Singular value correlation functions for products of Wishart random matrices

Gernot Akemann$^{1,2}$, Mario Kieburg$^{1,3}$ and Lu Wei$^{4,5}$

$^1$ Department of Physics, Bielefeld University, Postfach 100131, D-33501 Bielefeld, Germany
$^2$ Laboratori Nazionali di Frascati, Via E Fermi 40, I-00044 Frascati (Rome), Italy
$^3$ Department of Physics and Astronomy, SUNY, Stony Brook, New York 11794, NY, USA
$^4$ Department of Mathematics and Statistics, University of Helsinki, Helsinki, PO Box 68, FI-00014, Finland
$^5$ School of Electrical Engineering, Aalto University, PO Box 13000, FI-00076 Aalto, Finland

E-mail: akemann@physik.uni-bielefeld.de, mkieburg@physik.uni-bielefeld.de and lu.wei@helsinki.fi

Received 3 April 2013, in final form 29 May 2013
Published 21 June 2013
Online at stacks.iop.org/JPhysA/46/275205

Abstract

We consider the product of $M$ quadratic random matrices with complex elements and no further symmetry, where all matrix elements of each factor have a Gaussian distribution. This generalizes the classical Wishart–Laguerre Gaussian unitary ensemble with $M=1$. In this paper, we first compute the joint probability distribution for the singular values of the product matrix when the matrix size $N$ and the number $M$ are fixed but arbitrary. This leads to a determinantal point process which can be realized in two different ways. First, it can be written as a one-matrix singular value model with a non-standard Jacobian, or second, for $M \geq 2$, as a two-matrix singular value model with a set of auxiliary singular values and a weight proportional to the Meijer $G$-function. For both formulations, we determine all singular value correlation functions in terms of the kernels of biorthogonal polynomials which we explicitly construct. They are given in terms of the hypergeometric and Meijer $G$-functions, generalizing the Laguerre polynomials for $M=1$. Our investigation was motivated from applications in telecommunication of multi-layered scattering multiple-input and multiple-output channels. We present the ergodic mutual information for finite-$N$ for such a channel model with $M-1$ layers of scatterers as an example.

PACS numbers: 02.10.Yn, 02.30.Cj, 02.30.Ik, 02.50.Sk, 84.40.Ba, 84.40.Ua
Mathematics Subject Classification: 15B52, 33C20, 33C45, 94Axx

(Some figures may appear in colour only in the online journal)
1. Introduction

Random matrix theory (RMT) remains a very active field of research even after many decades of work. Originally conceived in the area of mathematical statistics and nuclear physics, today’s applications of RMT extend far beyond the mathematical and physical sciences, and we refer to [1] for a recent overview.

One of the topics in RMT that has caught recent attention is that of the products of random matrices. One of its original motivations in statistical physics was the description of chaotic and disordered systems [2], and among more recent applications are combinatorics [3] and telecommunications [4], which has also been one of our motivations. In particular, we consider MIMO (multiple-input and multiple-output) communication networks, where multi-antenna transceivers are utilized to improve the system capacity in a rich scattering environment.

Products of matrices loose much of the symmetry of the individual matrices and are generically complex. For simplicity, we will consider the individual matrices to be complex, too, with independent Gaussian distributions. The spectral properties of matrix ensembles carry important information, which is encoded in the eigenvalue decomposition as well as in the singular value decomposition. We will focus on the latter here.

A striking property of RMT is its universality, that is, the independence of the underlying distribution of the individual matrix elements. It is usually manifest in the limit of a large matrix size. However, if we study the local, microscopic behaviour of the spectrum on the scale of the mean level spacing between singular values, then it is often vital to have detailed knowledge of the joint distribution of singular values (or eigenvalues) at hand for a finite matrix size. In particular, to derive a determinantal or Pfaffian structure of the correlation functions of the random matrix ensemble has proven very useful for universality studies. Some of the most powerful proofs of universality start from the knowledge of orthogonal polynomials of these determinantal or Pfaffian point processes, in order to perform the asymptotic analysis. We refer the reader to [5] for a standard reference.

The aim of this work is to provide such a starting point, by deriving the joint probability density function (jpdf) of the singular values of the product matrix at finite matrix size $N$, for a finite product of $M$ matrices. We consider the simplest case of $M$ quadratic $N \times N$ matrices of Wishart type, independently and identically distributed by Gaussians with unit variance. In telecommunications, this is also the setting often encountered with both $N$ and $M$ finite. The singular value distribution of products of complex Wishart matrices is then the setup for the calculation of several information-theoretic quantities.

In previous works, the spectral density of singular values as well as the moments of such product matrices were derived in the macroscopic large-$N$ limit. Probabilistic methods such as free random variables [6–8], field-theoretic methods such as planar diagrams [9] and inverse Mellin transforms [3] have been used. The limit of infinitely many matrices in a product was studied in other works, either for finite-size [10–13] or for infinitely large matrices [14, 15], where the problem was mapped to a differential equation.

Very recently the jpdf and its correlation functions of the complex eigenvalues of the matrix ensemble we are considering have been derived for finite $N$ and $M$ [16–18]. Also, here the macroscopic large-$N$ density in the complex plane was known previously, see [19, 20, 9] for a collection of works. However, the corresponding question about the singular value correlation functions for finite $N$ and $M$ remained open and is addressed in our work.

This paper is organized as follows. In section 2, we determine the jpdf of singular values, in two different ways. Section 3 is devoted to the computation of the correlation functions, by first determining the biorthogonal polynomials associated with this problem in subsection 3.1, and then the kernel(s) leading to all $k$-point correlators in subsection 3.2. The spectral density
itself is discussed in more detail in subsection 3.3, where we compute all its moments and identify the correct rescaling for the macroscopic large-$N$ limit of our results. In particular, we compare our results to the known large-$N$ asymptotic spectral density e.g. from [9]. Section 4 illustrates how our results can help in applications in telecommunications, by computing the ergodic mutual information and comparing it to numerical simulations. After presenting our concluding remarks and open questions in section 5 we collect some technical tools in the two appendices.

2. Joint probability distribution of singular values

We are interested in the singular values of the product $P_M$ of $M$ independent matrices with complex matrix elements of size $N \times N$, $X_j \in \text{Gl}(N, \mathbb{C})$ for all $j = 1, \ldots, M$:

$$P_M \equiv X_M X_{M-1} \ldots X_1,$$

Furthermore, the matrices $X_j$ are distributed by identical, independent Gaussians,

$$P(X_j) = \exp \left[-\text{Tr}X_j^\dagger X_j\right].$$

The partition function $Z_N^{(M)}$ is then defined as

$$Z_N^{(M)} = C \int \prod_{j=1}^M \text{d}[X_j] P(X_j),$$

where $\text{d}[X_j] = \prod_{\alpha, \beta=1}^N d(X_j)_{\alpha \beta} d(X_j^\dagger)_{\alpha \beta}$ denotes the flat measure over all independent matrix elements. In this section, we compute the jpdf of the singular values of $P_M$. For $M = 1$, this is the well-known Wishart–Laguerre (also called chiral Gaussian) unitary ensemble, and throughout most of the following we will thus restrict ourselves to $M > 1$. We will not keep track of the normalization constant $C$ here, and will only specify it later, once we have changed to singular values.

In the first step, we perform the following successive change of variables from $X_j$ to $Y_j$:

$$Y_1 \equiv X_1, \quad \text{and} \quad Y_j \equiv X_j Y_{j-1} \quad \text{for} \quad j = 2, \ldots, M,$$

e.g. $Y_2 = X_2 X_1$, $Y_3 = X_3 X_2 X_1$ etc. In the new variables, $Y_M = P_M$ is the product matrix we are aiming for. While the first one, $Y_1 = X_1$, is a trivial relabelling, each subsequent change of variables carries a non-trivial Jacobian given by $1/\det[Y_{j-1}^\dagger Y_{j-1}]^N$ for $j = 2, \ldots, M$. This can be seen as follows. Due to $\text{d}(Y_j)_{\alpha \beta} = \sum_{\gamma=1}^N d(X_j)_{\alpha \gamma} (Y_{j-1})_{\gamma \beta}$, each column vector of the matrix $Y_j$ acquires a factor $1/\det[Y_{j-1}]$ from the change of variables, and likewise its complex conjugate. Taking into account all $N$ column vectors and their complex conjugates, we find the given Jacobian for each $j = 2, \ldots, M$, and we arrive at

$$Z_N^{(M)} = C \int \prod_{j=1}^M \text{d}[Y_j] \exp \left[-\text{Tr}Y_j^\dagger Y_j\right] \prod_{j=2}^M \frac{1}{\det \left[Y_{j-1}^\dagger Y_{j-1}\right]^N} \exp \left[-\text{Tr}Y_j^\dagger Y_j (Y_{j-1}^\dagger Y_{j-1})^{-1}\right].$$

(2.5)

In writing this, we assume that the matrices $X_j$ and thus their products $Y_j$ are invertible\(^6\).

