Products of rectangular random matrices: Singular values and progressive scattering

Gernot Akemann,[*] Jesper R. Ipsen,[†] and Mario Kieburg[‡]
Department of Physics, Bielefeld University, Postfach 100131, D-33501 Bielefeld, Germany
(Received 5 August 2013; published 11 November 2013)

We discuss the product of M rectangular random matrices with independent Gaussian entries, which have several applications, including wireless telecommunication and econophysics. For complex matrices an explicit expression for the joint probability density function is obtained using the Harish-Chandra-Itzykson-Zuber integration formula. Explicit expressions for all correlation functions and moments for finite matrix sizes are obtained using a two-matrix model and the method of biorthogonal polynomials. This generalizes the classical result for the so-called Wishart-Laguerre Gaussian unitary ensemble (or chiral unitary ensemble) at $M = 1$, and previous results for the product of square matrices. The correlation functions are given by a determinantal point process, where the kernel can be expressed in terms of Meijer G-functions. We compare the results with numerical simulations and known results for the macroscopic level density in the limit of large matrices. The location of the end points of support for the latter are analyzed in detail for general M. Finally, we consider the so-called ergodic mutual information, which gives an upper bound for the spectral efficiency of a MIMO communication channel with multifold scattering.

DOI: 10.1103/PhysRevE.88.052118 PACS number(s): 02.50.Sk, 02.10.Yn, 02.30.Cj, 84.40.Ba

I. INTRODUCTION

Random matrix theory has existed for more than half a century, and its success is undeniable. A vast number of applications is known within the mathematical and physical sciences and beyond; we refer to [1] for a recent overview. A direction within random matrix theory, which has recently caught renewed attention is the study of products of random matrices. Among others, products of matrices have been applied to disordered and chaotic systems [2], matrix-valued diffusions [3,4], quantum chromodynamics at finite chemical potential [5,6], Yang-Mills theory [7–9], finance [10], and wireless telecommunication [11]. In this paper, our attention is directed towards the latter.

When considering products of matrices we are faced with the fact that the product often possesses fewer symmetries than the individual matrices. For example, a product of symmetric matrices will not be symmetric, in general. For simplicity, we look at matrices with a minimum of symmetry. Our discussion concerns products of matrices drawn from the Wishart ensemble. Thus, the matrices have independently, identically distributed Gaussian entries. Also other proposals exist, e.g. by multiplying matrices that are chosen from a set of *fixed* matrices with a given probability. This problem has applications in percolation, as was pointed out in [12]. However, it considerably differs from our approach, notably due to the lack of invariance.

The statistical properties of the complex eigenvalues and real singular values of a product of matrices from the Wishart ensemble have been discussed in several papers (in the former case they are usually called Ginibre matrices). Macroscopic properties for eigenvalues of complex ($\beta = 2$) matrices have been discussed in the limit of large matrices using diagrammatic methods [4,13,14], while proofs are given in [15,16]. The macroscopic behavior of the singular values and their moments have also been discussed in the literature using probabilistic methods [17–19] as well as diagrammatic methods [14].

Recently, the discussion of products of matrices from Wishart ensembles has been extended to matrices of finite size [20–23], but this discussion has so far been limited to the case of square matrices. We want to extend this discussion to include products of rectangular matrices. In particular, we consider the product matrix

$$\mathbb{Y}_M = \mathbb{X}_M \mathbb{X}_{M-1} \cdots \mathbb{X}_1, \tag{1}$$

where $\mathbb{X}_m$ are $N_m \times N_{m-1}$ real ($\beta = 1$), complex ($\beta = 2$), or quaternion ($\beta = 4$) matrices from the Wishart ensemble. This paper is concerned with the singular values of such matrices, and the spectral correlation functions of $\mathbb{Y}_M \mathbb{Y}_M^\dagger$. A discussion of the complex eigenvalues is postponed to a future publication [24].

Matrix products like $\mathbb{Y}_M$ have direct applications in finance [10], wireless telecommunication [17], and quantum entanglement [25,26]. The importance of the generalization from square to rectangular matrices is evident from its applications to, e.g. wireless telecommunication. Let us consider a multiple-input–multiple-output (MIMO) communication channel from a single source to a single destination via $M - 1$ clusters of scatterers. The source and destination are assumed to be equipped with N_0 transmitting and N_M receiving antennas, respectively. Each cluster of scatterers is assumed to have N_m ($1 \leqslant m \leqslant M - 1$) scattering objects. Such a communication link is canonically described by a channel matrix identical to the complex version of the product matrix (1). Here the Gaussian nature of the matrix entries models a Rayleigh fading environment. This model was proposed in [17], while the single channel model ($M = 1$) goes back to [27–29]. There is no reason to assume that the number of scattering objects at each cluster in such a communication channel should be identical, which illustrates the importance of the generalization to rectangular matrices.

[*]akemann@physik.uni-bielefeld.de
[†]jipsen@math.uni-bielefeld.de
[‡]mkieburg@physik.uni-bielefeld.de

This paper is organized as follows. In Sec. II we find the joint probability density function for the singular values of the product matrix (1) in the complex case. Starting with general $\beta = 1,2,4$ it turns out that the restriction to complex ($\beta = 2$) matrices is necessary, since our method relies on the Harish-Chandra-Itzykson-Zuber integration formula for the unitary group [30,31]. An explicit expression for all k-point correlation functions for the singular values are derived in Sec. III using a two-matrix model and the method of biorthogonal polynomials. The spectral density and its moments are discussed further in Sec. IV, while we return to the above mentioned communication channel in Sec. V. Section VI is devoted to conclusions and outlook. Some properties and identities for the special functions we encounter are collected in the appendix.

II. JOINT PROBABILITY DISTRIBUTION OF SINGULAR VALUES

As mentioned in the introduction we are interested in the statistical properties of the singular values of the product matrix (1), which is governed by the following partition function,

$$Z_\beta^M = \prod_{m=1}^{M} \int |D\mathbb{X}_m| \exp[-\operatorname{Tr} \mathbb{X}_m \mathbb{X}_m^\dagger]. \tag{2}$$

Here $D\mathbb{X}_m$ denotes the Euclidean volume, i.e. the exterior product of all independent one forms, while $|D\mathbb{X}_m|$ is the corresponding unoriented volume element.

Let us assume that the smallest dimension is $N_0 = N_{\min}$. We stress that the properties of the nonzero singular values of $\mathbb{Y}_M$ are completely independent of this choice; see [24]. Thus, the product matrix, $\mathbb{Y}_M = \mathbb{X}_M \cdots \mathbb{X}_1$, has maximally rank N_0. It follows that the product matrix can be parametrized as [24]

$$\mathbb{Y}_M = \mathcal{U}_M \begin{pmatrix} Y_M \\ 0 \end{pmatrix}, \tag{3}$$

where Y_M is a square $N_0 \times N_0$ matrix with real, complex, or quaternion entries, while $\mathcal{U}_M$ is an orthogonal, a unitary, or a unitary symplectic matrix for $\beta = 1,2,4$, respectively. From Eq. (3) it is immediate that the nonzero singular values of the rectangular matrix $\mathbb{Y}_M$ are identical to the singular values of the square matrix Y_M. The ultimate goal is to derive the joint probability density function for these singular values. In [24] the invariance of the matrix measure for Y_M under permutations of the matrix dimensions, N_m, was shown. This invariance carries over to the joint probability density function of the singular values as we will see.

