pi$) positive imaginary component.

When $\mathrm{D_m} \geq 4$, except when L_{mz} is bi-quadratic for $\mathrm{D_m} = 4$, there is no choice but to manually isolate the correct solution among the numerically computed $\mathrm{D_m}$ roots of the polynomial $L_{\mathrm{mz}}^{(}m, z)$ at each $z = z_0$. The class of algebraic random matrices whose eigenvalue density function can be expressed in closed form is thus a much smaller subset of the class of algebraic random matrices. When the underlying density function is compactly supported, the boundary points will be singularities of the algebraic curve.

In particular, when the probability density function is compactly supported and the boundary points are not poles, they occur at points where some values of m coalesce. These points are the roots of the discriminant of L_{mz}, computed in MAPLE as:

```
> PossibleBoundaryPoints = solve(discrim(Lmz,m),z);
```

We suspect that "nearly all" algebraic random matrices with compactly supported eigenvalue distribution will exhibit a square root type behavior near boundary points at which there are no poles. In the generic case, this will occur whenever the boundary points correspond to locations where two branches of the algebraic curve coalesce.

For a class of random matrices that includes a subclass of algebraic random matrices, this has been established in [27]. This endpoint behavior has also been observed orthogonally/unitarily invariant random matrices whose distribution has the element-wise joint density function of the form

$$f(\mathbf{A}) = C_N \exp\left(-N \mathrm{Tr}\, V(\mathbf{A})\right) d\mathbf{A}$$

where V is an even degree polynomial with positive leading coefficient and $d\mathbf{A}$ is the Lebesgue measure on $N \times N$ symmetric/Hermitian matrices. In [9], it is shown that these random matrices have a limiting mean eigenvalue density in the $N \to \infty$ limit that is algebraic and compactly supported. The behavior at the endpoint typically vanishes like a square root, though higher order vanishing at endpoints is possible and a full classification is made in [10]. In [17] it is shown that square root vanishing is generic. A similar classification for the general class of algebraic random matrices remains an open problem. This problem is of interest because of the intimate connection between the endpoint behavior and the Tracy-Widom distribution. Specifically, we conjecture that "nearly all" algebraic random matrices with compactly supported eigenvalue distribution whose density function vanishes as the square root at the endpoints will, with appropriate re-centering and rescaling, exhibit Tracy-Widom fluctuations.

Whether the encoded distribution is compactly supported or not, the $-1/z$ behavior of the real part of Stieltjes transform (the principal value) as $z \to \pm\infty$ helps isolate the correct solution. In our experience,

while multiple solution curves might exhibit this behavior, invariably only one solution will have an imaginary branch that, when normalized, will correspond to a valid probability density. Why this always appears to be the case for the operational laws described is a bit of a mystery to us.

Example: Consider the Marčenko-Pastur density encoded by L_{mz} given in Table 1(b). The Puiseux expansion about the pole at $z = 0$ (the only pole!), has coefficient $(1 - 1/c)$ which corresponds to an atom only when $c > 1$ (as expected using a matrix theoretic argument). Finally, the branch points at $(1 \pm \sqrt{c})^2$ correspond to boundary points of the compactly supported probability density. Figure 4 plots the real and imaginary parts of the algebraic curve for $c = 2$.

8. Enumerating the moments and free cumulants

In principle, the moments generating function can be extracted from $L_{\mu\mathrm{z}}$ by a Puiseux expansion of the algebraic function $\mu(z)$ about $z = 0$. When the moments of an algebraic probability distribution exist, there is additional structure in the moments and free cumulants that allows us to enumerate them efficiently. For an algebraic probability distribution, we conjecture that the moments of all order exist if and only if the distribution is compactly supported.

Definition 8.1 (Rational generating function)**.** *Let $\mathbb{R}[[x]]$ denote the ring of formal power series (or generating functions) in x with real coefficients. A formal power series (or generating function) $v \in \mathbb{R}[[u]]$ is said to be rational if there exist polynomials in u, $P(u)$ and $Q(u)$, $Q(0) \neq 0$ such that*

$$v(u) = \frac{P(u)}{Q(u)}.$$

Definition 8.2 (Algebraic generating function)**.** *Let $\mathbb{R}[[x]]$ denote the ring of formal power series (or generating functions) in x with real coefficients. A formal power series (or generating function) $v \in \mathbb{R}[[u]]$ is said to be algebraic if there exist polynomials in u, $P_0(u), \ldots, P_{\mathrm{D_u}}(u)$, not all identically zero, such that*

$$P_0(u) + P_1(u)v + \ldots + P_{\mathrm{D_v}}(u)v^{\mathrm{D_v}} = 0.$$

The degree of v is said to be $\mathrm{D_v}$.

Definition 8.3 (D-finite generating function)**.** *Let $v \in \mathbb{R}[[u]]$. If there exist polynomials $p_0(u), \ldots, p_d(u)$, such that*

$$p_d(u)v^{(d)} + p_{d-1}(u)v^{(d-1)} + \ldots + p_1(u)v^{(1)} + p_0(u) = 0, \tag{8.1}$$

where $v^{(j)} = d^j v/du^j$. Then we say that v is a D-finite (short for differentiably finite) generating function (or power series). The generating function, $v(u)$, is also referred to as a holonomic function.

Definition 8.4 (P-recursive coefficients)**.** *Let a_n for $n \geq 0$ denote the coefficients of a D-finite series v. If there exist polynomials $P_0, \ldots, P_e \in \mathbb{R}[n]$ with $P_e \neq 0$, such that*

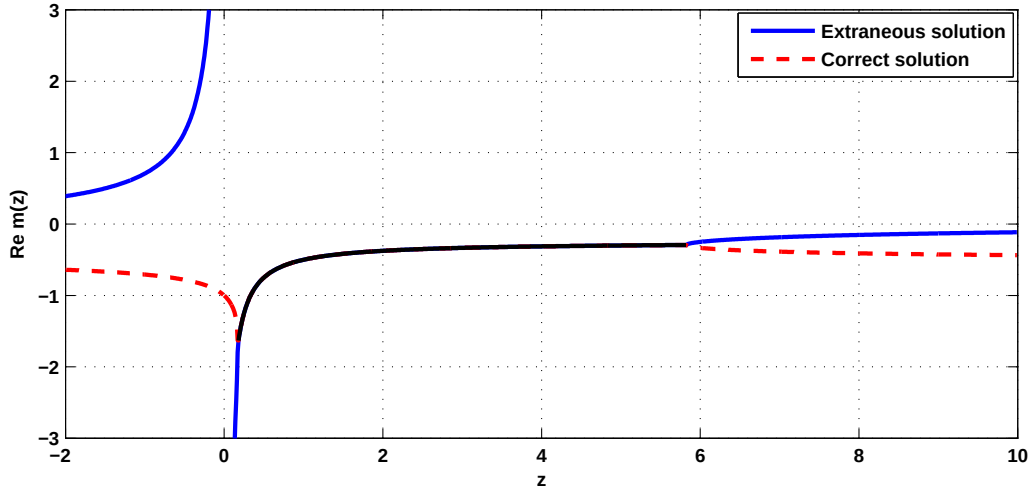
$$P_e(n)a_{n+e} + P_{e-1}(n)a_{n+e-1} + \ldots + P_0(n)a_n = 0,$$

for all $n \in \mathbb{N}$, then the coefficients a_n are said to be P-recursive (short for polynomially recursive).