In the second step, we decompose each matrix $Y_j = V_j \Lambda_j U_j$, $j = 1, \ldots, M$ into its angles and singular values, where $\Lambda_j = \text{diag}(\lambda_1^{(j)}, \ldots, \lambda_N^{(j)})$ contains the positive singular values $\lambda_a^{(j)} \in \mathbb{R}_+$, $a = 1, \ldots, N$, and $U_j \in \text{U}(N)$ and $V_j \in \text{U}(N)/\text{U}(1)^N$ are unitary. The Jacobian

\(^6\) Our restriction from general complex $N \times N$ matrices to $\text{Gl}(N, \mathbb{C})$ removes only a set of measure zero.
resulting from the singular value decomposition of each matrix \( Y_j \) is well known and is given in terms of the Vandermonde determinant,

\[
\Delta_N(\Lambda_j) \equiv \prod_{N > a > b > 1} (\lambda_a^{(i)} - \lambda_b^{(j)}) = \det_{1 \leq a, b \leq N} [(\lambda_a^{(i)})^{b-1}].
\]

(2.6)

Since we encounter the matrices \( Y_j \) in the combination \( Y_j^* Y_j \) only, the unitary matrices \( U_j \) completely drop out which leads to

\[
\begin{align*}
\mathcal{Z}_N^{(M)} &= C' \prod_{i=1}^M \left\{ \int \mathcal{D}V_i \mathcal{D}U_i \prod_{a=1}^d \mathcal{D}\lambda_a^{(i)} \right\} \exp \left[ -\text{Tr} \Lambda_1^2 \right] \Delta_N(\Lambda_1^2)
\times \prod_{j=2}^M \left\{ \int \mathcal{D}V_i \mathcal{D}U_i \prod_{a=1}^d \mathcal{D}\lambda_a^{(i)} \right\} \exp \left[ -\text{Tr} \Lambda_{j-1}^2 \right] \Delta_N(\Lambda_{j-1}^2)
\end{align*}
\]

(2.7)

In the second step, we employed the invariance of the Haar measure \( \mathcal{D}U_j \) under \( U_j \to U_j U_{j-1} \), which leads to the decoupling of the integrations over \( \mathcal{D}U_1 \) and all the \( \mathcal{D}V_j \) from the rest of the integrals. The remaining unitary integrations in the last line of equation (2.7) can be performed using the so-called Harish–Chandra–Itzykson–Zuber (HCIZ) integral \([21, 22]\)

\[
\int \mathcal{D}U_j \exp \left[ -\text{Tr} (U_j^* \Lambda_j^2 U_j \Lambda_j^{-2}) \right] = \frac{1}{\Delta_N(\Lambda_j^2) \Delta_N(\Lambda_{j-1}^2)} \det_{1 \leq a, b \leq N} \left[ \exp \left( -\frac{(\lambda_a^{(i)})^2}{(\lambda_b^{(i-1)})^2} \right) \right].
\]

(2.8)

The Vandermonde determinant of inverse powers is proportional to the ordinary one with positive powers due to the following identity:

\[
\Delta_N(\Lambda_j^{-2}) = \det_{1 \leq a, b \leq N} \frac{1}{(\lambda_a^{(i)})^{2b-2}} = (-1)^{W(-1)} \frac{\Delta_N(\Lambda_j^2)}{\Delta_N(\Lambda_{j-1}^2)}.
\]

(2.9)

This leads to many cancellations in equation (2.7), in particular, of almost all Vandermonde determinants:

\[
\begin{align*}
\mathcal{Z}_N^{(M)} &= C'' \prod_{i=1}^N \left\{ \int \mathcal{D}\lambda_a^{(i)} \prod_{j=1}^{M-1} \int \mathcal{D}\lambda_a^{(j)} \right\} \exp \left[ -\sum_{b=1}^N \left( \lambda_b^{(1)} \right)^2 \right] \Delta_N(\Lambda_1^2) \Delta_N(\Lambda_M^2)
\times \prod_{j=2}^M \left\{ \int \mathcal{D}\lambda_a^{(i)} \prod_{a=1}^d \int \mathcal{D}\lambda_a^{(i)} \right\} \exp \left( -\frac{(\lambda_a^{(i)})^2}{(\lambda_a^{(i-1)})^2} \right).
\end{align*}
\]

(2.10)

We expand the determinant comprising \( \lambda_a^{(i)} \) and \( \lambda_b^{(j)} \) in \( N! \) terms. Each of these terms yields the same contribution since the permutation involved in the definition of the determinant can be absorbed in the determinant comprising \( \lambda_a^{(2)} \) and \( \lambda_b^{(3)} \) due to the antisymmetry of determinants, and a relabelling of the integration variables. Next, we expand the determinant comprising \( \lambda_a^{(2)} \) and \( \lambda_b^{(3)} \) whose permutations can be absorbed in the determinant comprising \( \lambda_a^{(3)} \) and \( \lambda_b^{(4)} \), and so on. This interplay of expansion and absorption of the permutations of the determinants can be continued until all determinants stemming from the HCIZ integral are replaced by

\[\text{The Haar measure is normalized such that there is no further proportionality constant on the right-hand side.}\]
their diagonal part. Note that we do not require any symmetrization in the variables \( \lambda_a^{(M)} \) here. Hence, our argument applies to the correlation functions of the \( \lambda_a^{(M)} \) in section 3, too, where we integrate the jpdf only over a subset of these singular values.

Almost all remaining multiple integrals can be simplified as follows:

\[
\prod_{a=1}^{n} \left\{ \int_0^\infty \frac{d\lambda_a}{\lambda_a^{(1)}} \exp \left[ -\left( \lambda_a^{(1)} \right)^2 \right] \left( \prod_{j=2}^{N-1} \int_0^\infty \frac{d\lambda_j}{\lambda_j^{(j)}} \exp \left[ -\left( \lambda_j^{(j)} \right)^2 \right] \right) \right\} \times \det_{1 \leq c,d \leq N} \left[ \left( \lambda_c^{(1)} \right)^{2M-2} \right] = \det_{1 \leq c,d \leq N} \left[ \int_0^\infty \frac{d\lambda_c}{\lambda_c^{(1)}} \exp \left[ -\left( \lambda_c^{(1)} \right)^2 \right] \left( \prod_{j=2}^{N-1} \int_0^\infty \frac{d\lambda_j}{\lambda_j^{(j)}} \exp \left[ -\left( \lambda_j^{(j)} \right)^2 \right] \right) \times \exp \left[ -\left( \lambda_c^{(M)} \right)^2 \right] \right] = \det_{1 \leq c,d \leq N} \left[ \frac{1}{2^{M-1} N!} \prod_{a=0}^{M-1} \prod_{b=0}^{d-1} \left( \lambda_a^{(M)} \right)^{2M-2} \right].
\]  

(11.1)

Note that we have left out the integration over the variables \( \lambda_a^{(M)} \). In the second line of equation (11.1), we have pulled all the integrations into the corresponding rows of the determinants, and in the third line we have used the integral identity (A.10) in the squared singular values that are derived in appendix A. The special function appearing here is the so-called Meijer \( G \)-function, see equation (A.1) for its definition [23]. The number of zeros in the bottom line of the Meijer \( G \)-function is \( M - 1 \). Our first main result is thus the following singular value representation of the partition function, after changing to squared singular values \( s_a \equiv (\lambda_a^{(M)})^2 \) with \( ds_a = 2\lambda_a^{(M)}d\lambda_a^{(M)} \), \( a = 1, \ldots, N \), in equation (10.2):

\[
Z_N^{(M)} = C_N^{(M)} \int_0^\infty \prod_{a=1}^{N} ds_a \Delta_N(s) \det_{1 \leq c,d \leq N} \left[ G_{0,M}^{M,0}(0, \ldots, 0, -d-1 | s) \right] = \int_0^\infty \prod_{a=1}^{N} ds_a \mathcal{P}_{jpdf}(s),
\] 

(12.1)

\[
\mathcal{P}_{jpdf}(s) \equiv C_N^{(M)} \Delta_N(s) \det_{1 \leq c,d \leq N} \left[ G_{0,M}^{M,0}(0, \ldots, 0, -d-1 | s) \right],
\] 

(13.1)

where \( \mathcal{P}_{jpdf} \) is the jpdf. We will show later that it corresponds to a determinantal point process. The constant in front of equation (13.1),

\[
\left( C_N^{(M)} \right)^{-1} = N! \prod_{a=1}^{N} \Gamma(a)^{M+1},
\] 

(14.1)

has been chosen such that the partition function is normalized to unity. This can be seen as follows. Applying the Andrieuf integral identity,

\[
\int \prod_{a=1}^{N} ds_a \det_{1 \leq c,d \leq N} [\phi_c(s_a)] \det_{1 \leq c,d \leq N} [\psi_c(s_d)] = N! \det_{1 \leq c,d \leq N} \left[ \int ds \phi_c(s) \psi_c(s) \right],
\] 

(15.1)

which applies to any two sets of functions \( \phi_c \) and \( \psi_c \) such that all integrals exist, we obtain for the partition function

\[
Z_N^{(M)} = C_N^{(M)} N! \det_{1 \leq c,d \leq N} \left[ \int_0^\infty ds \psi_c(s)^{-1} G_{0,M}^{M,0}(0, \ldots, 0, -d-1 | s) \right] = C_N^{(M)} N! \det_{1 \leq c,d \leq N} [\Gamma(c)^{M-1} \Gamma(c + d - 1)] = C_N^{(M)} N! \prod_{c=1}^{N} \Gamma(c)^{M-1} \prod_{b=1}^{N} \Gamma(b)^2 = 1.
\] 

(16.1)
Here, we have used another integral identity from the appendix, equation (A.7), and pulled out factors of Gamma-functions from the rows of the determinant. The remaining determinant is nothing but the normalization of the Wishart–Laguerre ensemble at $M = 1$ which is well known [24].