The parametrization (3) follows directly from a parametrization of each individual matrix,

$$\mathbb{X}_m = \mathcal{U}_m \begin{pmatrix} X_m & A_m \\ 0 & B_m \end{pmatrix} \mathcal{U}_{m-1}^{-1}, \tag{4}$$

where $\mathcal{U}_0 = \mathbb{1}_{N_0}$. The matrices X_m, A_m, and B_m have the dimensions $N_0 \times N_0$, $N_0 \times (N_{m-1} - N_0)$ and $(N_m - N_0) \times (N_{m-1} - N_0)$, respectively. The entries of these matrices are real for $\beta = 1$, complex for $\beta = 2$, and quaternion for $\beta = 4$. Accordingly, we have

$$\mathcal{U}_m \in \begin{cases} \mathrm{O}(N_m)/[\mathrm{O}(N_0) \times \mathrm{O}(N_m - N_0)], \\ \mathrm{U}(N_m)/[\mathrm{U}(N_0) \times \mathrm{U}(N_m - N_0)], \\ \mathrm{USp}(2N_m)/\{\mathrm{USp}(2N_0) \times \mathrm{USp}[2(N_m - N_0)]\}, \end{cases} \tag{5}$$

for $\beta = 1,2,4$, respectively. The nonzero singular values of the rectangular product matrix (1) are identical to the singular values of the square product matrix $Y_M = X_M X_{M-1} \cdots X_1$ with Y_M and X_m, $m = 1, \ldots, M$, defined above. For this reason, we can safely replace the random matrix model containing rectangular matrices with a random matrix model containing square matrices, only. In terms of the new variables we get for the partition function, in analogy to [32] for $M = 1$,

$$Z_\beta^M \propto \prod_{m=1}^{M} \int |DX_m| \det{}^{\beta \nu_m/2}(X_m X_m^\dagger) \exp[-\operatorname{Tr} X_m X_m^\dagger], \tag{6}$$

where $\nu_m \equiv N_m - N_0 \geqslant 0$. A more general version of this result will be derived in [24]. In the partition function (6) and in most of this section we neglect an overall normalization constant, which is irrelevant for the computations. We reintroduce the normalization in Eq. (16) and give the explicit value in Eq. (21).

The Gaussian weight times a determinantal prefactor is sometimes referred to as the induced weight. For $M = 1$ its complex eigenvalues have been studied in [32].

In order to derive the joint probability density function for the singular values of the product matrix Y_M and thereby of Eq. (1), we follow the idea in [23], and reformulate the partition function (6) in terms of the product matrices $Y_m = X_m Y_{m-1} = X_m X_{m-1} \cdots X_1$, for $m = 1, \ldots, M$. In the following we assume that the product matrices, Y_m, are invertible (note that this restriction only removes a set of measure zero). We then know that [23]

$$\prod_{m=1}^{M} |DX_m| = |DY_1| \prod_{m=2}^{M} |DY_m| \det{}^{-\beta N_0/2}(Y_{m-1} Y_{m-1}^\dagger). \tag{7}$$

Changing variables from X_m to Y_m in the partition function Eq. (6) results in

$$\begin{aligned} Z_\beta^M \propto & \left[\prod_{m=1}^{M} \int |DY_m| \right] \det{}^{\beta \nu_M/2}(Y_M Y_M^\dagger) \exp[-\operatorname{Tr} Y_1 Y_1^\dagger] \\ & \times \left\{ \prod_{i=2}^{M} \det{}^{\beta(\nu_{i-1} - \nu_i - N_0)/2}(Y_{i-1} Y_{i-1}^\dagger) \right. \\ & \left. \times \exp[-\operatorname{Tr} Y_i Y_i^\dagger (Y_{i-1} Y_{i-1}^\dagger)^{-1}] \right\}. \end{aligned} \tag{8}$$

With this expression for the partition function we can express everything in terms of the singular values and a family of unitary matrices. We employ for each matrix Y_i a singular value decomposition [23] to write the product matrices as

$$Y_i = U_i \Sigma_i V_i^{-1}, \tag{9}$$

where $\Sigma_i = \operatorname{diag}\{\sigma_1^i, \sigma_2^i, \ldots, \sigma_{N_0}^i\}$ are positive definite diagonal matrices; the diagonal elements are the singular values of Y_i (for $\beta = 4$ the singular values show Kramer's degeneracy).

The unitary matrices, U_i and V_i, belong to

$$U_i \in \begin{cases} \mathrm{O}(N_0), \\ \mathrm{U}(N_0), \\ \mathrm{USp}(2N_0), \end{cases} \qquad V_i \in \begin{cases} \mathrm{O}(N_0), \\ \mathrm{U}(N_0)/\,\mathrm{U}(1)^{N_0}, \\ \mathrm{USp}(2N_0)/\,\mathrm{U}(1)^{N_0}, \end{cases} \tag{10}$$

for $\beta = 1,2,4$, respectively. It is well known that this change of variables yields the new measure

$$|DY_i| = |DU_i||DV_i| \prod_{k=1}^{N_0} d\sigma_k^i \left(\sigma_k^i\right)^{\beta-1} \left|\Delta_{N_0}[(\sigma^i)^2]\right|^\beta, \tag{11}$$

where $|DU_i|$ and $|DV_i|$ are the Haar measures for their corresponding groups and

$$\Delta_N(x) = \prod_{1\leqslant a<b\leqslant N} (x_a - x_b) = \det_{1\leqslant a,b\leqslant N} \left[x_a^{N-b}\right] \tag{12}$$

denotes the Vandermonde determinant. Inserting this parametrization into the partition function (8) and performing the shift $U_{\ell-1}^{-1} U_\ell \to U_\ell$ for $\ell = 2,\ldots,M$, we obtain

$$\begin{aligned} Z_\beta^M \propto & \left\{ \prod_{k=1}^{N_0} \left[\prod_{m=1}^{M} \int_0^\infty d\sigma_k^m \right] \left(\sigma_k^M\right)^{\beta(\nu_M+1)-1} e^{-(\sigma_k^1)^2} \right. \\ & \left. \times \prod_{i=2}^{M} \left(\sigma_k^{i-1}\right)^{\beta(\nu_{i-1}-\nu_i-N_0+1)-1} \right\} \prod_{j=1}^{M} \left|\Delta_{N_0}[(\sigma^j)^2]\right|^\beta \\ & \times \prod_{\ell=2}^{M} \int |DU_\ell||DV_\ell| \exp\left[-\operatorname{Tr} U_\ell \Sigma_\ell^2 U_\ell^{-1} \Sigma_{\ell-1}^{-2}\right]. \end{aligned} \tag{13}$$

The integrations over V_ℓ are trivial and only contribute to the normalization constant; the integration over U_ℓ is, however, more complicated. For $\beta = 2$, the integrals over U_ℓ are Harish-Chandra-Itzykson-Zuber integrals [30,31], while the integrals for $\beta = 1$ and $\beta = 4$ are still unknown in closed form. For this reason, we restrict ourselves to the complex case ($\beta = 2$), where we can carry out all integrals explicitly, and obtain an analytical expression for the joint probability density function. Recall that the complex ($\beta = 2$) product matrix is exactly the channel matrix used in wireless telecommunication to model MIMO channels with multiple scattering.

With the restriction to the $\beta = 2$ case, U_ℓ should be integrated over the unitary group, which yields [30,31]

$$\begin{aligned} & \int_{\mathrm{U}(N_0)} |DU_\ell| \exp\left[-\operatorname{Tr} U_\ell \Sigma_\ell^2 U_\ell^{-1} \Sigma_{\ell-1}^{-2}\right] \\ & \quad \propto \frac{\prod_{k=1}^{N_0} \left(\sigma_k^{\ell-1}\right)^{2(N_0-1)}}{\Delta_{N_0}[(\sigma^\ell)^2]\Delta_{N_0}[(\sigma^{\ell-1})^2]} \det_{1\leqslant a,b\leqslant N_0} \left[e^{-(\sigma_a^\ell)^2/(\sigma_b^{\ell-1})^2}\right], \end{aligned} \tag{14}$$

for $\ell = 2,\ldots,M$. Inserting this into the partition function (13) with $\beta = 2$ gives an expression for the partition function solely in terms of the singular values of the product matrices Y_i,

$$\begin{aligned} Z^M \equiv Z_{\beta=2}^M \propto & \left[\prod_{k=1}^{N_0} \int_0^\infty d\sigma_k^M \left(\sigma_k^M\right)^{2\nu_M+1} \right] \Delta_{N_0}[(\sigma^M)^2] \\ & \times \left\{ \prod_{i=1}^{M-1} \left[\prod_{\ell=1}^{N_0} \int_0^\infty d\sigma_\ell^i \left(\sigma_\ell^i\right)^{2(\nu_i-\nu_{i+1})-1} \right] \right. \\ & \left. \times \det_{1\leqslant a,b\leqslant N_0} \left[e^{-(\sigma_a^{i+1})^2/(\sigma_b^i)^2}\right] \right\} \left[\prod_{k=1}^{N_0} e^{-(\sigma_k^1)^2} \right] \Delta_{N_0}[(\sigma^1)^2]. \end{aligned} \tag{15}$$

For notational simplicity we change variables from the singular values to $s_a^i = (\sigma_a^i)^2$, i.e. the singular values (and eigenvalues) of the Wishart matrices $Y_i Y_i^\dagger$ (the singular values of $Y_M Y_M^\dagger$ are simply denoted by $s_a = s_a^M$). Furthermore, due to symmetrization we can replace the determinants of the exponentials with their diagonals, which will only change the partition function by a factor $(N_0!)^{M-1}$. Exploiting this, the partition function becomes

$$\begin{aligned} Z^M = & C_M^{-1} \left[\prod_{b=1}^{N_0} \int_0^\infty ds_b \, (s_b)^{\nu_M} \right] \Delta_{N_0}(s) \\ & \times \left\{ \prod_{a=1}^{N_0} \left[\prod_{i=1}^{M-1} \int_0^\infty \frac{ds_a^i}{s_a^i} \left(s_a^i\right)^{\nu_i-\nu_{i+1}} e^{-s_a^{i+1}/s_a^i} \right] e^{-s_a^1} \right\} \\ & \times \Delta_{N_0}(s^1), \end{aligned} \tag{16}$$

where C_M is a normalization constant.