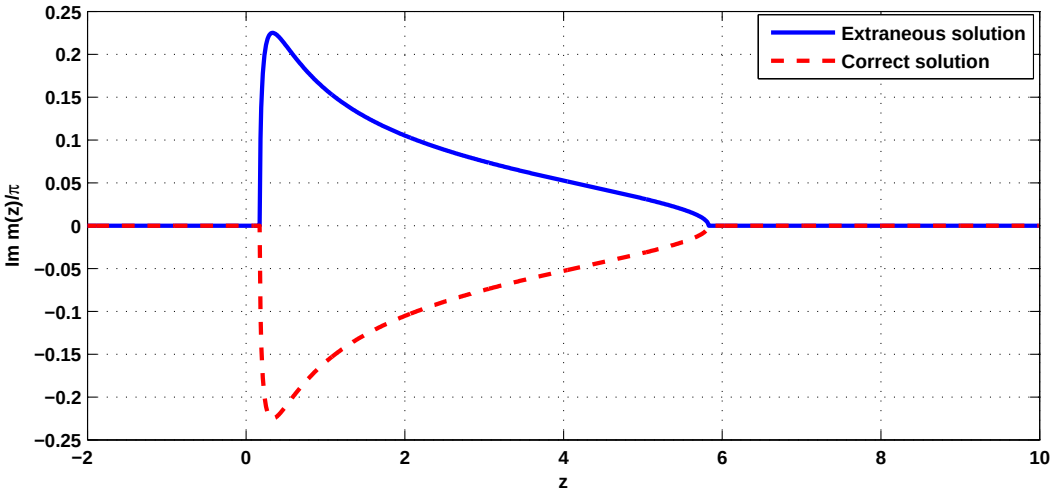
Proposition 8.5. *Let $v \in \mathbb{R}[[u]]$ be an algebraic power series of degree $\mathrm{D_v}$. Then v is D-finite and satisfies an equation (8.1) of order $\mathrm{D_v}$.*

Proof. A proof appears in Stanley [30, pp.187]. □

The structure of the limiting moments and free cumulants associated with algebraic densities is described next.



(a) Real component. The singularity at zero corresponds to an atom of weight 1/2. The branch points at $(1 \pm \sqrt{2})^2$ correspond to the boundary points of the region of support.



(b) Imaginary component normalized by π. The positive component corresponds to the encoded probability density function.

Figure 4: The real and imaginary components of the algebraic curve defined by the equation $L_{\mathrm{mz}}(m, z) = 0$, where $L_{\mathrm{mz}} \equiv czm^2 - (1 - c - z)\, m + 1$, which encodes the Marčenko-Pastur density. The curve is plotted for $c = 2$. The $-1/z$ behavior of the real part of the "correct solution" as $z \to \infty$ is the generic behavior exhibited by the real part of the Stieltjes transform of a valid probability density function.

Theorem 8.6. *If $f_A \in \mathcal{P}_{alg}$, and the moments exist, then the moment and free cumulant generating functions are algebraic power series. Moreover, both generating functions are D-finite and the coefficients are P-recursive.*

PROOF. If $f_A \in \mathcal{P}_{\text{alg}}$, then L^A_{mz} exists. Hence $L^A_{\mu z}$ and L^A_{rg} exist, so that $\mu_A(z)$ and $r_A(g)$ are algebraic power series. By Theorem 8.5 they are D-finite; the moments and free cumulants are hence P-recursive. □

There are powerful symbolic tools available for enumerating the coefficients of algebraic power series. The MAPLE based package `gfun` is one such example [24]. From the bivariate polynomial $L_{\mu z}$, we can obtain the series expansion up to degree `expansion_degree` by using the commands:

```
> with(gfun):
> MomentSeries = algeqtoseries(Lmyuz,z,myu,expansion_degree,'pos_slopes');
```

The option `pos_slopes` computes only those branches tending to zero. Similarly, the free cumulants can be enumerated from L_{rg} using the commands:

```
> with(gfun):
> FreeCumulantSeries = algeqtoseries(Lrg,g,r,expansion_degree,'pos_slopes');
```

For computing expansions to a large order, it is best to work with the recurrence relation. For an algebraic power series $v(u)$, the first `number_of_terms` coefficients can be computed from L_{uv} using the sequence of commands:

```
> with(gfun):
> deq := algeqtodiffeq(Luv,v(u));
> rec := diffeqtorec(deq,v(u),a(n));
> p_generator := rectoproc(rec,a(n),list):
> p_generator(number_of_terms);
```

Example: Consider the Marčenko-Pastur density encoded by the bivariate polynomials listed in Table 1(b). Using the above sequence of commands, we can enumerate the first five terms of its moment generating function as

$$\mu(z) = 1 + z + (c+1)\,z^2 + \left(3\,c + c^2 + 1\right) z^3 + \left(6\,c^2 + c^3 + 6\,c + 1\right) z^4 + O\left(z^5\right).$$

The moment generating function is a D-Finite power series and satisfies the second order differential equation

$$-z + zc - 1 + (-z - zc + 1)\,\mu(z) + \left(z^3c^2 - 2\,z^2c - 2\,z^3c + z - 2\,z^2 + z^3\right)\frac{d}{dz}\mu(z) = 0,$$

with initial condition $\mu(0) = 1$. The moments $M_n = a(n)$ themselves are P-recursive satisfying the finite depth recursion

$$\left(-2\,c + c^2 + 1\right) na(n) + ((-2 - 2\,c)\,n - 3\,c - 3)\,a(n+1) + (3+n)\,a(n+2) = 0$$

with the initial conditions $a(0) = 1$ and $a(1) = 1$. The free cumulants can be analogously computed.

What we find rather remarkable is that for algebraic random matrices, it is often possible to enumerate the moments in closed form even when the limiting density function cannot. The linear recurrence satisfied by the moments may be used to analyze their asymptotic growth.

When using the sequence of commands described, sometimes more than one solution might emerge. In such cases, we have often found that one can identify the correct solution by checking for the positivity of even moments or the condition $\mu(0) = 1$. More sophisticated arguments might be needed for pathological cases. It might involve verifying, using techniques such as those in [1], that the coefficients enumerated correspond to the moments a valid distribution function.

9. Computational free probability

9.1 Moments of random matrices and asymptotic freeness

Assume we know the eigenvalue distribution of two matrices $\mathbf{A}$ and $\mathbf{B}$. In general, using that information alone, we cannot say much about the the eigenvalue distribution of the sum $\mathbf{A}+\mathbf{B}$ of the matrices since eigenvalues of the sum of the matrices depend on the eigenvalues of $\mathbf{A}$ and the eigenvalues of $\mathbf{B}$, and also on the relation between the eigenspaces of $\mathbf{A}$ and of $\mathbf{B}$. However, if we pose this question in the context of $N\times N$-random matrices, then in many situations the answer becomes deterministic in the limit $N\to\infty$. Free probability provides the analytical framework for characterizing this limiting behavior.