We refer to the result (2.12) as a ‘one-matrix’ singular value model because it is of the form that would result from the singular value decomposition of a single random matrix, however with a non-standard Jacobian $\neq \Delta_N(s)^2$. As an easy check we can see that due to

$$G_{0,1}^{1,0}(a_{-1} | s_k) = e^{d-1} \exp[-s_k],$$

(2.17)

the expression in equation (2.12) reduces to the standard Wishart–Laguerre ensemble when setting $M = 1$, and taking the exponentials out of the second determinant.

In principle, we could now try to compute all singular value $k$-point correlation functions defined as

$$R_k^{(M)}(s_1, \ldots, s_k) \equiv \frac{N!}{(N-k)!} \int_0^\infty \prod_{a=k+1}^N ds_a \mathcal{P}_\text{pdf}(s).$$

(2.18)

For example for $k = 1$, this gives the spectral density which is normalized to $N$ in our convention following [24],

$$N = \int ds_1 R_{1,1}^{(M)}(s_1),$$

(2.19)

whereas for $k = N$ we have the $jpdf$ itself, $R_N^{(M)}(s_1, \ldots, s_N) = N! \mathcal{P}_\text{pdf}(s)$. However, due to the matrix inside the second determinant in equation (2.12) being labelled by the indices of the Meijer $G$-function, the computation of the $R_k^{(M)}$ is a highly nontrivial task. We postpone this computation to section 3 using a second ‘two-matrix’ formulation that is introduced in the following subsection.

2.1. An alternative jpdf with auxiliary variables

We introduce a formalism that is more convenient to handle when computing correlation functions. Let us step back by considering equation (2.10), taking for simplicity $M = 2$:

$$Z_N^{(M=2)} = C \int_0^\infty \prod_{a=1}^N d\lambda_a^{(2)} \prod_{b=1}^N d\lambda_b^{(1)} \exp \left[ -\left( \lambda_d^{(1)} \right)^2 \right] \Delta_N(\lambda_1^2) \Delta_N(\lambda_2^2)$$

$$\times \det_{1 \leq c \neq d \leq N} \left[ \exp \left( -\frac{\lambda_d^{(1)}}{\lambda_c^{(1)}} \right)^2 \right].$$

(2.20)

This is precisely in the form of a ‘two-matrix’ singular value model (2mm) that results from the singular value decomposition of a two-matrix model, see e.g. in [25]. The advantage is that we now have the standard form of the Jacobian given by one Vandermonde per set of variables and an additional determinant of a function that couples the two sets of variables. Such a setting can be tackled using the known biorthogonal polynomial technique, as reviewed in [26]. In order to apply this technique, we have to bring equation (2.10) into such a form, but for arbitrary values of $M \geq 2$. This can be readily achieved by taking the same steps as from equation (2.10) to equation (2.11), but this time excluding both sets of integrations over the variables $\lambda_d^{(M)}$ and $\lambda_d^{(1)}$.

The symmetry argument goes along the same lines replacing determinants by their diagonal parts, see the discussion after equation (2.10), but this time we keep the determinant containing the exponential with $\lambda_d^{(1)}$ and $\lambda_d^{(M)}$. The integrations are
where we have used again the identity (A.10) from appendix A, see also [16] for a related recurrence relation. We therefore arrive at our second main result, the following 2mm representation for the jpdf, after changing again to squared singular values \( s_a \equiv (\lambda_a^{(M)})^2 \), \( t_a \equiv (\lambda_a^{(1)})^2 \), \( a = 1, \ldots, N \):

\[
Z^{(M)}_N = \frac{C_N^{(M)}}{N!} \int_0^\infty \prod_{a=1}^N ds_t \prod_{a=1}^N dt_a \ e^{-s_t} \Delta_N(s) \Delta_N(t) \det_{1 \leq c,d \leq N} \left[ G_{0,M-1}^{M-1,0} \left( \begin{array}{c|c} s_c & \hline 0 & I_d \end{array} \right) \right] \frac{1}{2M-2} c_{0,M-1} \left( \begin{array}{c|c} \lambda_c^{(M)} & \hline \lambda_d^{(1)} & I_d \end{array} \right) \right],
\]

(22.22)

It is normalized to unity as we will check below.

The crucial advantage in comparison to \( P_{pdf} \) is the matrix inside the determinant which has indices that label the integration variables, and not the indices of the Meijer G-function as in equation (2.12). This 2mm describes the correlations among the singular values \( \lambda_a^{(1)} \) of a single matrix \( X_1 \) and the singular values \( \lambda_d^{(M)} \) of the entire product matrix \( P_M \), to be computed in section 3.

As a check for \( M = 2 \) we get back to equation (2.20), using

\[
G_{0,1}^{1,0} \left( \begin{array}{c|c} (\lambda_c^{(2)})^2 & \hline (\lambda_d^{(1)})^2 & I_d \end{array} \right) = \exp \left[ -\frac{(\lambda_c^{(2)})^2}{(\lambda_d^{(1)})^2} \right].
\]

(23.23)

Confirming the normalization in equation (2.22) is at the same time a check that this representation can be mapped back to equation (2.12) in a different way. Applying once again the Andréief formula (2.15) to equation (2.22), but this time only to the \( t \)-integration, we obtain the following:

\[
\int_0^\infty \prod_{a=1}^N dt_a \ e^{-s_t} \Delta_N(t) \det_{1 \leq c,d \leq N} \left[ G_{0,M-1}^{M-1,0} \left( \begin{array}{c|c} s_c & \hline 0 & I_d \end{array} \right) \right]\]

\[
= N! \det_{1 \leq c,d \leq N} \left[ \int_0^\infty dt_e^{d-2} e^{-t_e} G_{0,M-1}^{M-1,0} \left( \begin{array}{c|c} s_c & \hline 0 & I_d \end{array} \right) \right] \]

\[
= N! \det_{1 \leq c,d \leq N} \left[ \int_0^\infty dt_e^{d-2} e^{-t_e} G_{0,M-1}^{M-1,0} \left( \begin{array}{c} s_c \hline 0 \end{array} \right) \right] \]

(24.24)

upon using the identity (A.9) with \( m = M - 1 \) from appendix A. This brings us back to the ‘one-matrix’ model representation (2.12) in terms of a single set of singular values, with the proper normalization.
3. Singular value correlation functions

We are now prepared to compute arbitrary $k$-point correlation functions of the singular values. Rather than using definition (2.18), we will consider the more general correlation functions of the jpdf in the 2mm representation (2.22):

$$R^{(M)}_{k,l}(s_1, \ldots, s_k; t_1, \ldots, t_l) \equiv \frac{N!^2}{(N-k)!(N-l)!} \int_0^\infty \int_0^\infty N \prod_{a=k+1}^N ds_a \prod_{b=l+1}^N dt_b \mathcal{P}^{2mm}_{jpdf}(s, t).$$  

(3.1)

The $k$-point functions of the singular values of the product matrix $P_M$ in equation (2.18) can be obtained by integrating out all auxiliary variables or by setting $l = 0$: $R^{(M)}_{k,0}(s_1, \ldots, s_k; -) = R^{(M)}_k(s_1, \ldots, s_k)$.

Let us introduce the following set of biorthogonal polynomials (bOP) in monic normalization, $p_j(x) = x^j + \cdots$ and $q_j(x) = x^j + \cdots$,

$$\int_0^\infty ds \int dr w^{(M)}(s,t)p_j^{(M)}(s)q_j^{(M)}(t) = \delta_j R^{(M)}_i,$$

(3.2)

with squared norms $h_j^{(M)}$ and the weight function defined as

$$w^{(M)}(s,t) := \det M^{-1} e^{-M^{-1,0} \begin{pmatrix} s & t \end{pmatrix} \begin{pmatrix} s \end{pmatrix}}$$

(3.3)

for $M > 1$. These polynomials are guaranteed to exist following the general theory of bOP that was recently developed further [27], see also [26] for a recent review. We will explicitly construct the bOP for general $M > 1$. The general $(k, l)$-point correlation functions (3.1) are given in terms of four kernels that are constructed from the kernel of bOP [28–30]:

$$R^{(M)}_{k,l}(s_1, \ldots, s_k; t_1, \ldots, t_l) = \det \begin{bmatrix} H_0(s_a, s_b) & H_0(s_a, t_j) \\ H_0(s_b, t_j) & H_0(t_a, t_j) \\ & H_1(t_j, s_b) & H_1(t_j, t_j) \\ & & H_1(t_j, s_b) & H_1(t_j, t_j) \end{bmatrix}.$$

(3.4)

The kernel of bOP is defined as

$$K_N(s,t) \equiv \sum_{j=0}^{N-1} \frac{P_j^{(M)}(s)q_j^{(M)}(t)}{h_j^{(M)}}.$$

(3.5)

All four kernels $H_{ab}$ are based on this relation,

$$H_{01}(s_a, s_b) \equiv \int_0^\infty dr K_N(s_a, t) r^{-1} e^{-M^{-1,0} \begin{pmatrix} s_b & t \end{pmatrix} \begin{pmatrix} s_b \end{pmatrix}}.$$