The integrations over $s_a^1,\ldots,s_a^{M-1}$ have a similar structure. Hence, we can perform all these integrals in a similar fashion. We write the first exponential containing s_a^1 as a Meijer G-function using Eq. (A10), i.e.

$$\Delta_{N_0}(s^1) \prod_{a=1}^{N_0} e^{-s_a^1} = \det_{1\leqslant a,b\leqslant N_0} \left[G_{0,\,1}^{1,\,0}\left(\begin{matrix} - \\ b-1 \end{matrix} \middle| s_a^1 \right) \right]. \tag{17}$$

After a change of variables all the integrals can be performed inductively using the identities (A7) and (A5). These integrations finally give the joint probability density function, $\mathcal{P}_{\mathrm{jpdf}}$, for the singular values $s_1,\ldots,s_{N_0}$ of the Wishart matrix $Y_M Y_M^\dagger$,

$$\begin{aligned} \mathcal{P}_{\mathrm{jpdf}}^M(s_1,\ldots,s_{N_0}) = & C_M^{-1} \Delta_{N_0}(s) \\ & \times \det_{1\leqslant a,b\leqslant N_0} \left[G_{0,\,M}^{M,\,0}\left(\begin{matrix} - \\ \nu_M,\, \nu_{M-1},\, \ldots,\, \nu_2,\, \nu_1+b-1 \end{matrix} \middle| s_a \right) \right]. \end{aligned} \tag{18}$$

The partition function is thus given by

$$Z^M = \prod_{a=1}^{N_0} \int_0^\infty ds_a \, \mathcal{P}_{\mathrm{jpdf}}^M(s_1,\ldots,s_{N_0}). \tag{19}$$

This generalizes the joint probability density function for the product of square matrices from the Wishart ensemble given in [23] to the case of rectangular matrices. In principle, all k-point correlation functions for the singular values, $R_k^M(s_1,\ldots,s_k)$, can be calculated from the joint probability

density function (18) as

$$R_k^M(s_1,\ldots,s_k) = \frac{N_0!}{(N_0-k)!}\prod_{a=k+1}^{N_0}\int_0^\infty ds_a\,\mathcal{P}^M_{\text{jpdf}}(s_1,\ldots,s_{N_0}). \quad (20)$$

Due to the Meijer G-function inside the determinant (18) this is a nontrivial computation for $M \geqslant 2$. In complete analogy to the square case [23], it turns out that the correlation functions are more easily obtained using a two-matrix model and the method of biorthogonal polynomials. We discuss this in Sec. III, including other methods of derivation.

The normalization constant in Eqs. (15) and (18) is

$$C_M = N_0!\prod_{n=1}^{N_0}\prod_{m=0}^{M}\Gamma[n+\nu_m], \quad (21)$$

such that the partition function is equal to unity, which is straightforward to check using the Andréief integration formula. The one-point correlation function (or density) is normalized to the number of singular values,

$$\int_0^\infty ds\,R_1^M(s) = N_0, \quad (22)$$

which becomes evident in the following section.

III. TWO-MATRIX MODEL AND BIORTHOGONAL POLYNOMIALS

The purpose of this section is to find an explicit expression for the k-point correlation functions (20). We follow the idea in [23] and rewrite our problem as a two-matrix model by keeping the integrals over the s_a^1's and s_a^M's in Eq. (16) while integrating over the remaining variables. Within this model we exploit the method of biorthogonal polynomials to achieve our goal. First, we use the identity (A5) for the Meijer G-function to write the partition function (19) with $M \geqslant 2$ as

$$Z^M = \prod_{a=1}^{N_0}\int_0^\infty ds_a\prod_{i=1}^{N_0}\int_0^\infty dt_i\,\widetilde{P}^M_{\text{jpdf}}(s\,;t\,), \quad (23)$$

where the joint probability density function is given by

$$\widetilde{P}^M_{\text{jpdf}}(s\,;t\,) = C_M\Delta_{N_0}(s)\Delta_{N_0}(t)\det_{1\leqslant k,\ell,\leqslant N_0}\left[w_\nu^M(s_k,t_\ell)\right], \quad (24)$$

where $s_a \equiv s_a^M$ and $t_a \equiv s_a^1$, and the weight function depending on all indices ν_m collectively denoted by ν reads

$$w_\nu^M(s,t) = t^{\nu_1-1}e^{-t}G^{M-1,0}_{0,M-1}\left(\begin{matrix}-\\ \nu_M,\,\nu_{M-1},\,\ldots,\,\nu_2\end{matrix}\middle|\,\frac{s}{t}\right). \quad (25)$$

The structure of the joint probability density function (24) is similar to that of the two-matrix model discussed in [33]. Although the focus in [33] is on a multimatrix model with an Itzykson-Zuber interaction, the argument given is completely general and applies to our situation as well. The (k,ℓ)-point correlation functions for this two-matrix model are defined as

$$R^M_{k,\ell}(s\,;t\,) = \frac{(N_0!)^2}{(N_0-k)!(N_0-\ell)!}\times\prod_{a=k+1}^{N_0}\int_0^\infty ds_a\prod_{i=\ell+1}^{N_0}\int_0^\infty dt_i\,\widetilde{P}^M_{\text{jpdf}}(s;t). \quad (26)$$

Obviously, we can obtain the k-point correlation functions (20) by integrating out all t_i's, i.e. setting $\ell = 0$.

The benefit of the two-matrix model is that we can exploit the method of biorthogonal polynomials as in [33]. We choose a family of monic polynomials $q_j^M(t) = t^j + \cdots$ and $p_j^M(s) = s^j + \cdots$, which are biorthogonal with respect to the weight (25),

$$\int_0^\infty ds\int_0^\infty dt\,w_\nu^M(s,t)q_i^M(t)p_j^M(s) = h_j^M\delta_{ij}, \quad (27)$$

where h_j^M are constants. Furthermore, we introduce the functions $\psi_j^M(t)$ and $\varphi_j^M(s)$ defined as integral transforms of the biorthogonal polynomials,

$$\psi_j^M(t) \equiv \int_0^\infty ds\,w_\nu^M(s,t)p_j^M(s), \quad (28)$$

$$\varphi_j^M(s) \equiv \int_0^\infty dt\,w_\nu^M(s,t)q_j^M(t). \quad (29)$$

Note that $\psi_j^M(t)$ and $\varphi_j^M(s)$ are not necessarily polynomials. It is evident from the biorthogonality of the polynomials (27) that we have the orthogonality relations

$$\int_0^\infty dt\,q_i^M(t)\psi_j^M(t) = \int_0^\infty ds\,p_i^M(s)\varphi_j^M(s) = h_j^M\delta_{ij}. \quad (30)$$