Definition 9.1. *Let $\mathbf{A}=(\mathbf{A}_N)_{N\in\mathbb{N}}$ be a sequence of $N\times N$-random matrices. We say that $\mathbf{A}$ has a limit eigenvalue distribution if the limit of all moments*

$$\alpha_n := \lim_{N\to\infty} E[\mathrm{tr}(\mathbf{A}_N^n)] \qquad (n\in\mathbb{N})$$

exists, where E denotes the expectation and tr *the normalized trace.*

Using the language of limit eigenvalue distribution as in Definition 9.1, our question becomes: Given two random matrix ensembles of $N\times N$-random matrices, $\mathbf{A}=(\mathbf{A}_N)_{N\in\mathbb{N}}$ and $\mathbf{B}=(\mathbf{B}_N)_{N\in\mathbb{N}}$, with limit eigenvalue distribution, does their sum $\mathbf{C}=(\mathbf{C}_N)_{N\in\mathbb{N}}$, with $\mathbf{C}_N=\mathbf{A}_N+\mathbf{B}_N$, have a limit eigenvalue distribution, and furthermore, can we calculate the limit moments α_n^C of $\mathbf{C}$ out of the limit moments $(\alpha_k^{\mathbf{A}})_{k\geq 1}$ of $\mathbf{A}$ and the limit moments $(\alpha_k^B)_{k\geq 1}$ of $\mathbf{B}$ in a deterministic way. It turns out that this is the case if the two ensembles are in generic position, and then the rule for calculating the limit moments of $\mathbf{C}$ are given by Voiculescu's concept of "freeness."

Theorem 9.2 (Voiculescu [36])**.** *Let $\mathbf{A}$ and $\mathbf{B}$ be two random matrix ensembles of $N\times N$-random matrices, $\mathbf{A}=(\mathbf{A}_N)_{N\in\mathbb{N}}$ and $\mathbf{B}=(\mathbf{B}_N)_{N\in\mathbb{N}}$, each of them with a limit eigenvalue distribution. Assume that $\mathbf{A}$ and $\mathbf{B}$ are independent (i.e., for each $N\in\mathbb{N}$, all entries of $\mathbf{A}_N$ are independent from all entries of $\mathbf{B}_N$), and that at least one of them is unitarily invariant (i.e., for each N, the joint distribution of the entries does not change if we conjugate the random matrix with an arbitrary unitary $N\times N$ matrix). Then $\mathbf{A}$ and $\mathbf{B}$ are asymptotically free in the sense of the following definition.*

Definition 9.3 (Voiculescu [33])**.** *Two random matrix ensembles $\mathbf{A}=(\mathbf{A}_N)_{N\in\mathbb{N}}$ and $\mathbf{B}=(\mathbf{B}_N)_{N\in\mathbb{N}}$ with limit eigenvalue distributions are* asymptotically free *if we have for all $p\geq 1$ and all $n(1),m(1),\ldots,n(p),m(p)\geq 1$ that*

$$\lim_{N\to\infty} E\Big[\mathrm{tr}\big\{(\mathbf{A}_N^{n(1)} - \alpha_{n(1)}^A 1)\cdot(\mathbf{B}_N^{m(1)} - \alpha_{m(1)}^B 1)\cdots(\mathbf{A}^{n(p)} - \alpha_{n(p)}^A 1)\cdot(\mathbf{B}^{m(p)} - \alpha_{m(p)}^B 1)\big\}\Big] \;=\; 0$$

In essence, asymptotic freeness is actually a rule which allows to calculate all mixed moments in $\mathbf{A}$ and $\mathbf{B}$, *i.e.*, all expressions of the form

$$\lim_{N\to\infty} E[\mathrm{tr}(\mathbf{A}^{n(1)}\mathbf{B}^{m(1)}\mathbf{A}^{n(2)}\mathbf{B}^{m(2)}\cdots\mathbf{A}^{n(p)}\mathbf{B}^{m(p)})]$$

out of the limit moments of $\mathbf{A}$ and the limit moments of $\mathbf{B}$. In particular, this means that all limit moments of $\mathbf{A}+\mathbf{B}$ (which are sums of mixed moments) exist, thus $\mathbf{A}+\mathbf{B}$ has a limit distribution, and are actually determined in terms of the limit moments of $\mathbf{A}$ and the limit moments of $\mathbf{B}$. For more on free probability, including extensions to the setting where the moments do not exist, we refer the reader to [6, 15, 21, 37].

We now clarify the connection between the operational law of a subclass of algebraic random matrices and the convolution operations of free probability. This will bring into sharp focus how the polynomial method constitutes a framework for computational free probability theory.

Free additive convolution	$f_{A+B} = f_A \boxplus f_B$	$L_{\mathrm{rg}}^{A+B} = L_{\mathrm{rg}}^{A} \boxplus_{\mathrm{r}} L_{\mathrm{rg}}^{B}$
Free multiplicative convolution	$f_{A\times B} = f_A \boxtimes f_B$	$L_{\mathrm{sy}}^{A\times B} = L_{\mathrm{sy}}^{A} \boxtimes_{\mathrm{s}} L_{\mathrm{sy}}^{B}$

Table 10: Implicit representation of the free convolution of two algebraic probability densities.

Proposition 9.4. *Let $\mathbf{A}_N \overset{\mathrm{p}}{\longmapsto} A$ and $\mathbf{B}_N \overset{\mathrm{p}}{\longmapsto} B$ be two asymptotically free random matrix sequences as in Definition 9.1. Then $\mathbf{A}_N + \mathbf{B}_N \overset{\mathrm{p}}{\longmapsto} A + B$ and $\mathbf{A}_N \times \mathbf{B}_N \overset{\mathrm{p}}{\longmapsto} AB$ (where the product is defined whenever $\mathbf{A}_N \times \mathbf{B}_N$ has real eigenvalues for every $\mathbf{A}_N$ and $\mathbf{B}_N$) with the corresponding limit eigenvalue density functions, f_{A+B} and f_{AB} given by*

$$f_{A+B} = f_A \boxplus f_B \tag{9.1a}$$

$$f_{AB} = f_A \boxtimes f_B \tag{9.1b}$$

where $\boxplus$ denotes free additive convolution and $\boxtimes$ denotes free multiplicative convolution. These convolution operations can be expressed in terms of the R and S transforms as described in Propositions 5.22 and 5.24 respectively.

Proof. This result appears for density functions with compact support in [34,35]. It was later strengthened to the case of density functions with unbounded support. See [15] for additional details and references. □

In Theorems 5.23 and 5.25 we, in effect, showed that the free convolution of algebraic densities produces an algebraic density. This stated succinctly next.

Corollary 9.5. *Algebraic probability distributions form a semi-group under free additive convolution.*

Corollary 9.6. *Algebraic distributions with positive semi-definite support form a semi-group under free multiplicative convolution.*

This establishes a framework for computational free probability theory by identifying the class of distributions for which the free convolution operations produce a "computable" distribution.

9.2 Implicitly encoding the free convolution computations

The computational framework established relies on being able to implicitly encode free convolution computations as a resultant computation on appropriate bivariate polynomials as in Table 10. This leads to the obvious question: Are there other more *effective* ways to implicitly encode free convolution computations? The answer to this rhetorical question will bring into sharp focus the reason why the bivariate polynomial encoding at the heart of the polynomial method is indispensable for any symbolic computational implementation of free convolution. First, we answer the analogous question about the most effective encoding for classical convolution computations.