(3.6)

$$H_{00}(s_a, t_j) \equiv t_j^{-1} e^{-t_j} K(s_a, t_j),$$

(3.7)

$$H_{11}(t_j, s_b) \equiv \int_0^\infty ds \int_0^\infty dr K_N(s, t) r^{-1} e^{-M^{-1,0} \begin{pmatrix} s & t \end{pmatrix} \begin{pmatrix} s \end{pmatrix}} e^{-M^{-1,0} \begin{pmatrix} s & t \end{pmatrix} \begin{pmatrix} s \end{pmatrix}}.$$

(3.8)

$$H_{10}(t_j, t_j) \equiv t_j^{-1} e^{-t_j} \int_0^\infty ds K_N(s, t) e^{-M^{-1,0} \begin{pmatrix} s & t \end{pmatrix} \begin{pmatrix} s \end{pmatrix}}.$$

(3.9)

Note that equation (3.4) implies that both the one- and two-matrix model jpdf represent determinantal point processes, i.e. for the former, equation (2.13) becomes

$$\mathcal{P}_{jpdf}(s) = \frac{1}{N!} \prod_{1 \leq a,b \leq N} [H_{01}(s_a, s_b)].$$

(3.10)
3.1. The biorthogonal polynomials

In order to compute the bOP, let us first determine the bimoment matrix

\[
I_{ij} \equiv \int_0^\infty \int_0^\infty ds \, t^{i+j} e^{-s} w^{(M)}(s, t) \frac{s^j t^i}{i!}
\]

\[
= \int_0^\infty ds \, t^i e^{-s} \int_0^\infty \frac{ds}{s} G_{0,M-1}^{M-1,0} \left( \frac{s}{t} \right)
\]

\[
= \int_0^\infty ds \, t^{i+j} e^{-s} (i!)^{M-1}
\]

\[
= (i+j)! (i!)^{M-1},
\]

which follows again from an identity for Meijer G-functions, see equation (A.7) from appendix A. The bOP as well as their norms are determined by this bimoment matrix (see e.g. [27]):

\[
p_n^{(M)}(s) = \frac{1}{D_n^{(M)}} \det \begin{bmatrix}
I_{00} & I_{10} & \cdots & I_{n0} \\
I_{01} & I_{11} & \cdots & I_{n1} \\
\vdots & \vdots & \ddots & \vdots \\
I_{0n-1} & I_{1n-1} & \cdots & I_{nn-1}
\end{bmatrix},
\]

\[
q_n^{(M)}(t) = \frac{1}{D_n^{(M)}} \det \begin{bmatrix}
I_{00} & I_{10} & \cdots & I_{n0} & 1 \\
I_{01} & I_{11} & \cdots & I_{n1} & t \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
I_{0n-1} & I_{1n-1} & \cdots & I_{nn-1} & t^n
\end{bmatrix},
\]

where

\[
D_n^{(M)} = \prod_{0 \leq i, j \leq n-1} [I_{ij}!] = \prod_{i=0}^{n-1} (i!)^{M-1} \det_{0 \leq i, j \leq n-1} [(i+j)!],
\]

\[
h_n^{(M)} = D_n^{(M)} / D_{n+1}^{(M)}.
\]

In order to have more explicit expressions, it is instructive to compare these equations with the standard Laguerre polynomials \(L_n(x)\). We need them in monic normalization denoted by \(\tilde{L}_n(x)\):

\[
\tilde{L}_n(x) \equiv (-1)^n n! L_n(x) \equiv \sum_{k=0}^{n} \frac{(-1)^{n-k}(n!)^2}{(n-k)!} x^k,
\]

with squared norms

\[
\int_0^\infty dx \, e^{-x} \tilde{L}_n(x) \tilde{L}_m(x) = \delta_{nm} (n!)^2 \equiv \delta_{nm} h_n^{(M=1)},
\]

and a symmetric bimoment matrix

\[
I_{ij} |_{M=1} \equiv \int_0^\infty dt \, t^{i+j} e^{-t} = (i+j)!, \quad \tilde{L}_n(x) = \frac{1}{D_n^{(M=1)}} \det \begin{bmatrix}
I_{00} |_{M=1} & \cdots & I_{0n} |_{M=1} \\
\vdots & \ddots & \vdots \\
I_{0n-1} |_{M=1} & \cdots & I_{nn-1} |_{M=1}
\end{bmatrix}.
\]
The former equals equation (3.11) with \( M = 1 \). As can be seen, the monic Laguerre polynomials also have a determinant representation from the Gram–Schmidt procedure, which is exactly that in equation (3.12) at \( M = 1 \) (or equation (3.13) as they become equal then).

From the comparison of equations (3.12) and (3.13) to equation (3.18), we can read off the following. For the \( q_n(t) \), we take out the common factors \((\frac{\vec{t}}{n!})^{M-1}\) from the first \( n \) columns of the determinant, \( i = 0, 1, \ldots, n - 1 \), with the remaining determinant being identical to that of the monic Laguerre polynomials. The determinant of the bimoment matrix (3.14) is already written to be proportional to the corresponding one of the Laguerre ensemble. We thus have

\[
g_n^{(M)}(t) = \frac{\prod_{i=0}^{n-1}(\vec{t})^{M-1}}{\prod_{i=0}^{n-1}(\vec{t})^{M-1}} \tilde{L}_n(t) = \tilde{L}_n(t),
\]

(3.19)

for all values of \( M \). Likewise we can read off the squared norms by comparing them to the Laguerre case (3.17):

\[
h_n^{(M)} = \frac{\prod_{i=0}^{n}(\vec{t})^{M-1}}{\prod_{i=0}^{n}(\vec{t})^{M-1}} h_n^{(M-1)} = (n!)^{M+1}.
\]

(3.20)

This equation is formally redundant for \( M = 1 \).

For the polynomials \( p_n(s) \), the case is slightly more complicated. Also, for these polynomials we can take out common factors from all \( n + 1 \) columns; however, this will modify the arguments in the last row of the determinant in the numerator: \( s' \rightarrow s'/(\frac{\vec{t}}{n!})^{M-1} \).

Expanding with respect to the last row we obtain a polynomial with the same coefficients as the monic Laguerre polynomials, but now with \( s'/(\frac{\vec{t}}{n!})^{M-1} \) instead of the monomial \( s' \) alone, resulting in

\[
p_n^{(M)}(s) = \frac{\prod_{i=0}^{n}(\vec{t})^{M-1}}{\prod_{i=0}^{n}(\vec{t})^{M-1}} \sum_{k=0}^{n} (-1)^{n-k} \frac{n!}{(n-k)!} \left( \frac{k!}{\vec{k}^{M-1}} \right)^2 \frac{s^k}{(k!)^{M-1}}
\]

\[
= \sum_{k=0}^{n} (-1)^{n-k} \frac{n!}{(n-k)!} \left( \frac{k!}{\vec{k}^{M-1}} \right)^2 \frac{s^k}{(k!)^{M-1}}
\]

(3.21)

This function can be interpreted as a generalization of the monic Laguerre polynomials reobtained when setting \( M = 1 \), see equation (3.16). Moreover, we could express it in terms of the generalized hypergeometric function \(_1F_M\) which has \( M \) arguments equal to 1 in the second set of its indices.

The kernel (3.5) is now completely determined. We proceed by computing the various kernels (3.6), (3.8) and (3.9) by integrating the kernel of bOP.

### 3.2. The kernels and all correlation functions

We start with the kernel \( H_{01} \) that is relevant for the correlation functions of all singular values \( s_a = (\lambda_{,a}^{(M)})^2 \). We have

\[
H_{01}(s_a, s_b) = \sum_{j=0}^{N-1} \frac{1}{H_j^{(M)}} \overline{p_j^{(M)}}(s_a) \int_0^\infty dt \bar{\tilde{L}}_j(t') t'^{-1} e^{-\frac{t}{\bar{T}}} G_{0,m-1}^{M} \left( \frac{\bar{T}}{T} \right)
\]

\[
\sum_{j=0}^{N-1} \frac{1}{H_j^{(M)}} \overline{p_j^{(M)}}(s_a) \overline{q_j^{(M)}}(s_b).
\]