Moreover, it follows from the discussion in [33] that the (k,ℓ)-point correlation functions are given by a determinantal point process,

$$R^M_{k,\ell}(s\,;t\,) = \det_{\substack{1\leqslant a,b\leqslant k\\ 1\leqslant i,j\leqslant\ell}}\begin{bmatrix}K_{11}^M(s_a,s_b) & K_{12}^M(s_a,t_j)\\ K_{21}^M(t_i,s_b) & K_{22}^M(t_i,t_j)\end{bmatrix}, \quad (31)$$

where the four subkernels are defined in terms of the biorthogonal polynomials and the weight function as

$$\begin{aligned}
K_{11}^M(s_a,s_b) &= \sum_{n=0}^{N_0-1}\frac{p_n^M(s_a)\varphi_n^M(s_b)}{h_n^M},\\
K_{12}^M(s_a,t_j) &= \sum_{n=0}^{N_0-1}\frac{p_n^M(s_a)q_n^M(t_j)}{h_n^M},\\
K_{21}^M(t_i,s_b) &= \sum_{n=0}^{N_0-1}\frac{\psi_n^M(t_i)\varphi_n^M(s_b)}{h_n^M} - w_\nu^M(s_b,t_i),\\
K_{22}^M(t_i,t_j) &= \sum_{n=0}^{N_0-1}\frac{\psi_n^M(t_i)q_n^M(t_j)}{h_n^M}.
\end{aligned} \quad (32)$$

In particular we have that the k-point correlation functions (20) for the singular values of the product matrix $Y_M Y_M^\dagger$ are given by

$$R_k^M(s_1,\ldots,s_k) = \det_{1\leqslant a,b\leqslant k}\left[K_{11}^M(s_a,s_b)\right]. \quad (33)$$

The goal is to find the biorthogonal polynomials, $q_j^M(t)$ and $p_j^M(s)$, and the norms, h_j^M, and thereby all correlation functions for the singular values of the product matrix, $\mathbb{Y}_M$. Note that we use a slightly different notation for the subkernels than in [23]; the notation in this paper is chosen to emphasize the fact that all the statistical properties of the singular values are determined

by the biorthogonal polynomials, $q_j^M(t)$ and $p_j^M(s)$, and the weight function, $w_v^M(s,t)$.

In order to find the biorthogonal polynomials we follow the approach in [23] and start by computing the bimoments,

$$I_{ij}^M \equiv \int_0^\infty ds \int_0^\infty dt\, w_v^M(s,t) s^i\, t^j = (i+j+\nu_1)! \prod_{m=2}^{M} (i+\nu_m)!, \tag{34}$$

for $M \geqslant 2$. Here the integration has been performed using integral identities for the Meijer G-function, see Eqs. (A4) and (A5). Using Cramer's rule, the biorthogonal polynomials, as well as the norms, can be expressed in terms of the bimoments as [34,35],

$$q_n^M(t) = \frac{1}{D_{n-1}^M} \det \begin{bmatrix} I_{00}^M & I_{10}^M & \cdots & I_{(n-1)0}^M & 1 \\ I_{01}^M & I_{11}^M & \cdots & I_{(n-1)1}^M & t \\ \vdots & \vdots & & \vdots & \vdots \\ I_{0n}^M & I_{1n}^M & \cdots & I_{(n-1)n}^M & t^n \end{bmatrix},$$
$$p_n^M(s) = \frac{1}{D_{n-1}^M} \det \begin{bmatrix} I_{00}^M & I_{01}^M & \cdots & I_{0(n-1)}^M & 1 \\ I_{10}^M & I_{11}^M & \cdots & I_{1(n-1)}^M & s \\ \vdots & \vdots & & \vdots & \vdots \\ I_{n0}^M & I_{n1}^M & \cdots & I_{n(n-1)}^M & s^n \end{bmatrix}, \tag{35}$$

where

$$D_n^M \equiv \det_{0\leqslant i,j\leqslant n} \left[I_{ij}^M\right] = \prod_{i=0}^{n} \prod_{m=0}^{M} (i+\nu_m)!. \tag{36}$$

The norms can be expressed as

$$h_n^M = D_n^M / D_{n-1}^M = \prod_{m=0}^{M} (n+\nu_m)!. \tag{37}$$

Recall that $\nu_i \equiv N_i - N_0 \geqslant 0$ are non-negative integers by definition ($\nu_0 = 0$).

In order to get more explicit expressions for the biorthogonal polynomials, we define the bimoment matrix (34) for $M = 1$ as the bimoments with respect to the Laguerre weight,

$$I_{ij}^{M=1} \equiv \int_0^\infty ds\, e^{-s} s^{\nu_1+i+j} = (i+j+\nu_1)!. \tag{38}$$

It follows that the polynomials (35) for $M = 1$ are the Laguerre polynomials in monic normalization,

$$p_n^{M=1}(s) = q_n^{M=1}(s) = \widetilde{L}_n^{\nu_1}(s) \equiv (-1)^n n! L_n^{\nu_1}(s), \tag{39}$$

where $L_n^{\nu_1}(s)$ are the associated Laguerre polynomials. We recall that the Laguerre polynomials are defined as

$$\widetilde{L}_n^{\nu_1}(s) = \sum_{k=0}^{n} \frac{(-1)^{n+k}}{(n-k)!} \frac{(n+\nu_1)!}{(k+\nu_1)!} \frac{n!}{k!} s^k \tag{40}$$

and satisfy the orthogonality relation

$$\int_0^\infty ds\, e^{-s} s^{\nu_1} \widetilde{L}_k^{\nu_1}(s) \widetilde{L}_\ell^{\nu_1}(s) = h_k^{M=1} \delta_{k\ell}, \tag{41}$$

with $h_k^{M=1} = k!(k+\nu_1)!$.

The bimoment matrix, $[I_{ij}^M]_{0\leqslant i,j\leqslant n}$, with $M \geqslant 2$ given by Eq. (34) differs from the bimoment matrix, $[I_{ij}^1]_{0\leqslant i,j\leqslant n}$, given by Eq. (38) by multiplication of a diagonal matrix. It directly follows from this fact that the polynomials $q_n^M(t)$ are related to the Laguerre polynomials as

$$q_n^M(t) = \prod_{i=0}^{n-1} \prod_{m=2}^{M} (i+\nu_m)! \frac{D_{n-1}^1}{D_{n-1}^M} \widetilde{L}_n^{\nu_1}(t) = \widetilde{L}_n^{\nu_1}(t). \tag{42}$$

The evaluation of the polynomials $p_n^M(s)$ is slightly more complicated. For the polynomials $q_n^M(t)$, the factorization is the same for all powers of t, but for the polynomials $p_n^M(s)$ we have to treat the powers differently; in particular we substitute $s^k \to s^k / \prod_{m=2}^M (k+\nu_m)!$. Using the explicit expression for the Laguerre polynomials (40) we find

$$p_n^M(s) = \sum_{k=0}^{n} \frac{(-1)^{n+k} n!}{(n-k)!} \left[\prod_{m=1}^{M} \frac{(n+\nu_m)!}{(k+\nu_m)!}\right] \frac{s^k}{k!}, \tag{43}$$

which is a generalized hypergeometric polynomial [see Eq. (A2) in the Appendix)

$$p_n^M(s) = (-1)^n \prod_{m=1}^{M} \frac{(n+\nu_m)!}{\nu_m!} \times {}_1F_M\left(\begin{matrix} -n \\ 1+\nu_M, \ldots, 1+\nu_1 \end{matrix} \middle| s\right). \tag{44}$$

For $\nu_M = \cdots = \nu_1 = 0$ this polynomial reduces to the result presented in [23], while the monic Laguerre polynomials are reobtained by setting $M = 1$. Alternatively we may write $p_n^M(s)$ as a Meijer G-function,

$$p_n^M(s) = (-1)^n \prod_{m=0}^{M} (n+\nu_m)! \times G_{1,M+1}^{1,0}\left(\begin{matrix} n+1 \\ 0, -\nu_M, \ldots, -\nu_1 \end{matrix} \middle| s\right). \tag{45}$$

This expression will be particularly useful in Sec. IV, where we discuss the asymptotic behavior of the end points of support of the spectral density. In Eq. (45) we have used the relation (A9) between generalized hypergeometric polynomials and Meijer G-functions. It might not be immediately clear that the Meijer G-function in Eq. (45) is a polynomial. To see this, one writes the Meijer G-function as a contour integral using its definition (A3). The integrand has exactly n simple poles and the contour is closed such that these poles are encircled. The residue for each pole gives a monomial, such that the complete contour integral yields a polynomial.