Recall that classical convolution can be expressed in terms of the Laplace transform of the distribution function. In what follows, we assume that the distributions have finite moments[1]. Hence the Laplace transform can be written as a formal exponential moment generating function. Classical additive and multiplicative convolution of two distributions produces a distribution whose exponential moment generating function equals the series (or Cauchy) product and the coefficient-wise (or Hadamard) product of the individual exponential moment generating functions, respectively. Often, however, the Laplace transform of

[1] In the general case, tools from complex analysis can be used to extend the argument.

either or both the individual distributions being convolved cannot be written in closed form. The next best thing to do then is to find an implicit way to encode the Laplace transform and to do the convolution computations via this representation.

When this point of view is adopted, the task of identifying candidate encodings is reduced to finding the class of representations of the exponential generating function that remains closed under the Cauchy and Hadamard product. Clearly, rational generating functions (see Definition 8.1) satisfy this requirement. It is shown in Theorem 6.4.12 [30, pp.194], that D-finite generating functions (see Definition 8.3) satisfy this requirement as well.

Proposition 8.5 establishes that all algebraic generating functions (see Definition 8.2) and by extension, rational generating functions, are also D-finite. However, not all D-finite generating functions are algebraic (see Exercise 6.1 [30, pp. 217] for a counter-example) so that algebraic generating functions do not satisfy the closure requirement. Furthermore, from Proposition 6.4.3 and Theorem 6.4.12 in [30], if the *ordinary* generating function is D-finite then so is the *exponential* generating function and vice versa. Thus D-finite generating functions are the largest class of generating functions for which classical convolution computations can be performed via an implicit representation.

In the context of developing a computational framework based on the chosen implicit representation, it is important to consider computability and algorithmic efficiency issues. The class of D-finite functions is well-suited in that regard as well [24] so that we regard it as the most *effective* class of representations in which the classical convolution computations may be performed implicitly.

However, this class is inadequate for performing free convolution computations implicitly. This is a consequence of the prominent role occupied in this theory by ordinary generating functions. Specifically, the ordinary formal R and S power series, are obtained from the ordinary moment generating function by functional inversion (or reversion), and are the key ingredients of free additive and multiplicative convolution (see Propositions 9.4, 5.22 and 5.24). The task of identifying candidate encodings is thus reduced to finding the class of representations of the ordinary moment generating function that remains closed under addition, the Cauchy product, *and* reversion. D-finite functions only satisfy the first two conditions and are hence unsuitable representations.

Algebraic functions do, however, satisfy all three conditions. The algorithmic efficiency of computing the resultant (see Section 4.2) justifies our labelling of the bivariate polynomial encoding as the most *effective* way of implicitly encoding free convolution computations. The candidacy of constructibly D-finite generating functions [5], which do not contain the class of D-finite functions but do contain the class of algebraic functions, merits further investigation since they are closed under reversion, addition and multiplication. Identifying classes of representations of generating functions for which *both* the classical and free convolution computations can be performed implicitly and effectively remains an important open problem.

10. Applications

We illustrate the use of the computational techniques developed in Section 6. with some examples. Documented MATLAB implementation of the polynomial method is available via the RMTool package [22] from `http://www.mit.edu/~raj/rmtool/`; the examples considered in this article, along with many more, appear there and in [20].

10.1 The Jacobi random matrix

The Jacobi matrix ensemble is defined in terms of two independent Wishart matrices $\mathbf{W}_1(c_1)$ and $\mathbf{W}_2(c_2)$ as $\mathbf{J} = (\mathbf{I} + \mathbf{W}_2(c_2)\,\mathbf{W}_1^{-1}(c_1))^{-1}$. The subscripts are not to be confused for the size of the matrices. Listing the computational steps needed to generate a realization of this ensemble, as in Table 11, is the easiest way to identify the sequence of random matrix operations needed to obtain L^J_{mz}.

Transformation	Numerical MATLAB code	Symbolic MATLAB code
Initialization	% Pick n, c1, c2 N1=n/c1; N2=n/c2;	% Define symbolic variables syms m c z;
$\mathbf{A}_1 = \mathbf{I}$	A1 = eye(n,n);	Lmz1 = m*(1-z)-1;
$\mathbf{A}_2 = \mathbf{W}_1(c_1) \times \mathbf{A}_1$	G1 = randn(n,N1)/sqrt(N1); W1 = G1*G1'; A2 = W1*A1;	Lmz2 = AtimesWish(Lmz1,c1);
$\mathbf{A}_3 = \mathbf{A}_2^{-1}$	A3 = inv(A2);	Lmz3 = invA(Lmz2);
$\mathbf{A}_4 = \mathbf{W}_2(c_2) \times \mathbf{A}_3$	G2 = randn(n,N2)/sqrt(N2); W2 = G2*G2'; A4 = W2*A3;	Lmz4 = AtimesWish(Lmz3,c2);
$\mathbf{A}_5 = \mathbf{A}_4 + \mathbf{I}$	A5 = A4+I;	Lmz5 = shiftA(Lmz4,1);
$\mathbf{A}_6 = \mathbf{A}_5^{-1}$	A6 = inv(A5);	Lmz6 = invA(Lmz5);

Table 11: Sequence of MATLAB commands for sampling the Jacobi ensemble. The functions used to generate the corresponding bivariate polynomials symbolically are listed in Table 7

.

We first start off with $\mathbf{A}_1 = \mathbf{I}$. The bivariate polynomial that encodes the Stieltjes transform of its eigenvalue distribution function is given by

$$L^1_{\rm mz}(m,z) = (1-z)m - 1. \tag{10.1}$$

For $\mathbf{A}_2 = \mathbf{W}_1(c_1) \times \mathbf{A}_1$, we can use (5.30) to obtain the bivariate polynomial

$$L^2_{\rm mz}(m,z) = z\,c_1\,m^2 - (-c_1 - z + 1)\,m + 1. \tag{10.2}$$

For $\mathbf{A}_3 = \mathbf{A}_2^{-1}$, from (5.6), we obtain the bivariate polynomial

$$L^3_{\rm mz}(m,z) = z^2 c_1 m^2 + (c_1\,z + z - 1)\,m + 1. \tag{10.3}$$

For $A_4 = \mathbf{W}_2(c_2) \times \mathbf{A}_3$. We can use (5.30) to obtain the bivariate polynomial

$$L^4_{\rm mz}(m,z) = \left(c_1\,z^2 + c_2\,z\right)m^2 + (c_1\,z + z - 1 + c_2)\,m + 1. \tag{10.4}$$

For $\mathbf{A}_5 = \mathbf{A}_4 + \mathbf{I}$, from (5.6), we obtain the bivariate polynomial

$$L^5_{\rm mz}(m,z) = \left((z-1)^2\,c_1 + c_2\ (z-1)\right)m^2 + (c_1\ (z-1) + z - 2 + c_2)\,m + 1. \tag{10.5}$$

Finally, for $\mathbf{J} = \mathbf{A}_6 = \mathbf{A}_5^{-1}$, from (5.6), we obtain the required bivariate polynomial

$$\begin{aligned} L^J_{\rm mz}(m,z) \equiv L^6_{\rm mz}(m,z) = &\left(c_1\,z + z^3 c_1 - 2\,c_1\,z^2 - c_2\,z^3 + c_2\,z^2\right)m^2 \\ &+ \left(-1 + 2\,z + c_1 - 3\,c_1\,z + 2\,c_1\,z^2 + c_2\,z - 2\,c_2\,z^2\right)m - c_2\,z - c_1 + 2 + c_1\,z. \end{aligned} \tag{10.6}$$