(3.22)
where we introduce the following integral transform:
\[
\chi_j^{(M)}(s_b) = \int_0^\infty \mathrm{d}t' \tilde{L}_j(t') t'^{-1} e^{-t'} G_{0,M-1}^{1,0} \left( \frac{s_b}{t'} \right)
\]
\[
= \sum_{j=0}^J (-1)^{J-i} \left( \frac{j!}{i!} \right)^2 \int_0^\infty \mathrm{d}t' (t')^{j-1} e^{-t'} G_{0,M-1}^{1,0} \left( \frac{s_b}{t'} \right)
\]
\[
= \sum_{j=0}^J (-1)^{J-i} \left( \frac{j!}{i!} \right)^2 G_{0,M}^{1,0} \left( \frac{s_b}{t' \cdots t_0} \right),
\tag{3.23}
\]
upon using the identity (A.9) for \(d - 1 = i\) and \(m = M - 1\). A more compact expression of \(\chi_j\) can be found by combining the Rodrigues formula
\[
\tilde{L}_j(t') = e^{-t'} \left( -\frac{\mathrm{d}}{\mathrm{d}t'} \right)^j (t'^j e^{-t'}),
\tag{3.24}
\]
and the identity (A.13). We substitute \(t' \to s_b/t'\) in equation (3.23) and express the derivative in equation (3.24) as a derivative in \(s_b\). The integration over \(t'\) yields
\[
\chi_j^{(M)}(s_b) = \left( -\frac{\mathrm{d}}{\mathrm{d}s_b} \right)^j (s_b \tilde{G}_{0,M}^{1,0}(0, \ldots | s_b)) = (-1)^{j} G_{1,1,M+1}^{1,1,0} \left( \frac{s_b}{s} \right).
\tag{3.25}
\]
The second equality can be obtained by applying definition (A.1) of Meijer’s G-function. We can thus write down the full answer for all \(k\)-point correlation functions of the singular values:
\[
R_k^{(M)}(s_1, \ldots, s_k) = R_{k,0}^{(M)}(s_1, \ldots, s_k; -) = \det_{1 \leq a,b \leq k} [H_{01}(s_a, s_b)]
\]
\[
= \det_{1 \leq a,b \leq k} \left[ \sum_{j=0}^{N-1} \frac{1}{j!} \frac{1}{F_M(-j; 1, \ldots, 1; s_b)} G_{1,1,M+1}^{1,1,0} \left( \frac{s_b}{s} \right) \right].
\tag{3.26}
\]
The simplest example is the density or one-point correlation function of singular values which is given by
\[
R_1^{(M)}(s) = H_{01}(s, s) = \sum_{j=0}^{N-1} \frac{1}{j!} \frac{1}{F_M(-j; 1, \ldots, 1; s)} G_{1,1,M+1}^{1,1,0} \left( \frac{s_b}{s} \right).
\tag{3.27}
\]
This example will be further discussed in subsection 3.3.

The next kernel \(H_{00}\) is readily given from its definition (3.7) together with equations (3.21), (3.19) and (3.20). We therefore turn to \(H_{00}\) from equation (3.9)
\[
H_{00}(t_i, t_j) = t_j^{-1} e^{-t_j} \sum_{i=0}^{N-1} \frac{1}{H_{ii}^{(M)}} \left( \int_0^\infty \mathrm{d}s p_i^{(M)}(s) G_{0,M-1}^{1,0} \left( \frac{s}{t_i} \right) \right) \tilde{L}_i(t_j)
\]
\[
= t_j^{-1} e^{-t_j} \sum_{i=0}^{N-1} \frac{1}{H_{ii}^{(M)}} \psi_i^{(M)}(t_i) \tilde{L}_i(t_j),
\tag{3.28}
\]
where we define the following integral transform:
\[
\psi_i^{(M)}(t) = \int_0^\infty \mathrm{d}s p_i^{(M)}(s) G_{0,M-1}^{1,0} \left( \frac{s}{t_i} \right)
\]
\[
= \sum_{i=0}^{I-1} \frac{(-1)^{I-i}}{(I-i)!} \left( \frac{I}{I} \right)^{M+1} \int_0^\infty \mathrm{d}s G_{0,M-1}^{1,0} \left( \frac{s}{t_i} \right)
\]
\[
= \sum_{i=0}^{I-1} \frac{(-1)^{I-i}}{(I-i)!} \left( \frac{I}{I} \right)^{M+1} t_i^{I+1} \left( \frac{I}{I} \right)^{M-1}
\]
\[
= (I)^{M-1} t_i \tilde{L}_i(t).
\tag{3.29}
\]
In the second step, we have used the identity (A.7), which leads us back to the standard Laguerre polynomials. Taking into account the normalization (3.20), we arrive at the following final result:

\[
H_{10}(t_1, t_j) = \frac{t_j}{t_j} e^{-t_j} \sum_{j=0}^{N-1} \frac{1}{(j!)^2} \tilde{L}_j(t_1) \tilde{L}_j(t_j),
\]

which is proportional to the kernel of ordinary Laguerre polynomials (3.17). The remaining kernel \(H_{11}\) can be expressed in terms of the two integral transforms which we have already computed,

\[
H_{11}(t, s) = \sum_{l=0}^{N-1} \frac{1}{H_{11}^{(M)}} \psi_{l}^{(M)}(t) \chi_{l}^{(M)}(s) = G_{0,M-1}^{1,0} \left( \frac{s}{7} \right).
\]

This completes the computation of all \((k, l)\)-point correlation functions in the 2mm, together with equation (3.4).

Although we will postpone the detailed analysis of the large-\(N\) limit to future work, let us mention the following nontrivial identity with respect to the kernel \(H_{11}\):

\[
\sum_{l=0}^{N-1} \frac{1}{(l!)^2} \tilde{L}_l(t) \chi_{l}^{(M)}(s) = G_{0,M-1}^{1,0} \left( \frac{s}{7} \right),
\]

implying that \(\lim_{N \to \infty} H_{11}(t, s) = 0\). Assuming that the sum converges and can be integrated piecewise this can be shown as follows. In appendix B, we verify that \(p_{l}^{(M)}(s)\) and \(\chi_{l}^{(M)}(s)\) form a set of orthogonal functions with respect to the flat measure, see equation (3.34). After multiplying both sides of equation (3.32) with \(p_{l}^{(M)}(s)\) and integrating \(s\) over \(\mathbb{R}\), we obtain

\[
(j!)^{M-1} t_j \tilde{L}_j(t) = \int_{0}^{\infty} ds \ p_{l}^{(M)}(s) G_{0,M-1}^{1,0} \left( \frac{s}{7} \right) = \psi_{l}^{(M)}(t),
\]

which is consistent with equation (3.29).

The identity (3.32) most likely implies that in the naive large-\(N\) limit, meaning \(s, t\) and \(M\) fixed, the correlation function \(H_{11}^{(M)} \) factorizes into \(s\)- and \(t\)-dependent parts regardless whether \(H_{00}\) vanishes or not since the determinant (3.4) factorizes into a \(k \times k\) determinant incorporating \(H_{01}\) and an \(l \times l\) determinant comprising the block \(H_{10}\). Therefore, the correlation functions of the singular values \(t_j\) of the matrix \(X_j\) decouple and become the ones of the standard Wishart–Laguerre type. There may be other ways to obtain a nontrivial coupling between singular values \(s_a\) and \(t_j\) in a more sophisticated large-\(N\) limit (like in the so-called weak limit in [25]).

We finally make contact again with the one-matrix model formulation (2.12). The following orthogonality relation which follows from equation (3.2) is explicitly verified in appendix B:

\[
\int_{0}^{\infty} ds p_{l}^{(M)}(s) \chi_{l}^{(M)}(s) = (l!)^{M+1} \delta_{l} ;
\]

in other words \(p_{l}^{(M)}(s)\) and \(\chi_{l}^{(M)}(s)\) constitute a set of biorthogonal functions with respect to a single variable with flat measure. In particular, this relation results in the following property of the kernel \(H_{01}\), see equation (3.22), that contains the two:

\[
\int_{0}^{\infty} ds_{1} H_{01}(s_{1}, s_{1}) H_{01}(s_{1}, s_{2}) = H_{01}(s_{1}, s_{2}).
\]

We have thus closed the circle back to the jpdf (2.13) where we can directly replace

\[
\Delta_{N}(s) \det_{1 \leq c, d \leq N} \left[ G_{0,M}^{a,b \to a-1,d-1} | x_{c} \right] = \det_{1 \leq a, b \leq N} \left[ p_{a-1}^{(M)}(s_b) \right] \det_{1 \leq c, d \leq N} \left[ \chi_{c-1}^{(M)}(s_d) \right] \]

\[
= \prod_{j=0}^{N-1} h_{1}^{(M)} \det_{1 \leq a, b \leq N} \left[ H_{01}(s_a, s_b) \right].
\]
This uses the invariance property of determinants under the addition of columns, and then proceeds with standard techniques to deduce the correlation functions. This directly leads from equation (2.13) to equation (3.10) for the jpdf. With the property (3.35) of the kernel we deduce equation (3.26) from Dyson’s theorem [24].