With the explicit expressions for the biorthogonal polynomials (42) and (44), we are ready to compute the functions $\psi_n^M(t)$ and $\varphi_n^M(s)$ defined in Eq. (29) and thereby implicitly find all the subkernels (32). The functions $\psi_n^M(t)$ turn out to be polynomials, too,

$$\psi_n^M(t) = \prod_{m=2}^{M} (n+\nu_m)!\, t \widetilde{L}_n^{\nu_1}(t), \tag{46}$$

which can be directly obtained from the definition (29) using the integral identity (A4).

Likewise, we can obtain an explicit expression for the functions $\varphi_n^M(s)$ by inserting the polynomial (42) into the definition (29). It follows from the integral identity (A5) that

$$\varphi_n^M(s) = \sum_{k=0}^{n} \frac{(-1)^{n+k}}{(n-k)!} \frac{(n+\nu_1)!}{(k+\nu_1)!} \frac{n!}{k!} \times G_{0,\,M}^{M,\,0}\left(\begin{matrix} - \\ \nu_M, \ldots, \nu_2, \nu_1 + k \end{matrix} \middle| s \right). \tag{47}$$

However, it is possible to get a more compact expression. Recall that the Laguerre polynomials can be expressed using Rodrigues' formula,

$$\widetilde{L}_n^{\nu_1}(t) = (-1)^n t^{-\nu_1} e^t \frac{d^n}{dt^n}(t^{n+\nu_1} e^{-t}). \tag{48}$$

We insert Rodrigues' formula into the definition for $\varphi_n^M(s)$; see Eq. (29). The differentiation in Eq. (48) can easily be changed to a differentiation of the Meijer G-function (stemming from the weight function) using integration by parts, since all boundary terms are zero. Then the differentiation can be computed using Eq. (A8), while the final integration over t can be performed using the identity (A5). This finally leads to

$$\varphi_n^M(s) = (-1)^n G_{1,\,M+1}^{M,\,1}\left(\begin{matrix} -n \\ \nu_M, \nu_{M-1}, \ldots, \nu_1, 0 \end{matrix} \middle| s \right). \tag{49}$$

In addition to the fact that Eq. (49) is a more compact expression than the representation (47), it is also immediate that $\varphi_n^M(s)$ is symmetric in all the indices ν_m, which is far from obvious in Eq. (47).

Now we have explicit expressions for all components contained in the formula for the (k,ℓ)-point correlation functions (31), which completes the derivation. In particular, combining Eqs. (37), (44), and (49) the subkernel $K_{11}^M(s_a,s_b)$ is given by

$$K_{11}^M(s_a,s_b) = \sum_{n=0}^{N_0-1} \frac{1}{n!} \prod_{m=1}^{M} \frac{1}{\nu_m!}\, {}_1F_M\left(\begin{matrix} -n \\ 1+\nu_M, \ldots, 1+\nu_1 \end{matrix} \middle| s_a \right) \times G_{1,\,M+1}^{M,\,1}\left(\begin{matrix} -n \\ \nu_M, \ldots, \nu_1, 0 \end{matrix} \middle| s_b \right). \tag{50}$$

It provides a direct generalization of the formula given in [23] for square matrices to the case of rectangular matrices. If we use the alternative formula (45) for $p_n^M(s)$ we obtain

$$K_{11}^M(s_a,s_b) = \sum_{n=0}^{N_0-1} G_{1,\,M+1}^{1,\,0}\left(\begin{matrix} n+1 \\ 0, -\nu_M, \ldots, -\nu_1 \end{matrix} \middle| s_a \right) \times G_{1,\,M+1}^{M,\,1}\left(\begin{matrix} -n \\ \nu_M, \ldots, \nu_1, 0 \end{matrix} \middle| s_b \right). \tag{51}$$

The k-point correlation functions for the singular values are immediately found from Eq. (33). Note that the kernel and thereby all k-point correlation functions are symmetric in all the indices ν_m. This symmetry reflects the invariance of the singular values of the product matrix, $Y_M = X_M \cdots X_1$, under reordering of the matrices X_m which we prove in a more general setting in [24]. The normalization of the spectral density (22) is immediately clear from the orthogonality relation (30).

Finally, we would like to mention an alternative derivation for the correlation functions (20) in terms of the kernel K_{11}^M. Given the orthogonality relation (30) of the polynomials p_i^M (43) and the functions φ_j^M (47) we can generate these by adding columns in the two determinants in the joint probability density function (18) and then proceed with the standard Dyson theorem. This is in complete analogy as described in [23]. Alternatively, the kernel can be derived by using biorthogonal functions and explicitly inverting the bimoment matrix [36]. Furthermore, a construction using multiple orthogonal polynomials exist [37,38], too.

IV. MOMENTS AND ASYMPTOTICS

In this section we take a closer look at the spectral density. First we use the density to find an explicit expression for the moments. Second we discuss the macroscopic large-N_0 limit of the density.

We know from the previous section that the density, or one-point correlation function, is given as a sum over Meijer G-functions,

$$R_1^M(s) = \sum_{n=0}^{N_0-1} G_{1,\,M+1}^{1,\,0}\left(\begin{matrix} n+1 \\ 0, -\nu_M, \ldots, -\nu_1 \end{matrix} \middle| s \right) \times G_{1,\,M+1}^{M,\,1}\left(\begin{matrix} -n \\ \nu_M, \ldots, \nu_1, 0 \end{matrix} \middle| s \right), \tag{52}$$

which is normalized to the number of singular values, N_0. Figure 1 shows a comparison between the analytical expression and numerical simulations for an example. The expectation value for the singular values is defined in terms of the density (52) as

$$\mathbb{E}\{f(s)\} \equiv \frac{1}{N_0} \int_0^\infty ds\, R_1^M(s)\, f(s), \tag{53}$$

where the factor $1/N_0$ is included since the density (52) is normalized to the number of singular values.

We first look at the moments, $\mathbb{E}\{s^\ell\}$. Note that we do not assume that ℓ is an integer and that the half-integer values of ℓ are interesting, too, since the singular values, σ_a, of the product matrix, $\mathbb{Y}_M$, are given by the square roots of the eigenvalues of the Wishart matrix; i.e. $\sigma_a = \sqrt{s_a}$. In order to calculate the moments, we explicitly write the first Meijer G-function in Eq. (52) as a polynomial [see Eqs. (43) and (45)] and rewrite the moments as

$$N_0 \mathbb{E}\{s^\ell\} = \sum_{n=0}^{N_0-1} \sum_{k=0}^{n} \frac{(-1)^k}{(n-k)!} \prod_{m=0}^{M} \frac{1}{(k+\nu_m)!} \times \int_0^\infty ds\, s^{\ell+k} G_{1,\,M+1}^{M,\,1}\left(\begin{matrix} -n \\ \nu_M, \ldots, \nu_1, 0 \end{matrix} \middle| s \right). \tag{54}$$

The integral over s can be performed using an identity for the Meijer G-function (A4). After reordering the sums and applying Euler's reflection formula for the gamma function we get

$$N_0 \mathbb{E}\{s^\ell\} = \sum_{k=0}^{N_0-1} \prod_{m=0}^{M} \frac{\Gamma[\ell+k+\nu_m+1]}{(k+\nu_m)!} \times \sum_{n=0}^{N_0-k-1} \frac{(-1)^n}{n!\,\Gamma[\ell-n+1]}, \tag{55}$$