Using matrix theoretic arguments, it is clear that the random matrix ensembles $\mathbf{A}_3, \ldots \mathbf{A}_6$ are defined only when $c_1 < 1$. There will be an atomic mass of weight $(1 - 1/c_2)$ at 1 whenever $c_2 > 1$. The non-atomic component of the distribution will have a region of support $\mathcal{S}^{\cap} = (a_-, a_+)$. The limiting density function for each of these ensembles can be expressed as

$$\boxed{f_{A_i}(x) = \frac{\sqrt{(x-a_-)(a_+-x)}}{2\,\pi\,l_2(x)} \qquad \text{for } a_- < x < a_+,} \tag{10.7}$$

for $i = 2, \ldots, 6$, where a_-, a_+ , where the polynomials $l_2(x)$ are listed in Table 12. The moments for

	$l_2(x)$	$a_\pm$
A_2	$x\,c_1$	$(1 \pm \sqrt{c_1})^2$
A_3	$x^2\,c_1$	$\dfrac{1}{(1 \mp \sqrt{c_1})^2}$
A_4	$c_1 x^2 + c_2 x$	$\dfrac{1 + c_1 + c_2 - c_1 c_2 \pm 2\sqrt{c_1 + c_2 - c_1 c_2}}{(1 - c_1)^2}$
A_5	$c_1(x-1)^2 + c_2(x-1)$	$\dfrac{c_1^2 - c1 + 2 + c_2 - c_1 c_2 \pm 2\sqrt{c_1 + c_2 - c_1 c_2}}{(1 - c_1)^2}$
A_6	$\left(c_1\,x + x^3 c_1 - 2\,c_1\,x^2 - c_2\,x^3 + c_2\,x^2\right)$	$\dfrac{(1 - c_1)^2}{c_1^2 - c1 + 2 + c_2 - c_1 c_2 \mp 2\sqrt{c_1 + c_2 - c_1 c_2}}$

Table 12: Parameters for determining the limiting eigenvalue density function using (10.7).

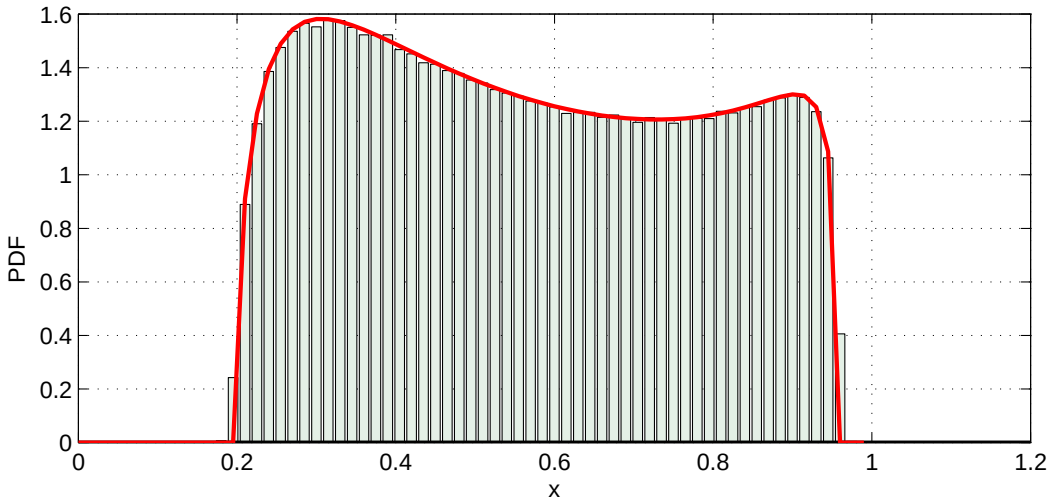


Figure 5: The limiting density (solid line), $f_{A_6}(x)$, given by (10.7) with $c_1 = 0.1$ and $c_2 = 0.625$ is compared with the normalized histogram of the eigenvalues of a Jacobi matrix generated using the code in Table 11 over 4000 Monte-Carlo trials with $n = 100$, $N_1 = n/c_1 = 1000$ and $N_2 = n/c_2 = 160$.

the general case when $c_1 \neq c_2$ can be enumerated using the techniques described; they will be quite messy. Instead, consider the special case when $c_1 = c_2 = c$. Using the tools described, the first four terms of the moment series, $\mu(z) = \mu_J(z)$, can be computed directly from $L^J_{\mu z}$ as

$$\mu(z) = \frac{1}{2} + \left(\frac{1}{8}\,c + \frac{1}{4}\right) z + \left(\frac{3}{16}\,c + \frac{1}{8}\right) z^2 + \left(\frac{1}{32}\,c^2 + \frac{3}{16}\,c - \frac{1}{128}\,c^3 + \frac{1}{16}\right) z^3 + \left(-\frac{5}{256}\,c^3 + \frac{5}{64}\,c^2 + \frac{5}{32}\,c + \frac{1}{32}\right) z^4 + O\left(z^5\right).$$

The moment generating function satisfies the differential equation

$$-3\,z + 2 + zc + \left(-6\,z^2 + z^3 + 10\,z + z^3 c^2 - 2\,z^3 c - 4\right)\mu\left(z\right) + \left(z^4 - 5\,z^3 - 2\,z^4 c + 8\,z^2 + z^4 c^2 + 2\,z^3 c - 4\,z - z^3 c^2\right)\frac{d}{dz}\mu\left(z\right) = 0,$$

with the initial condition $\mu(0) = 1$. The moments $a(n) = M_n$ themselves are P-recursive and obtained by

the recursion

$$\left(-2\,c+c^2+1+\left(-2\,c+c^2+1\right)n\right)a\left(n\right)+\left(\left(-5+2\,c-c^2\right)n-11+2\,c-c^2\right)a\left(n+1\right)$$
$$+\left(26+8\,n\right)a\left(n+2\right)+\left(-16-4\,n\right)a\left(n+3\right)=0,$$

with the initial conditions $a(0)=1/2$, $a(1)=1/8\,c+1/4$, and $a(2)=3/16\,c+1/8$. We can similarly compute the recursion for the free cumulants, $a(n)=K_{n+1}$, as

$$nc^2a\left(n\right)+\left(12+4\,n\right)a\left(n+2\right)=0,$$

with the initial conditions $a(0)=1/2$, and $a(1)=1/8\,c$.

10.2 Random compression of a matrix

Theorem 10.1. *Let $\mathbf{A}_N \mapsto A \in \mathcal{P}_{alg}$. Let $\mathbf{Q}_N$ be an $N \times N$ Haar unitary/orthogonal random matrix independent of $\mathbf{A}_N$. Let $\mathbf{B}_n$ be the upper $n \times n$ block of $\mathbf{Q}_N\mathbf{A}_N\mathbf{Q}_N^{'}$. Then*

$$\mathbf{B}_n \mapsto B \in \mathcal{P}_{alg}$$

as $n/N \to c$ for $n, N \to \infty$.