3.3. Spectral density, its moments and large- \( N \) scaling

In this subsection, we discuss in more detail the implications of our results for the spectral density of singular values. Starting from expression (3.27) which we repeat here in two equivalent forms,

\[
R_1^{(M)}(s) = \sum_{l=0}^{N-1} \sum_{i,j=0}^l \frac{(-1)^{j+i} (l!)^2}{(l-j)! (l-i)! (j!) (j!)^{M+1}} C_{0,M}^{M,0}(\ldots; i+j) \mid s \rangle^n.
\]

we can explicitly compute expectation values for the moments for finite- \( N \). Starting from the first expression, we obtain

\[
\mathbb{E}[s^k] = \frac{1}{N} \int_0^\infty ds \ s^k R_1^{(M)}(s)
\]

\[
= \frac{1}{N} \sum_{l=0}^{N-1} \sum_{i,j=0}^l \frac{(-1)^{j+i} (l!)^2 (i+j+k)! ((j+k)!)^{M+1}}{(l-j)! (l-i)! (j!) (j!)^{M+1}}.
\]

Here, we have normalized by equation (2.19) and we have used again the identity (A.7) for moments of the Meijer \( G \)-function. On the other hand, using the compact expression for \( \chi_j^{(M)}(s) \) in the second formulation of the density we can obtain a more concise result in the following way:

\[
\mathbb{E}[s^k] = \frac{1}{N} \sum_{l=0}^{N-1} \sum_{j=0}^l \frac{(-1)^{j-i} (j!)^{M+1}}{(j-i)! (j!) (j!)^{M+1}} \int_0^\infty ds s^{j+k} (-1)^j G_{1,M+1}^{M,1}(0_{j},0|s)
\]

\[
= \frac{1}{N} \sum_{l=0}^{N-1} \sum_{j=0}^l \frac{(-1)^{j-i} (j!)^{M+1}}{(j-i)! (j!) (j!)^{M+1}} \sum_{j=0}^{j+k} \frac{(-1)^{j} (j!)^{M+1}}{j! (j-i)!}
\]

\[
= \frac{1}{N} \sum_{l=0}^{N-1} \frac{(j+k)!}{l!} \sum_{j=0}^{j+k} \frac{(-1)^{j} (j!)^{M+1}}{j! (j-i)!}
\]

\[
= \frac{1}{N} \sum_{l=0}^{N-1} \frac{(j+k)!}{l!} \int_0^{2\pi} d\varphi \frac{1}{2\pi k!} (1 - e^{i\varphi})^k \left( 1 - e^{-i(N-l-1)\varphi} \right)
\]

\[
= \frac{1}{N} \sum_{l=0}^{N-1} \frac{(-1)^{N-l-1}}{k!} (k+l)! \left( \frac{k+l}{l!} \right)^{M+1} \left( \frac{k-l}{N-l-1} \right).
\]

We use the convention that inverse powers of factorials of negative integers give zero, rather than using the Gamma-function everywhere. In the first step, we have employed equation (A.8). Interestingly, the remaining sum can be further expressed in terms of a hypergeometric function if \( k \geq N \).

\[
\mathbb{E}[s^k] = (-1)^{N-1} \frac{(k!)^{M-1} (k-1)!}{N! \Gamma(k-N+1)} \times_{M+2} F_{M+1}(k+1, \ldots, k+1, 1-N; 1, \ldots, 1, k-N+1; 1),
\]

\[
(3.40)
\]
by extending the sum to infinity and comparing their Taylor series. Indeed, this relation can be
generalized to \( k < N \). Note that in this case the singular contributions in the hypergeometric
function cancel with those in the Gamma-function in the denominator.

In the ensuing discussion, we will need in particular the first moment \( F_N^{(M)} \) for \( k = 1 \)
when rescaling the density, which can be readily read off

\[
E[x] \equiv F_N^{(M)} = NM.
\] (3.42)

It agrees with the known case for \( M = 1 \). An alternative short derivation for the first moment
is sketched in appendix B. Higher moments easily follow from equation (3.40), e.g. for the
second moment we have

\[
E[x^2] = \frac{1}{2} NM ((N+1)^{M+1} - (N-1)^{M+1}).
\] (3.43)

We illustrate our results for the density (3.27) by plotting it for various values of \( N \) and \( M \). As a first example in figure 1 the density is shown for \( M = 1, 2, 3 \) at fixed \( N = 4 \). Clearly, it is
mandatory to know the right scale dependence of the correlation functions on \( N \) and \( M \). This
means that after properly rescaling the bulk of the singular values is of order 1, in order to be
able to compare the density for finite \( N \) at different values of \( N \). In particular, it is important
to check the finite \( N \)-results against the limiting large-\( N \) behaviour for different \( M \), which has
been derived for the products of quadratic [6, 7] and rectangular matrices [8, 9].

Let us explain our procedure. First, we normalize our density to unity, using
equation (2.19). Then, we rescale the density by its first moment,

\[
\hat{K}_1^{(M)}(x) \equiv \frac{1}{N} F_N^{(M)} R_1^{(M)}(F_N^{(M)} x),
\] (3.44)

so that the new density \( \hat{K}_1^{(M)}(x) \) has a norm and first moment equal to unity. Note that this
rescaling is an alternative to the unfolding procedure onto the scale of the local mean level
spacing. Instead of fixing the mean distance between two successive singular values, we fix
here the singular values themselves, such that they are always of order 1. This is exactly the
macroscopic limit.

Inserting equation (3.42), we thus obtain a limiting density, which we denote by,

\[
\hat{\rho}^{(M)}(x) \equiv \lim_{N \to \infty} \hat{K}_1^{(M)}(x) = \lim_{N \to \infty} N^{M-1} R_1^{(M)}(N^M x),
\] (3.45)

8 All densities satisfying this property will be denoted with a hat ' \( \hat{\cdot} \) '.

Figure 1. Comparison of the density (3.27) \( R^{(M)}(s) \) for fixed \( N = 4 \) without rescaling. The values
\( M = 1, 2, 3 \) correspond to the top (blue), middle (red) and bottom (black) curve, respectively.
which is chosen subject to boundary conditions and to yield a real density on the support equation (3).

The resolvent is related to the limiting spectral density by

\[
\lim_{N\to\infty} R_1^{(M=1)}(Nx) \approx \frac{1}{2\pi} \sqrt{\frac{4N - s}{s}} \Theta(4N - x),
\]

which is the $N$-dependent Marchenko–Pastur density and $\Theta(x)$ denotes the Heaviside function. With the first moment being given by $F_N^{(M=1)} = N$, we thus have

\[
\hat{\rho}^{(M=1)}(x) = \lim_{N\to\infty} N^0 R_1^{(M=1)}(Nx) = \frac{1}{2\pi} \sqrt{\frac{4 - x}{x}} \Theta(4 - x),
\]

for the rescaled and normalized density. This is the Marchenko–Pastur density with compact support on $(0, 4]$. A comparison between $\hat{\rho}^{(M=1)}(x)$ and $\bar{R}_1^{(M=1)}(x)$ for various values of $N = 3, 4, 5$ and 10 is given in figure 2 (left).

For $M > 1$, the limiting expression for the density is not as explicit as in equation (3.47). Here, we will follow the notation of [9] where a polynomial equation for the resolvent $G(z)$ was derived, which we display for the case of quadratic matrices only:

\[
(zG^{(M)}(z))^{M+1} = z(zG^{(M)}(z) - 1).
\]

The resolvent is related to the limiting spectral density by

\[
G^{(M)}(z) \equiv \int_0^\infty d\lambda \frac{\rho^{(M)}(\lambda)}{z - \lambda},
\]

where $z$ is outside the support of the spectral density. This relation can be inverted as follows:

\[
\rho^{(M)}(\lambda) = -\frac{1}{\pi} \lim_{\epsilon \to 0^+} \Im \{G^{(M)}(\lambda + i\epsilon)\}.
\]

For $M = 1$, equation (3.48) reduces to a quadratic equation, which after taking the discontinuity along the support equation (3.50) leads to equation (3.47), without further rescaling.

Increasing to $M = 2$ the equation becomes cubic and we can still write out its solution, which is chosen subject to boundary conditions and to yield a real density on the support $(0, 3^{1/2})$.

We thank Z. Burda for indicating how to determine the limiting support.
\[ G^{(M=2)}(z) = \frac{1}{\sqrt{3\varepsilon}} \left( A_1^{1/3}(z) + A_2^{1/3}(z) \right) = \frac{1}{\sqrt{3\varepsilon}} \left( (-A_+(z))^{1/3} + A_1^{1/3}(z) \right), \]

\[ A_\pm(z) = \sqrt{\frac{27}{4\varepsilon}} - 1 \pm \sqrt{\frac{27}{4\varepsilon}}. \] (3.51)

The density \( \hat{\rho}^{(M=2)}(s) \) that is obtained from equation (3.51) by taking the discontinuity according to equation (3.50) (which happens to have the first moment equal to unity without further rescaling) is shown in figure 2 (right) in comparison to our rescaled finite-\( M \) density \( \hat{\rho}_s^{(M=2)}(s) \) for various values of \( N \). As in the known case \( M = 1 \) we obtain a nice agreement for \( M = 2 \). An alternative derivation of \( \hat{\rho}^{(M=2)}(s) \) via multiple orthogonal polynomials was quite recently presented in [31], with which our density agrees.

As a final remark, for larger \( M \) it might be useful to resolve the singularity of the density at the origin [9], \( \lim_{s \to 0} \hat{\rho}^{(M)}(s) \sim s^{-M/(M+1)} \), by changing variables, just as it is well known that for \( M = 1 \) a change to squared variables maps the Marchenko–Pastur density equation (3.47) to the semi-circle. A more detailed comparison to existing large-\( N \) results for the density, its moments and support would require a careful asymptotic analysis of the special functions constituting our finite-\( N \) density, and is postponed to future work.

4. Applications to telecommunication

Consider a MIMO network with a single source and destination, both equipped with \( N \) antennas. Information transmitted by the source is conveyed to the destination via \( M - 1 \) successive clusters of scatterers, where each cluster (layer) is assumed to have \( N \) scattering objects. Such a channel model proposed in [8] is typical in modelling the indoor propagation of information between different floors [32].