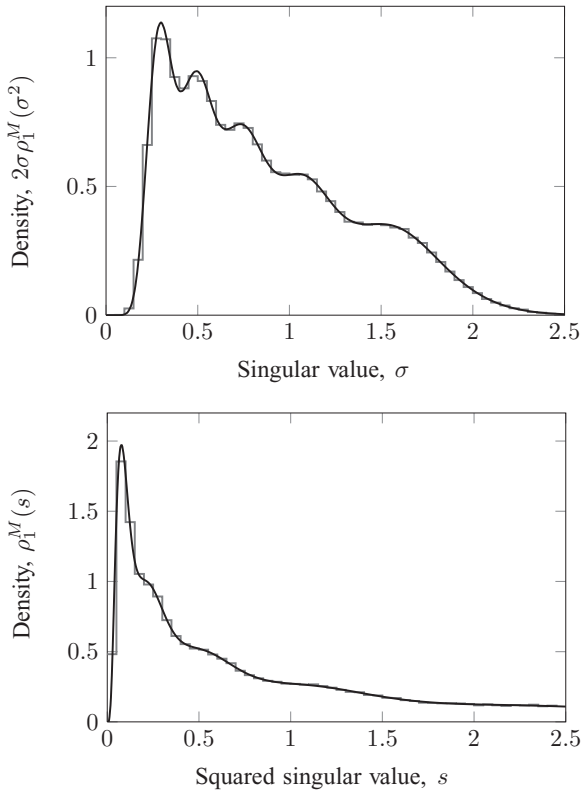

FIG. 1. The histograms (bin width is 0.05) show the distributions of singular values (top) and squared singular values (bottom) for 50 000 realizations of the product matrix $\mathbb{Y}_3 = \mathbb{X}_3\mathbb{X}_2\mathbb{X}_1$ for $M = 3$, with $\nu_1 = 5$, $\nu_2 = 10$, $\nu_3 = 15$, and $N_0 = 5$. The solid curves are the analytical predictions for the rescaled densities of singular values, $2\sigma\rho_1^3(\hat{\sigma}^2)$, and of squared singular values, $\rho_1^3(\hat{s})$, respectively; cf. Eq. (61).

where ℓ may also take noninteger values. For integer values of ℓ some of the terms vanish due to the poles of the gamma function. Note that the moments are divergent whenever $\ell \leqslant -\nu_{\min} - 1$ is an integer ($\nu_{\min} \equiv \min\{\nu_1, \ldots, \nu_M\}$), but well defined for all other values of ℓ. The second sum in Eq. (55) can be evaluated by a relation for the (generalized) binomial series,

$$\sum_{n=0}^{N}(-1)^n\binom{z}{n} = (-1)^N\binom{z-1}{N}, \quad z \in \mathbb{C}. \tag{56}$$

We write the first sum in Eq. (55) in reverse order ($k \to N_0 - k - 1$) and perform the second sum using the identity (56), yielding

$$N_0\mathbb{E}\{s^\ell\} = \sum_{k=0}^{N_0-1}\frac{(-1)^k}{k!\Gamma[\ell-k]\ell}\prod_{m=0}^{M}\frac{\Gamma[\ell+N_m-k]}{\Gamma[N_m-k]}. \tag{57}$$

Alternatively, the moments can be written as

$$N_0\mathbb{E}\{s^\ell\} = \sum_{k=0}^{N_0-1}\frac{(-1)^{1+k}\prod_{j=0}^{N_0-1}(j-\ell-k)}{k!(N_0-1-k)!\ell} \times \prod_{m=1}^{M}\frac{\Gamma[\ell+\nu_m+k+1]}{\Gamma[\nu_m+k+1]}, \tag{58}$$

which is useful when considering the limit of negative integer ℓ. Recall that N_m are the different matrix dimensions of the original product (1) and $\nu_m = N_m - N_0$.

For $\ell \to 0$ all terms in the sum are equal to one and we recover the normalization. Simplifications also occur when ℓ is an integer; here most of the terms in the sum vanish, due to the gamma function in the denominator. In particular, the first positive moment and the first negative moment are given by

$$\mathbb{E}\{s\} \equiv \mathcal{N}_M = \prod_{m=1}^{M}N_m \quad \text{and} \quad \mathbb{E}\{s^{-1}\} = \prod_{m=1}^{M}\frac{1}{\nu_m}. \tag{59}$$

The second moment is slightly more complicated,

$$\mathbb{E}\{s^2\} = \frac{1}{2}\prod_{m=1}^{M}N_m\left[\prod_{m=0}^{M}(N_m+1) - \prod_{m=0}^{M}(N_m-1)\right]. \tag{60}$$

When $M = 1$ these formulas reduce to the well-known results for the Wishart-Laguerre ensemble (e.g. see [11]), while we get the result [23] for square matrices by setting $N_0 = \cdots = N_M$. Note that any negative moment is divergent if $\nu_m = 0$ for any $1 \leqslant m \leqslant M$.

The first moment, $\mathcal{N}_M$, provides us with a natural scaling of the spectral density,

$$\rho_1^M(\hat{s}) \equiv \frac{\mathcal{N}_M}{N_0}R_1^M(\hat{s}\,\mathcal{N}_M), \tag{61}$$

such that the rescaled density has a finite first moment of unity also in the large-N_0 limit. In Eq. (61) and the following, we use a hat "$\hat{}$" to denote rescaled variables.

The expectation value with respect to the rescaled density (61) is related to the definition (53) by a simple scaling of the variable,

$$\hat{\mathbb{E}}\{f(\hat{s})\} \equiv \int_0^\infty d\hat{s}\,\rho_1^M(\hat{s})f(\hat{s}) = \mathbb{E}\left\{f\left(\frac{\hat{s}}{\mathcal{N}_M}\right)\right\}, \tag{62}$$

for any observable $f(\hat{s})$. The rescaling ensures that we have a well-defined probability density with compact support in the large-N_0 limit; in particular, the density $\rho_1^1(\hat{s})$ for a single matrix $M = 1$ reduces to the celebrated Marčenko-Pastur density for $N_0 \to \infty$.

An algebraic way to obtain the macroscopic behavior of the spectral density (61) for arbitrary M was provided in [14], using the resolvent also known as the Stieltjes transform, $G^M(\hat{z})$, defined as

$$G^M(\hat{z}) \equiv \int_0^\infty d\hat{s}\lim_{N_0\to\infty}\frac{\rho_1^M(\hat{s})}{\hat{z}-\hat{s}}, \tag{63}$$

with $\hat{z}$ outside the limiting support of ρ_1^M. It was shown that in the large-N_0 limit the resolvent satisfies a polynomial equation [14],

$$\hat{z}\,G^M(\hat{z})\prod_{m=1}^{M}\frac{\hat{z}\,G^M(\hat{z})+\hat{\nu}_m}{\hat{\nu}_m+1} = \hat{z}[\hat{z}\,G^M(\hat{z})-1], \tag{64}$$

where $\hat{z}$ lies outside the support of the singular values and $\hat{\nu}_m$ denotes the rescaled differences in matrix dimensions; i.e. $\hat{\nu}_m \equiv \nu_m/N_0$ for $m = 1, \ldots, M$. In general, one needs to solve an $(M+1)$st order equation in order to find the resolvent, $G^M(\hat{z})$. It is clear that such an equation can generically only

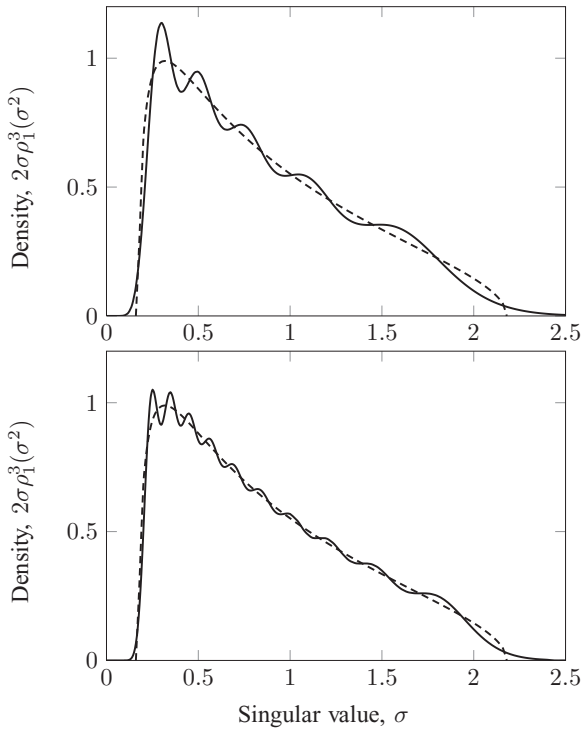

FIG. 2. The solid lines show the $M = 3$ rescaled spectral densities for the singular values for $N_0 = 5$ (top) and $N_0 = 10$ (bottom) both with $\hat{\nu}_1 = 1$, $\hat{\nu}_2 = 2$, $\hat{\nu}_3 = 3$. The dashed curves indicate the corresponding macroscopic limit [14].

be solved analytically for $M \leqslant 3$ (see also the discussions in [37,39]).