Proof. Let $\mathbf{P}_N$ be an $N \times N$ projection matrix

$$\mathbf{P}_N \equiv \mathbf{Q}_N \begin{bmatrix} \mathbf{I}_n & \\ & \mathbf{0}_{N-n} \end{bmatrix} \mathbf{Q}_N^{'}.$$

By definition, $\mathbf{P}_N$ is an atomic matrix so that $\mathbf{P}_N \to P \in \mathcal{M}_{\text{alg}}$ as $n/N \to c$ for $n, N \to \infty$. Let $\widetilde{\mathbf{B}}_N = \mathbf{P}_N \times \mathbf{A}_N$. By Corollary 5.27, $\widetilde{\mathbf{B}}_N \to \widetilde{B} \in \mathcal{M}_{\text{alg}}$. Finally, from Theorem 5.11, we have that $\mathbf{B}_n \to B \in \mathcal{M}_{\text{alg}}$. □

The proof above provides a recipe for computing the bivariate polynomial L^B_{mz} explicitly as a function of L^A_{mz} and the compression factor c. For this particular application, however, one can use first principles [29] to directly obtain the relationship

$$r_B(g) = r_A(c\,g),$$

expressed in terms of the R transform. This translates into the operational law

$$\boxed{L^B_{\text{rg}}(r,g) = L^A_{\text{rg}}(r, c\,g).} \tag{10.8}$$

Example: Consider the atomic matrix $\mathbf{A}_N$ half of whose eigenvalues are equal to one while the remainder are equal to zero. Its eigenvalue distribution function is given by (3.6). From the bivariate polynomial, L^A_{rg} in Table 1(a) and (10.8) it can be show that the limiting eigenvalue distribution function of $\mathbf{B}_n$, constructed from $\mathbf{A}_N$ as in Theorem 10.1, is encoded by the polynomial

$$L^B_{\text{mz}} = \left(-2\,cz^2+2\,cz\right)m^2-\left(-2\,c+4\,cz+1-2\,z\right)m-2\,c+2,$$

where c is the limiting compression factor. Poles occur at $z=0$ and $z=1$. The leading terms of the Puiseux expansion of the two branches about the poles at $z=z_0$ are

$$\left\{\left(\frac{z-z_0}{-2\,c+4\,c^2}+\frac{1-2\,c}{2c}\right)\frac{1}{z-z_0},\frac{2\,c-2}{-1+2\,c}\right\}.$$

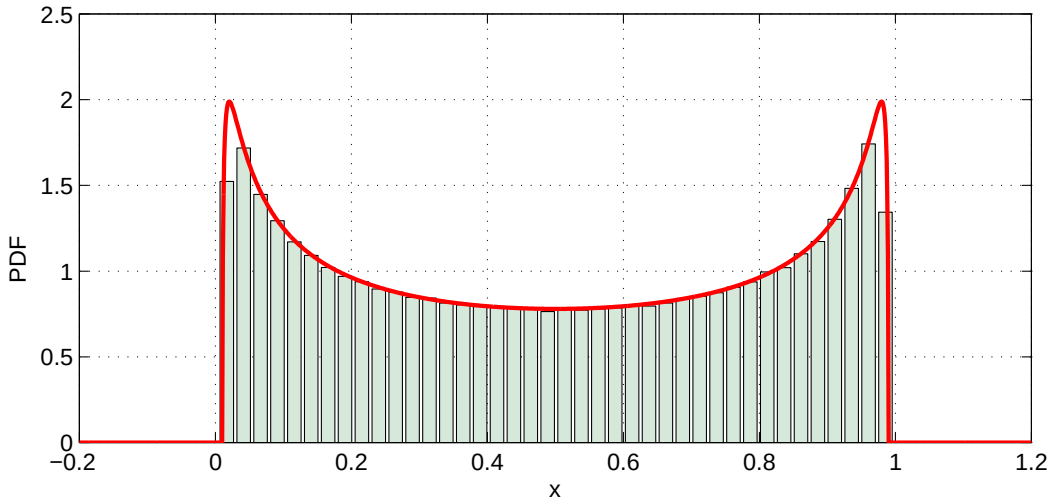


Figure 6: The limiting eigenvalue density function (solid line) of the top $0.4N \times 0.4N$ block of a randomly rotated matrix is compared with the experimental histogram collected over 4000 trials with $N = 200$. Half of the eigenvalues of the original matrix were equal to one while the remainder were equal to zero.

It can be easily seen that when $c > 1/2$, the Puiseux expansion about the poles $z = z_0$ will correspond to an atom of weight $w_0 = (2c-1)/2c$. Thus the limiting eigenvalue distribution function has density

$$\boxed{f_B(x) = \max\left(\frac{2c-1}{2c}, 0\right)\delta(x) + \frac{1}{\pi}\frac{\sqrt{(x-a_-)(a_+-x)}}{2\,xc - 2\,cx^2} I_{[a_-,a_+]} + \max\left(\frac{2c-1}{2c}, 0\right)\delta(x-1),} \tag{10.9}$$

where $a_\pm = 1/2 \pm \sqrt{-c^2+c}$. Figure 10.2 compares the theoretical prediction in (10.9) with a Monte-Carlo experiment for $c = 0.4$. From the associated bivariate polynomial

$$L^B_{\mu z} \equiv (-2\,c + 2\,cz)\,\mu^2 + (z - 2 - 2\,cz + 4\,c)\,\mu - 2\,c + 2,$$

we obtain two series expansions whose branches tend to zero. The first four terms of the series are given by

$$1 + \frac{1}{2}z + \frac{1+c}{4}z^2 + \frac{3+c}{8}z^3 + O\left(z^4\right), \tag{10.10}$$

and,

$$\frac{c-1}{c} + \frac{c-1}{2c}z - \frac{(c-1)\,(-2+c)}{4c}z^2 - \frac{(c-1)\,(3\,c-4)}{8c}z^3 + O\left(z^4\right), \tag{10.11}$$

respectively. Since $c \leq 1$, the series expansion in (10.11) can be eliminated since $\mu(0) := \int dF^B(x) = 1$. Thus the coefficients of the series in (10.10) are the correct moments of the limiting eigenvalue distribution. A recursion for the moments can be readily derived using the techniques developed earlier.

10.3 Free additive convolution of equilibrium measures

Equilibrium measures are a fascinating topic within random matrix theory. They arise in the context of research that examines why very general random models for random matrices exhibit universal behavior in the large matrix limit. Suppose we are given a potential $V(x)$ then we consider a sequence of Hermitian, unitarily invariant random matrices $\mathbf{A}_N$, the joint distribution of whose elements is of the form

$$P(\mathbf{A}_N) \propto \exp\left(-N \operatorname{Tr} V(\mathbf{A}_N)\right) d\mathbf{A}_N,$$

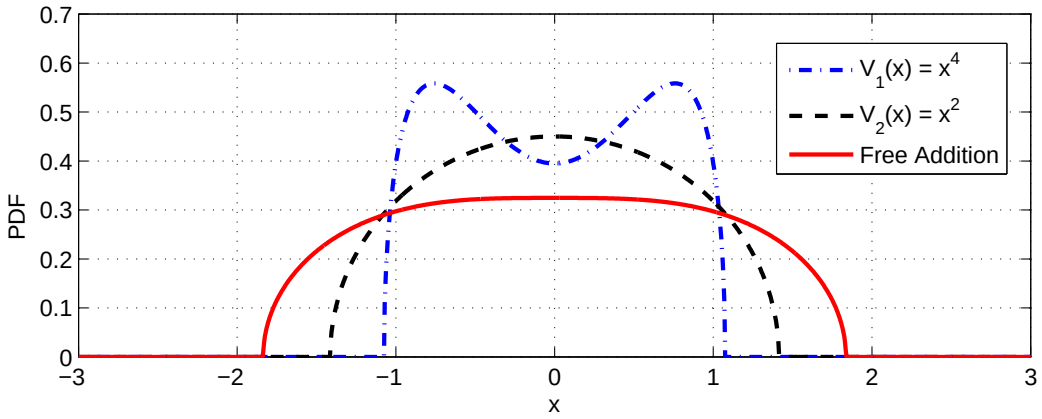


Figure 7: Additive convolution of equilibrium measures corresponding to potentials $V_1(x)$ and $V_2(x)$.