We assume that the vector-valued transmitted signal propagates from the transmitter array to the first cluster, from the first to the second cluster and so on, until it is received from the \((M-1)\)st cluster by the receiver antenna array. Each communication channel is described by a random complex Gaussian matrix, and as a result the effective channel of this multi-layered model equals the product matrix \( P_M \), see equation (2.1).

For the described communication channels, the mutual information measured in units of the natural logarithm (nats) per second per Hertz is defined as

\[ I(\gamma) \equiv \ln \det (I_N + \frac{\gamma}{N^M} P_M P_M^\dagger) = \sum_{i=1}^{N} \ln \left( 1 + \frac{\gamma}{N^M s_i} \right), \] (4.1)

where \( \gamma \) defines the average received signal-to-noise ratio per antenna which is a constant. We employ the distribution (3.27) of squared singular values to compute its average. The quantity of interest is called the ergodic mutual information of such channels. It is given by the expectation value of the random variable \( I(\gamma) \). Using the analogue of expression (3.38) from the previous section we have

\[ \mathbb{E}[I(\gamma)] = N \mathbb{E} \left[ \ln \left( 1 + \frac{\gamma}{N^M s^2} \right) \right] \]

\[ = \sum_{i=0}^{N-1} \sum_{\{i,j\} \neq \emptyset} \frac{(-1)^{i+j} (I!)^2}{(i-j)! (i-I)! (j+I)! (j-I)!} \int_0^\infty ds G_{0,M}^{(M)} \left( j, \ldots, i, i+j \mid s \right) \]

\[ \times \ln \left( 1 + \frac{\gamma}{N^M s} \right) \] (4.2)
Figure 3. The ergodic mutual information of multi-layered scattering MIMO channels with a fixed number of 2 clusters ($M = 3$) with a different number of scatters $N = 2, 4, 8$.

\[
E[I(\gamma)] = \sum_{l=0}^{N-1} \sum_{i,j=0}^{l} \frac{(-1)^{l+j}(l!)^2}{(l-j)!(l-i)!(j!)^{M+1}} 
\times G_{2,M+2}^{M+2,1} \left( \begin{array}{c} 0, 1 \\ 0, 0, j+1, \ldots, j+1, i+j+1 \end{array} \middle| \frac{N^M}{\gamma} \right).
\] (4.3)

The last step is obtained by first replacing the logarithm by a Meijer $G$-function, equation (A.5), and then applying the integral identity (A.14) from the appendix. Note that the corresponding ergodic mutual information for the traditional MIMO channel model, i.e. $M = 1$, was derived in [33].

In order to get an independent confirmation of our analytical result (4.3), we compare it to numerical simulations as follows. We plot the ergodic mutual information (4.3) in figure 3 against Monte Carlo simulations as a function of $\gamma$ in decibel (dB) for a given number of clusters, $M - 1$, as an example, with different numbers of scatters per cluster, $N$. Each simulated curve is obtained by averaging over $10^6$ independent realizations of $P_M$. The statistical error bar is smaller than the symbol for the simulation in our plots. The comparison in figure 3 shows a two-layered scattering channel, i.e. $M = 3$, with the number of scatters per layer varying from $N = 2, 4$ to 8. We have also compared our results to simulations for other values of $N$ and $M$.

5. Conclusions and open questions

In this paper, we have derived the joint probability distribution function (jpdf) of singular values for any finite product of $M$ quadratic random matrices of finite size $N \times N$, with complex elements distributed according to a Gaussian distribution. This generalizes the Wishart–Laguerre (also called chiral Gaussian) unitary ensemble which we recover for $M = 1$. Starting from the jpdf we have computed all $k$-point density correlation functions of the singular values, by taking a detour over a two-matrix model-like representation of the same model. In that way, we showed that the jpdf being proportional to a Vandermonde times the determinant of Meijer $G$-functions represents a determinantal point process. Its kernel of orthogonal functions generates all $k$-point functions in the standard way using Dyson’s theorem. We also solved the auxiliary two-matrix model that couples a single matrix to the product of $M$ matrices, by constructing the biorthogonal polynomials explicitly, as well as the corresponding four kernels.
with their integral transforms. On the way we found some nontrivial identities and integral representations of the Meijer G-function.

The density of the singular values are discussed in more detail at finite $N$ and $M$, including all its moments. We identified the macroscopic scaling to match the density with the known large-$N$ results for the macroscopic density of singular values. As a further application we have computed the averaged mutual information for multi-layered scattering of MIMO channels and have compared them to Monte Carlo simulations for small $M$ and $N$.

Previous results for the macroscopic large-$N$ density of the singular values of quadratic or rectangular matrices and its expectation values of traces have mainly been obtained from probabilistic methods, in particular, using free random variables. The explicit results that we have obtained for the jpdf and all correlation functions thereof open up the possibility of another direction. One can now investigate the microscopic scaling limits zooming into various parts of the spectrum, by performing the asymptotic analysis of the orthogonal polynomials and their integral transforms that we computed. Since the ensemble represents a determinantal point process, one can also investigate the limiting distribution of individual eigenvalues, e.g. by considering their Fredholm determinant representation. Moreover, one can now study the distributions of linear statistics such as the trace of the derived ensemble as well.

Based on known results for the universality of the spectrum of random matrices we expect the following outcome for such an analysis. The bulk and the soft edge behaviour of the spectrum should be governed by the universal Sine- and Airy-kernels, respectively, after unfolding and scaling appropriately. This includes the fact that at the soft edge we expect to find the Tracy–Widom distribution, and it will be very interesting to identify the right scaling for that. In contrast, at the hard edge we expect to find new universality classes labelled by $M$. The way the macroscopic density diverges at the origin already depends on it. This would fit into what was found recently for the complex eigenvalue spectrum of products of independent Ginibre matrices.

Other open problems include a generalization of our construction to the products of rectangular matrices, which seems quite feasible. Also, the inclusion of determinants (or characteristic polynomials) into the weight, which should be related to rectangular matrices, seems within reach. On the other hand, moving on to non-Gaussian weight functions or investigating other symmetry classes with real or quaternion real matrix elements are very challenging problems. The reason is that our method depends crucially on the Harish-Chandra–Itzykson–Zuber group integral to eliminate the angular variables. For the other symmetry classes, no such explicit tool is yet known.

Acknowledgments

We would like to thank Zdzisław Burda and Ralf Müller for useful discussions and correspondence. GA thanks the Laboratori Nazionali di Frascati for their hospitality where part of this work was carried out. We acknowledge partial support by the SFB | TR12 ‘Symmetries and Universality in Mesoscopic Systems’ of the German research council DFG (GA), by the Alexander von Humboldt foundation (MK), as well as by the Finnish Centre of Excellence in Computational Inference Research and the Nokia Foundation (LW).

Appendix A. Some integral identities for Meijer G-functions

In this appendix, we collect a few integral representations and identities for the so-called Meijer G-function. It is defined as [23]
functions related to the \(G_m\) \(J\). Phys. A: Math. Theor. given by \(\[23\]\)

of the path. We are, in particular, interested in the case

Note that the function is symmetric in all its indices

In the first line, we simply substituted due to the following shift \(\[23\]\):

Macdonald function. It is possible to absorb powers of the argument of the Meijer

\(K\)

The contour of integration \(C\) goes from \(-i\infty\) to \(+i\infty\) such that all poles of the Gamma functions related to the \(b_j\) lie to the right of the path, and all poles related to the \(a_j\) to the left of the path. We are, in particular, interested in the case

The rest follows from the definition of the Meijer \(G\)-function and the Gamma-function. Note that for
$d = 1$ and $b_1 = \ldots = b_m = 0$, this recursion for the Meijer G-function was already derived in [16].

We are now prepared to show the multiple integral representation of the Meijer G-function that we need in the derivation of the pdf in section 2. It slightly generalizes the representation found in [16]. The statement is that

$$
G_{0,m} \left( \begin{array}{c} -x_m \\ 0, \ldots , 0 \end{array} \right| \frac{x_0}{x_0} ) = \int_0^{\infty} \frac{dx_1}{x_1} \left( \frac{x_1}{x_0} \right)^b \int_0^{\infty} \frac{dx_2}{x_2} \ldots \int_0^{\infty} \frac{dx_m}{x_m} \prod_{j=1}^{m} e^{-x_j/x_{j-1}}, \quad (A.10)
$$

for $m > 1$. Comparing this equation to equation (2.11) with $b = d - 1$, we work with $m = M$ squared singular values, $x_j = \left( \lambda^{(j)}_c \right)^2$, where we introduce a dummy variable $x_0$. The variable $x_0$ will be set to unity for our original purposes. However, it will be useful when applying the identity to equation (2.21). Our proof goes by induction in $m$. For $m = 2$ [23], we have

$$
\int_0^{\infty} \frac{dx_1}{x_1} \left( \frac{x_1}{x_0} \right)^b e^{-\frac{a}{x_0} - \frac{\gamma}{x_1}} = 2 \left( \frac{x_2}{x_0} \right)^{b/2} K_{b/2}(2\sqrt{x_2/x_0}) = G_{0,2}^{2,0} \left( \begin{array}{c} 0 \\ 0, b \end{array} \right| \frac{x_2}{x_0} ) , \quad (A.11)
$$

where the last step is due to equation (A.4). This leads to

$$
\int_0^{\infty} \frac{dx_2}{x_2} G_{0,2}^{2,0} \left( \begin{array}{c} 0 \\ 0, b \end{array} \right| \frac{x_3}{x_0} ) e^{-x_2/x_3} = G_{0,3}^{3,0} \left( \begin{array}{c} -x_3 \\ 0, 0, b \end{array} \right| \frac{x_3}{x_0} ) . \quad (A.12)
$$

for $m = 3$, where we have applied the identity (A.9) in its second form shown in the first line, with $d = 1$ and $v = x_2/x_0$. For the induction step $m - 1 \rightarrow m$, we simply have to repeat the same procedure, which follows easily from the very same identity

$$
\int_0^{\infty} \frac{dx_m}{x_m} G_{0,m}^{m,0} \left( \begin{array}{c} -x_m \\ 0, \ldots , 0, b \end{array} \right| \frac{x_m}{x_0} ) e^{-x_{m+1}/x_m} = G_{0,m+1}^{m+1,0} \left( \begin{array}{c} -x_{m+1} \\ 0, \ldots , 0, b \end{array} \right| \frac{x_{m+1}}{x_0} ) . \quad (A.13)
$$

which completes the proof.