The correct resolvent is chosen by its asymptotic behavior, $\hat{z}G^M(\hat{z}) \to 1$ for $\hat{z} \to \infty$. When an expression for the resolvent is known, then the spectral density can be directly obtained from the resolvent using

$$\rho_1^{M,\infty}(\hat{s}) \equiv \lim_{N_0\to\infty} \rho_1^M(\hat{s}) = \frac{1}{\pi} \lim_{\varepsilon\to 0^+} \operatorname{Im} G^M(\hat{s} - \imath\varepsilon). \quad (65)$$

In Fig. 2 we compare this macroscopic limit with the rescaled density (52) at finite N_m.

For the case $M = 1$ one can readily derive the well-known Marčenko-Pastur law. Another particular case in which the spectral density $\rho_1^{M,\infty}$ can be directly calculated is $M = 2$ with $\hat{\nu}_1$ and $\hat{\nu}_2$ arbitrary. This case plays an important role when studying cross correlation matrices of two different sets of time series as it appears in forecasting models [10,40], where time-lagged correlation matrices are nonsymmetric. Our random matrix model then corresponds to the case of two time series which are uncorrelated. Despite the independence of the distribution of the matrix elements, correlations among the singular values of the cross correlation matrix follow. The solution of Eq. (64) yields the level density

$$\rho_1^{M,\infty}(\hat{s}) = \frac{\sqrt{3(\hat{\nu}_1+1)(\hat{\nu}_2+1)\hat{s} + \hat{\nu}_1^2 - \hat{\nu}_1\hat{\nu}_2 + \hat{\nu}_2^2}}{3\pi\hat{s}} \times \operatorname{Im}[A^{-1/3}(f(\hat{s})) + A^{1/3}(f(\hat{s}))], \quad (66)$$

with

$$f\left[\frac{z}{(\hat{\nu}_1+1)(\hat{\nu}_2+1)}\right] = 3\frac{\left[3z + \hat{\nu}_1^2 - \hat{\nu}_1\hat{\nu}_2 + \hat{\nu}_2^2\right]^3}{\left[3(3+\hat{\nu}_1+\hat{\nu}_2)z + \hat{\nu}_1^3 - (\hat{\nu}_1+\hat{\nu}_2)^3/3 + \hat{\nu}_2^3\right]^2} \quad (67)$$

and

$$A(z) = \sqrt{\frac{27}{4z} - 1} - \sqrt{\frac{27}{4z}}. \quad (68)$$

Indeed, the special case $\hat{\nu}_1 = \hat{\nu}_2 = 0$ agrees with the result derived in [23,26,37] because $f(\hat{s})|_{\hat{\nu}_1=\hat{\nu}_2=0} = \hat{s}$.

It is also desirable to know where the end points of support of the macroscopic spectrum are located. These edges can be found from the algebraic formula for the resolvent (64) using a simple trick. We assume that the resolvent behaves as $|G^M(\hat{z})| \sim |\hat{z} - \hat{s}_\pm|^{\alpha_\pm}$, with $\alpha_\pm < 1$ and $\alpha_\pm \neq 0$ in the vicinity of the edges, $\hat{s}_\pm$. This edge behavior of the resolvent is known to hold in certain cases; e.g. $M = 1$ yields $\alpha_\pm = 1/2 < 1$ (except when the inner edge is zero, $\hat{s}_- = 0$, then $\alpha_- = -1/2 < 1$). Due to known universality results for random matrices, it is expected that $\alpha_\pm < 1$ and $\alpha_\pm \neq 0$ in general. With this particular edge behavior, it is clear that $|dG^M/d\hat{z}| \to \infty$ for $\hat{z} \to \hat{s}_\pm$, or equivalently $d\hat{z}/dG^M \to 0$ for $\hat{z} \to \hat{s}_\pm$. Differentiating both sides of Eq. (64) with respect to G^M and evaluating them at $d\hat{z}/dG^M = 0$ yields an equation for the extrema of $\hat{z}$,

$$\hat{z}_0 = \left[1 + \sum_{j=1}^{M} \frac{\hat{z}_0 G^M(\hat{z}_0)}{\hat{z}_0 G^M(\hat{z}_0) + \hat{\nu}_j}\right] \prod_{m=1}^{M} \frac{\hat{z}_0 G^M(\hat{z}_0) + \hat{\nu}_m}{\hat{\nu}_m + 1}. \quad (69)$$

Two of these extrema are the inner edge, $\hat{z}_0 = \hat{s}_-$, and the outer edge, $\hat{z}_0 = \hat{s}_+$. The edges, $\hat{s}_\pm$, also satisfy Eq. (64). Combining both equations, we get an expression for the edges,

$$\hat{s}_\pm = \frac{\hat{u}_0}{1+\hat{u}_0} \prod_{m=1}^{M} \frac{\hat{\nu}_m - \hat{u}_0}{\hat{\nu}_m + 1}, \quad (70)$$

in terms of $\hat{u}_0 \equiv -\hat{z}_0 G^M(\hat{z}_0)$, which is given by

$$\sum_{m=1}^{M} \frac{\hat{u}_0(\hat{u}_0+1)}{\hat{\nu}_m - \hat{u}_0} = 1. \quad (71)$$

This equation is equivalent to a polynomial equation of $(M+1)$st order, as is the case for the resolvent; see Eq. (64). However, in certain cases Eq. (71) simplifies. In particular, Eq. (71) reduces to an Mth order equation if $\hat{\nu}_i = \hat{\nu}_j$ for $i \neq j$, if $\hat{\nu}_i \to 0$ or if $\hat{\nu}_i \to \infty$. The latter means that $N_i \gg N_0$, meaning that the matrix dimension N_i decouples from the macroscopic theory.

In general, the set of equations (70) and (71) yields $(M+1)$ solutions, of which two correspond to the inner edge and the outer edge of the spectral density. In the special case where $\hat{\nu} \equiv \hat{\nu}_1 = \cdots = \hat{\nu}_M$, there are only two solutions (see Fig. 3),

$$\hat{s}_\pm(\hat{\nu}) = \frac{M+1+2\hat{\nu} \pm \sqrt{(M+1)^2 + 4M\hat{\nu}}}{2(\hat{\nu}+1)} \times \left[\frac{M+1+2M\hat{\nu} \pm \sqrt{(M+1)^2+4M\hat{\nu}}}{2M+2M\hat{\nu}}\right]^M. \quad (72)$$

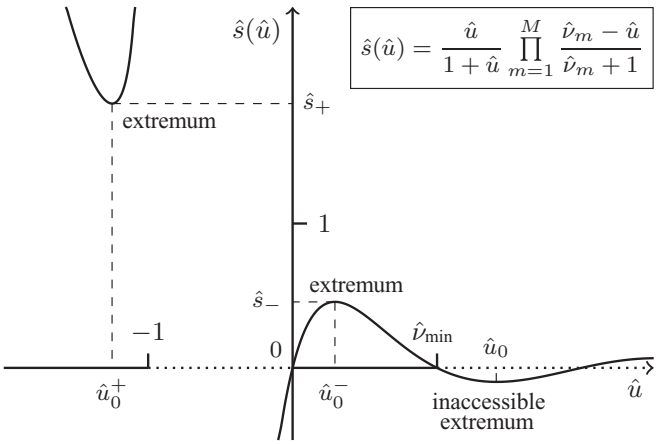

FIG. 3. Illustration of the optimization problem given by Eq. (76). Extrema within the intervals $(-1,0)$ and $(\hat{\nu}_{\min},\infty)$ must be disregarded due to the cuts in the complex $(-\hat{v})$-plane and complex $\hat{u}$-plane, respectively. This leaves only two valid extrema which correspond to the inner edge and the outer edge, respectively. Note that the solutions for the inner edge and the outer edge are separated by the pole at -1.

Note that for $M = 1$ this result reduces to the known values for the edges of the Marčenko-Pastur density (e.g. see [11]), while the limit $\hat{s}_\pm(\hat{v} \to 0)$ reproduces the result for the product of square matrices; see [23,26,37]. It is easy to numerically verify that the result holds in general.