where $d\mathbf{A}_N = \prod_{i \leq j} (d\mathbf{A}_N)_{ij}$. The equilibrium measure, when it exists, is the unique probability distribution function that minimizes the logarithmic energy (see [11] for additional details). The resulting equilibrium measure depends explicitly on the potential $V(x)$ and can be explicitly computed for some potentials. In particular, for potentials of the form $V(x) = t\,x^{2m}$, the Stieltjes transform of the resulting equilibrium measure is an algebraic function [11, Chp. 6.7, pp. 174-175] so that the equilibrium measure is an algebraic distribution. Hence we can formally investigate the additive convolution of equilibrium measures corresponding to two different potentials. For $V_1(x) = x^2$, the equilibrium measure is the (scaled) semi-circle distribution encoded by the bivariate polynomial

$$L^A_{\mathrm{mz}} \equiv m^2 + 2\,m\,z + 2.$$

For $V_2(x) = x^4$, the equilibrium measure is encoded by the bivariate polynomial

$$L^B_{\mathrm{mz}} \equiv 1/4\,m^2 + mz^3 + z^2 + 2/9\,\sqrt{3}.$$

Since $\mathbf{A}_N$ and $\mathbf{B}_N$ are unitarily invariant random matrices, if $\mathbf{A}_N$ and $\mathbf{B}_N$ are independent, then the limiting eigenvalue distribution function of $\mathbf{C}_N = \mathbf{A}_N + \mathbf{B}_N$ can be computed from L^A_{mz} and L^B_{mz}. The limiting eigenvalue density function $f_C(x)$ is the free additive convolution of f_A and f_B. The MATLAB command `LmzC = AplusB(LmzA,LmzB);` will produce the bivariate polynomial

$$L^C_{\mathrm{mz}} = -9\,m^4 - 54\,m^3 z + \left(-108\,z^2 - 36\right) m^2 - \left(72\,z^3 + 72\,z\right) m - 72\,z^2 - 16\,\sqrt{3}.$$

Figure 10.3 plots the probability density function for the equilibrium measure for the potentials $V_1(x) = x^2$ and $V_2(x) = x^4$ as well as the free additive convolution of these measures. The interpretation of the resulting measuring in the context of potential theory is not clear. The matrix $\mathbf{C}_N$ will no longer be unitarily invariant so it is pointless to look for a potential $V_3(x)$ for which F^C is an equilibrium measure. The tools and techniques developed in this article might prove useful in further explorations.

10.4 Algebraic sample covariance matrices

The (broader) class of algebraic Wishart sample covariance matrices for which this framework applies is described next.

Theorem 10.2. *Let $\mathbf{A}_n \xmapsto{\text{p}} A \in \mathcal{M}_{alg}$, and $\mathbf{B}_N \xmapsto{\text{p}} B \in \mathcal{M}_{alg}$ be algebraic covariance matrices with $\mathbf{G}_{n,N}$ denoting an $n \times N$ (pure) Gaussian random matrix (see Definition 5.13). Let $\mathbf{X}_{n,N} = \mathbf{A}_n^{1/2}\mathbf{G}_{n,N}\mathbf{B}_N^{1/2}$. Then*

$$\mathbf{S}_n = \mathbf{X}_{n,N}\mathbf{X}_{n,N}^{'} \xmapsto{\text{p}} S \in \mathcal{M}_{alg},$$

as $n, N \to \infty$ and $c_N = n/N \to c$.

Proof. Let $\mathbf{Y}_{n,N} \equiv \mathbf{G}_{n,N}\mathbf{B}_N^{1/2}$, $\mathbf{T}_n \equiv \mathbf{Y}_{n,N}\mathbf{Y}'_{n,N}$ and $\widetilde{\mathbf{T}}_N = \mathbf{Y}'_{n,N}\mathbf{Y}_{n,N}$. Thus $\mathbf{S}_n = \mathbf{A}_n \times \mathbf{T}_n \equiv \mathbf{A}_n^{1/2}\mathbf{T}_n\mathbf{A}_n^{1/2}$. The matrix $\mathbf{T}_n$, as defined, is invariant under orthogonal/unitary transformations, though the matrix $\widetilde{\mathbf{T}}_N$ is not. Hence, by Corollary 5.27, and since $\mathbf{A}_n \mapsto A \in \mathcal{M}_{\text{alg}}$, $\mathbf{S}_n \mapsto S \in \mathcal{M}_{\text{alg}}$ whenever $\mathbf{T}_n \mapsto T \in \mathcal{M}_{\text{alg}}$.

From Theorem 5.8, $\mathbf{T}_n \mapsto T \in \mathcal{M}_{\text{alg}}$ if $\widetilde{\mathbf{T}}_N \mapsto \widetilde{T} \in \mathcal{M}_{\text{alg}}$. The matrix $\widetilde{\mathbf{T}}_N = \mathbf{B}_N^{1/2}\mathbf{G}'_{n,N}\mathbf{G}_{n,N}\mathbf{B}_N^{1/2}$ is clearly algebraic by application of Corollary 5.27 and Theorem 5.6 since $\mathbf{B}_N$ is algebraic and $\mathbf{G}'_{n,N}\mathbf{G}_{n,N}$ is algebraic and unitarily invariant. □

The proof of Theorem 10.2 provides us with a recipe for computing the polynomials that encode the limiting eigenvalue distribution of $\mathbf{S}$ in the far more general situation where the observation vectors are modelled as samples of a multivariate Gaussian with spatio-temporal correlations. The limiting eigenvalue distribution of $\mathbf{S}$ depends on the limiting (algebraic) eigenvalue distributions of $\mathbf{A}$ and $\mathbf{B}$ and may be called using the `AtimesWishtimesB` function in the RMTool [22] package. See [23] for the relevant code.

10.5 Other applications

There is often a connection between well-known combinatorial numbers and random matrices. For example, the even moments of the Wigner matrix are the famous Catalan numbers. Similarly, if $\mathbf{W}_N(c)$ denotes the Wishart matrix with parameter c, other combinatorial correspondences can be easily established using the techniques developed. For instance, the limiting moments of $\mathbf{W}_N(1) - \mathbf{I}_N$ are the Riordan numbers, the large Schröder numbers correspond to the limiting moments of $2\mathbf{W}_N(0.5)$ while the small Schröder numbers are the limiting moments of $4\mathbf{W}_N(0.125)$. Combinatorial identities along the lines of those developed in [13] might result from these correspondences.

11. Some open problems

- If we are given a single realization of an $N \times N$ sized algebraic random matrix $\mathbf{A}_N$, is it possible to reliably infer the minimal bivariate polynomial L_{mz}, *i.e.*, with combined degree $\text{D}_\text{m} + \text{D}_\text{z}$ as small as possible, that encodes its limiting eigenvalue distribution?
- This is closely related to the following problem. Define the set of "admissible" real-valued coefficients c_{ij} for $0 \le i \le \text{D}_\text{m}$ and $0 \le j \le \text{D}_\text{z}$. Here admissibility implies that (a branch of) a solution $m(z)$ of the equation $L^A_{\text{mz}}(m, z) := \sum_{i=0}^{\text{D}_\text{m}} \sum_{j=0}^{\text{D}_\text{z}} c_{ij} u^i v^j = 0$ is globally the Stieltjes transform of a positive probability distribution.