Note that the same identity (A.10) can be used to provide the second step in equation (2.21), when setting $b = 0$, $m = M - 1$ and shifting the indices of the variables $x_{j-1} = \left( \lambda^{(j)}_c \right)^2$ for $j = 1, \ldots , M$. This is the reason why $x_0$ is useful.

Finally, we state an integral identity needed in section 4, concerning the integral of two Meijer G-functions,

$$
\int_0^{\infty} ds \ G_{2,2}^{1,2} \left( \begin{array}{c} 1 \end{array} \right| \begin{array}{c} 1 \end{array} \right| \frac{y}{N} ) G_{0,M}^{M,0} \left( \begin{array}{c} - \\ j, \ldots , j, i + j \end{array} \right| \begin{array}{c} 0 \\ 0 \end{array} ) = \frac{N!}{y^{M+2}} \ G_{2,M+2}^{M+2,1} \left( \begin{array}{c} -1, -1, j, \ldots , j, i + j \end{array} \right| \begin{array}{c} 0 \\ y \end{array} ) . \quad (A.14)
$$

Note that it is a particular choice of a general formula [34]. In order to arrive at equation (4.3) we apply the shift (A.6).

Appendix B. Orthogonality check and first moment

In this appendix, we explicitly confirm both the orthogonality (3.2) of the bOP $p_n^{(M)}(s)$ and $q_i^{(M)}(r)$ with respect to two variables, as well as the fact that $p_n^{(M)}(s)$ and $x_i^{(M)}(s)$ constitute a set of biorthogonal functions with respect to one variable with flat measure, equation (3.34). Although being true by construction, we will see the orthogonality ultimately boils down to the standard orthogonality of Laguerre polynomials.

The biorthogonal polynomials that were constructed in subsection 3.1 using the bimoment matrix must automatically satisfy the orthogonality relation (3.2). We will check this here independently, which implies at the same time that one of the polynomials and the integral
transform (3.23) of the other are orthogonal functions (as they should be, in order to constitute proper kernels):

\[
\int_0^\infty ds \, p^{(M)}_j(s) \hat{X}_j^{(M)}(s) = \int_0^\infty ds \int_0^\infty dt \, w^{(M)}(s, t) p^{(M)}_j(s) q^{(M)}_j(t)
\]

\[
= \int_0^\infty \int_0^\infty dt \, e^{-s t} \tilde{L}_j(t) \sum_{k=0}^j \frac{(-1)^{j-k}}{(i-k)!} \left( \frac{it}{k!} \right)^{M+1} \int_0^\infty \frac{ds}{s} e^{s-1} (\sum_{0, \ldots, 0} \frac{s}{t})^{k}
\]

\[
= \int_0^\infty \int_0^\infty dt \, e^{-s t} \tilde{L}_j(t) \sum_{k=0}^j \frac{(-1)^{j-k}}{(i-k)!} \left( \frac{it}{k!} \right)^{M+1} t^k (k!)^{M-1}
\]

\[
= (it)^{M+1} \int_0^\infty dt \, e^{-1} \tilde{L}_j(t) \tilde{L}_0(t)
\]

\[
= (it)^{M+1} \delta_{ij}.
\]

(B.1)

Here, we have used the identity (A.7) for moments of the Meijer G-function, which cancels the extra factorials in the generalized Laguerre polynomials \( p^{(M)}_j(s) \), see equation (3.21), after the first integration over \( s' = s/t \). The last step follows from the known orthogonality of Laguerre type.

In the second part of this appendix, we provide a much simpler, probabilistic argument that leads to the first moment equation (3.42). Noting that

\[
\mathbb{E}[X_j X_j'] = N I_N, \quad \text{for } j = 1, \ldots, L,
\]

we have that

\[
\mathbb{E}[s] = \frac{1}{N} \mathbb{E} \left[ \sum_{a=1}^N s_a \right] = \frac{1}{N} \mathbb{E} \left[ \text{Tr}(P_M^a P_M^a) \right]
\]

\[
= \frac{1}{N} \text{Tr} \left( \prod_{j=1}^M \mathbb{E}[X_j X_j'] \right) = \frac{1}{N} N^M \text{Tr}(I_N) = N^M.
\]

(B.3)

where we have reordered successively the \( X_j \) under the trace in the first line.

References

[1] Akemann G, Baik J and Di Francesco P (ed) 2011 The Oxford Handbook of Random Matrix Theory (Oxford: Oxford University Press)
[2] Crisanti A, Paladin G and Vulpiani A 1993 Products of Random Matrices in Statistical Physics (Springer Series in Solid-State Sciences vol 104) (Heidelberg: Springer)
[3] Penson K A and Życzkowski K 2011 Product of Ginibre matrices: Fuss–Catalan and Raney distributions Phys. Rev. E 83 061118 (arXiv:1103.3453v3 [math-ph])
[4] Tulino A M and Verdú S 2004 Random matrix theory and wireless communications Found. Trends Commun. Inform. Theory 1 1–182
[5] Tulino A M and Verdú S 2011 Asymptotic singular value distributions in information theory The Oxford Handbook of Random Matrix Theory (Oxford: Oxford University Press) chapter 41, pp 851–72
[6] Deift P A 1999 Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach (Courant Lecture Notes in Mathematics vol 3) (New York: New York University, Courant Institute of Mathematical Sciences)
[7] Bónac T, Belinschi S, Capitaine M and Collins B 2011 Free Bessel laws Can. J. Math. 63 3–37 (arXiv:0910.5931 [math.PR])
[8] Benaych-Georges F 2010 On a surprising relation between the Marchenko–Pastur law, rectangular and square free convolutions Ann. Inst. Henri Poincaré Probab. Stat. 46 644–52 (arXiv:0808.3938 [math.PR])
[9] Mütter R 2002 On the asymptotic eigenvalue distribution of concatenated vector-valued fading channels IEEE Trans. Inform. Theory 48 2086–91
[9] Burda Z, Jarosz A, Livin G, Nowak M A and Swiech A 2010 Eigenvalues and singular values of products of rectangular Gaussian random matrices (the extended version) *Phys. Rev. E* **82** 066114 (arXiv:1007.3594 [cond-mat.stat-mech])

Burda Z, Jarosz A, Livin G, Nowak M A and Swiech A 2011 *Acta Phys. Pol. B* **42** 939–85 (arXiv:1103.3964 [cond-mat.stat-mech])

[10] Newman C M 1986 The distribution of Lyapunov exponents: exact results for random matrices *Commun. Math. Phys.* **103** 121–6

[11] Iosip M and Newman C M 1992 The triangle law for Lyapunov exponents of large random matrices *Commun. Math. Phys.* **143** 591–8

[12] Jackson A D, Lautrup B, Johansen P and Nielsen M 2002 Products of random matrices *Phys. Rev. E* **66** 066124 (arXiv:physics/0202037 [physics.data-an])

[13] Forrester P J 2013 Lyapunov exponents for products of complex Gaussian random matrices *J. Stat. Phys.* **151** 796–808 (arXiv:1206.2001 [math.PR])

[14] Akemann G and Strahov E 2013 Hole probabilities and overcrowding estimates for products of complex Gaussian matrices *J. Phys. A: Math. Theor.* **45** 465201 (arXiv:1208.0187 [math-ph])

[15] Bertola M 2011 Two-matrix models and biorthogonal polynomials *The Oxford Handbook of Random Matrix Theory* (Oxford: Oxford University Press) chapter 15, pp 310–28

[16] Bertola M, Gekhtman M and Szmigielski J 2012 Cauchy–Laguerre two-matrix model and the Meijer-G random point field arXiv:1211.5369 [math.PR]

[17] Zhang L 2013 A note on limiting mean distribution of singular values for products of two Wishart random matrices arXiv:1305.0726 [math-ph]

[18] Saunders S R and Aragón-Zavala A 2007 *Antennas and Propagation for Wireless Communication Systems* 2nd edn (New York: Wiley)

[19] Telatar I E 1999 Capacity of multi-antenna Gaussian channels *Eur. Trans. Telecommun.* **10** 585–95

[20] Prudnikov A P, Brychkov Y A and Marichev O I 1990 *Integrals and Series: More Special Functions* vol 3 (London: Gordon and Breach)