Acknowledgements

Our work has greatly benefitted from interactions with colleagues. It is our pleasure to acknowledge their contribution. Jack Silverstein kindled the first author's involvement in the subject with a manuscript of unpublished work and several email discussions. Richard Stanley introduced us to D-finite series while

we learned about resultants from Pablo Parillo. We particularly wish to thank the anonymous referees for carefully proof-reading the original manuscript and challenging us to improve the organization of the paper.

We owe Roland Speicher thanks for numerous conversations, feedback, encouragement and for his patience in answering our queries. Our enthusiasm for free probability can be directly traced to his inspiring exposition in his lectures we have attended and the survey papers posted on his website. So much so, that when we sat down to write Section 9.1, we decided that it was pointless attempting to emulate the clarity of his writing. We are grateful to Roland for allowing us to borrow his writing in Section 9.1.

The authors were supported by a National Science Foundation grant DMS-0411962 and an Office of Naval Research Special Post-Doctoral Award N00014-07-1-0269.

Bibliography

[1] N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, Hafner Publishing Co., New York, New York, 1965. Translated by N. Kemmer.

[2] A. G. Akritas, *Sylvester's forgotten form of the resultant*, Fibonacci Quart., 31 (1993), pp. 325–332.

[3] G. Anderson and O. Zeitouni, *A law of large numbers for finite-range dependent random matrices*. `http://arxiv.org/abs/math/0609364`, September 2006. Preprint.

[4] Z. D. Bai and J. W. Silverstein, *On the empirical distribution of eigenvalues of a class of large dimensional random matices*, Journal of Multi. Analysis, 54 (1995), pp. 175–192.

[5] F. Bergeron and C. Reutenauer, *Combinatorial resolution of systems of differential equations. III. A special class of differentially algebraic series*, European J. Combin., 11 (1990), pp. 501–512.

[6] P. Biane, *Free probability for probabilists*, in Quantum probability communications, Vol. XI (Grenoble, 1998), QP-PQ, XI, World Sci. Publishing, River Edge, NJ, 2003, pp. 55–71.

[7] P. Billingsley, *Convergence of probability measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, second ed., 1999. A Wiley-Interscience Publication.

[8] B. Collins, *Product of random projections, Jacobi ensembles and universality problems arising from free probability*, Probab. Theory Related Fields, 133 (2005), pp. 315–344.

[9] P. Deift, T. Kriecherbauer, and K. T.-R. McLaughlin, *New results on the equilibrium measure for logarithmic potentials in the presence of an external field*, J. Approx. Theory, 95 (1998), pp. 388–475.

[10] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides, and X. Zhou, *Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory*, Comm. Pure Appl. Math., 52 (1999), pp. 1335–1425.

[11] P. A. Deift, *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, vol. 3 of Courant Lecture Notes in Mathematics, New York University Courant Institute of Mathematical Sciences, New York, 1999.

[12] W. B. Dozier and J. W. Silverstein, *On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices*. `http://www4.ncsu.edu/~jack/infnoise.pdf`, 2004.

[13] I. Dumitriu and E. Rassart, *Path counting and random matrix theory*, Electronic Journal of Combinatorics, 7 (2003). R-43.

[14] P. Flajolet and R. Sedgewick, *Analytic combinatorics: Functional equations, rational and algebraic functions*, Research Report 4103, INRIA, 2001. `http://algo.inria.fr/flajolet/Publications/FlSe01.pdf`.

[15] F. Hiai and D. Petz, *The semicircle law, free random variables and entropy*, vol. 77, American Mathematical Society, 2000.

[16] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1991.

[17] A. Kuijlaars and K. T.-R. McLaughlin, *Generic behavior of the density of states in random matrix theory and equilibrium problems in the presence of real analytic external fields*, Commun. Pure Appl. Math., 53 (2000), pp. 736–785.

[18] V. A. Marčenko and L. A. Pastur, *Distribution of eigenvalues in certain sets of random matrices*, Mat. Sb. (N.S.), 72 (114) (1967), pp. 507–536.

[19] B. D. McKay, *The expected eigenvalue distribution of a large regular graph*, Linear Algebra And Its Applications, 40 (1981), pp. 203–216.

[20] R. R. Nadakuditi, *Applied Stochastic Eigen-Analysis*, PhD thesis, Massachusetts Institute of Technology, February 2007. Department of Electrical Engineering and Computer Science.

[21] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, London Mathematical Society Lecture Note Series, New York: Cambridge University Press, 2006.

[22] N. R. Rao, *RMTool: A random matrix and free probability calculator in MATLAB*. `http://www.mit.edu/~raj/rmtool/`.

[23] N. R. Rao and A. Edelman, *The polynomial method for the eigenvectors of random matrices*. Preprint.

[24] B. Salvy and P. Zimmermann, *Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable*, ACM Trans. on Math. Software, 20 (1994), pp. 163–177.

[25] J. W. Silverstein, *The limiting eigenvalue distribution of a multivariate F matrix*, SIAM Journal on Math. Anal., 16 (1985), pp. 641–646.

[26] ——, *Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices*, J. of Multivariate Anal., 55(2) (1995), pp. 331–339.

[27] J. W. Silverstein and S.-I. Choi, *Analysis of the limiting spectral distribution of large-dimensional random matrices*, J. Multivariate Anal., 54 (1995), pp. 295–309.

[28] R. Speicher, *Free probability theory and non-crossing partitions*, Sém. Lothar. Combin., 39 (1997), pp. Art. B39c, 38 pp. (electronic).

[29] ——, *Free probability theory and random matrices*, in Asymptotic combinatorics with applications to mathematical physics (St. Petersburg, 2001), vol. 1815 of Lecture Notes in Math., Springer, Berlin, 2003, pp. 53–73.

[30] R. P. Stanley, *Enumerative combinatorics. Vol. 2*, vol. 62 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[31] B. Sturmfels, *Introduction to resultants*, in Applications of computational algebraic geometry (San Diego, CA, 1997), vol. 53 of Proc. Sympos. Appl. Math., Amer. Math. Soc., Providence, RI, 1998, pp. 25–39.

[32] A. M. Tulino and S. Verdú, *Random matrices and wireless communications*, Foundations and Trends in Communications and Information Theory, 1 (2004).

[33] D. Voiculescu, *Symmetries of some reduced free product C^*-algebras*, in Operator algebras and their connections with topology and ergodic theory (Buşteni, 1983), vol. 1132 of Lecture Notes in Math., Springer, Berlin, 1985, pp. 556–588.

[34] ——, *Addition of certain noncommuting random variables*, J. Funct. Anal., 66 (1986), pp. 323–346.

[35] ——, *Multiplication of certain noncommuting random variables*, J. Operator Theory, 18 (1987), pp. 223–235.

[36] ——, *Limit laws for random matrices and free products*, Invent. Math., 104 (1991), pp. 201–220.

[37] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free random variables*, vol. 1 of CRM Monograph Series, American Mathematical Society, Providence, RI, 1992.

[38] E. P. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Annals of Math., 62 (1955), pp. 548–564.