Looking at Eqs. (70) and (71), an obvious question is as follows: Which solutions correspond to the edges of the spectrum? In order to answer this question, we derive the same equations through a different route. The rescaled spectral density (61) serves as the starting point, and the locations of the edges are determined using a saddle point approximation for large N_0. This also illustrates the point that the finite N_m expression discussed in this paper is equivalent to the result presented in [14] in the macroscopic limit.

In the large-N_0 limit we may approximate the sum over n [see Eq. (52)] by an integral. Moreover, we write the Meijer G-functions as contour integrals (A3) and approximate the gamma functions using Stirling's formula. The rescaled density (61) becomes

$$\rho_1^M(\hat{s}) \approx \frac{\mathcal{N}_M}{N_0} \int_0^1 d\hat{n}\, \frac{N_0}{2\pi\imath} \int_{L_1} d\hat{v}\, e^{-N_0 S(-\hat{v},\hat{n})} \times \frac{N_0}{2\pi\imath} \int_{L_2} d\hat{u}\, e^{N_0 S(\hat{u},\hat{n})}, \tag{73}$$

where the action, S, is given by

$$S(\hat{u},\hat{n}) = \hat{u} \ln \mathcal{N}_M \hat{s} + \sum_{m=1}^{M} (\hat{\nu}_m - \hat{u})[\ln N_0(\hat{\nu}_m - \hat{u}) - 1] + (\hat{n} + \hat{u})[\ln N_0(\hat{n} + \hat{u}) - 1] - \hat{u}\,(\ln N_0 \hat{u} - 1) \tag{74}$$

with $\hat{n} = n/N_0$, $\hat{u} = u/N_0$ and $\hat{\nu}_m = \nu_m/N_0$. It is important to note that the integrand in the definition of the Meijer G-function (A3) contains poles which lie on the real axis. The contours L_1 and L_2 encircle the poles of the original Meijer G-functions in accordance to definition (A3). In the large-N_0 limit these poles condense into cuts, such that the complex $\hat{u}$-plane has a cut on the interval $(\hat{\nu}_{\min},\infty)$ and the complex $(-\hat{v})$-plane has a cut on the interval $(-1,0)$. The contours L_1 and L_2 encircle these cuts in the $\hat{v}$-plane and the $\hat{u}$-plane, respectively. Both contour integrals can be evaluated by a saddle point approximation. Furthermore, variation with respect to $\hat{n}$ yields $\hat{u} = -\hat{v}$ at the saddle point, and due to the symmetry between the two saddle point equations we can restrict our attention to one of them. The saddle point equation for $\hat{u}$ yields

$$\hat{s} = \frac{\hat{u}_0}{\hat{n} + \hat{u}_0} \prod_{m=1}^{M} \frac{\hat{\nu}_m - \hat{u}_0}{\hat{\nu}_m + 1}, \quad 0 \leqslant \hat{n} \leqslant 1. \tag{75}$$

Equation (75) gives the saddle points, $\hat{u}_0$, for any given $\hat{s}$. In order to find the saddle points for the edges of the spectrum, we have to find the values of $\hat{n}$ and $\hat{u}_0$ which give the extremal values of $\hat{s}$.

Optimizing with respect to $\hat{n}$, we see that $\hat{n}$ has no optimal value within the interval $(0,1)$; hence, $\hat{n}$ must lie on the boundary due to the Laplace approximation (saddle point approximation on a real support). The only nontrivial result comes from $\hat{n} = 1$. Inserting this condition into the saddle point equation (75) we reproduce formula (70). The condition for $\hat{u}_0$ is given by differentiating the left hand side of the saddle point equation (75) and setting this result equal to zero,

$$\frac{d}{d\hat{u}_0} \left[\frac{\hat{u}_0}{1 + \hat{u}_0} \prod_{m=1}^{M} \frac{\hat{\nu}_m - \hat{u}_0}{\hat{\nu}_m + 1} \right] = 0. \tag{76}$$

This condition is identical to formula (71). Hence, the saddle point method reproduces the result obtained from the algebraic equation (64) for the resolvent.

The saddle points, which satisfy Eq. (76), are the extrema of the function within the square brackets. This function has a pole at -1 and goes to $+\infty$ for $\hat{u}_0 \to -\infty$ such that there is exactly one minimum to the left of the pole; see Fig. 3. On the right of the pole the function oscillates such that it has zeros at $0, \hat{\nu}_1, \ldots, \hat{\nu}_M$. Since the rational function on the right hand side of Eq. (75) is continuous, it has extrema between neighboring zeros [see Fig. 3], yielding M additional extrema. It follows that the optimization problem (76) has $M + 1$ solutions for $\hat{u}_0$, which are all real: One solution $\hat{u}_0^+ < -1$ which gives the outer edge of the spectrum $\hat{s}_+$, one solution $0 \leqslant \hat{u}_0^- \leqslant \hat{\nu}_{\min}$ which gives the inner edge of the spectrum $\hat{s}_-$, and $M - 1$ solutions $\hat{u}_0 \geqslant \nu_{\min}$ which must be disregarded due to the cut in the complex $\hat{u}$ plane mentioned above. It is clear that Eq. (76) cannot have more than $M + 1$ solutions implying that we have found all solutions. With this result we know how to choose the correct solution of Eq. (71), which was what we wanted to establish.

Before ending the discussion about the edges of the spectral density, it is worth noting that Eq. (71) is an $(M + 1)$st order equation, and the general case, for this reason, cannot be solved analytically. However, it is possible to set up some analytical bounds for the edges. The starting point are the conditions $0 \leqslant \hat{u}_0^- \leqslant \hat{\nu}_{\min}$ and $-\infty < \hat{u}_0^+ < -1$ for the saddle points. We will analyze step by step first the bounds on the inner edge, $\hat{s}_-$, and then on the outer edge, $\hat{s}_+$.

Let us consider the inner edge, $\hat{s}_-$. Since $0 \leqslant \hat{\nu}_{\min} \leqslant \hat{\nu}_m$, $m = 1, \ldots, M$, we can readily estimate

$$\min\left\{\frac{\hat{\nu}_m}{\hat{\nu}_m+1}, \frac{\hat{\nu}_{\max}-\hat{u}_0}{\hat{\nu}_{\max}+1}\right\} \geqslant \frac{\hat{\nu}_m-\hat{u}_0}{\hat{\nu}_m+1} \geqslant \frac{\hat{\nu}_{\min}-\hat{u}_0}{\hat{\nu}_{\min}+1} \tag{77}$$

for any $\hat{u}_0 \geqslant 0$. Note that these bounds hold since the rational function, $(\hat{\nu}_m - \hat{u}_0)/(\hat{\nu}_m + 1)$, is strictly monotonously increasing in $\hat{\nu}_m$ for $\hat{u}_0 \geqslant 0$. We plug Eq. (77) into Eq. (70) and extremize the lower and upper bound, which yields

$$0 \leqslant \hat{s}_-(\hat{\nu}_{\min}) \leqslant \hat{s}_- \leqslant \min\left\{\prod_{m=1}^{M}\frac{\hat{\nu}_m}{\hat{\nu}_m+1}, \hat{s}_-(\hat{\nu}_{\max})\right\} < 1, \tag{78}$$

where we made use of the result (72) for the case when all $\hat{\nu}$ are equal to $\hat{\nu}_{\min}$ or to $\hat{\nu}_{\max}$. The bounds (78) are not at all optimal. However, they immediately reflect the fact that the inner edge vanishes if and only if $\hat{\nu}_{\min}$ vanishes.

For the outer edge we have to employ the condition $\hat{u}_0 < -1$, which yields the estimates

$$\frac{\hat{\nu}_{\min}-\hat{u}_0}{\hat{\nu}_{\min}+1} \geqslant \frac{\hat{\nu}_m-\hat{u}_0}{\hat{\nu}_m+1} \geqslant \frac{\hat{\nu}_{\max}-\hat{u}_0}{\hat{\nu}_{\max}+1}. \tag{79}$$

Hereby we used the fact that the rational function, $(\hat{\nu}_m - \hat{u}_0)/(\hat{\nu}_m + 1)$, is monotonously decreasing in $\hat{\nu}_m$ in the considered regime. Employing the result (72) we find the bounds

$$1 < \hat{s}_+(\hat{\nu}_{\max}) \leqslant \hat{s}_+ \leqslant \hat{s}_+(\hat{\nu}_{\min}) \leqslant \frac{(M+1)^{M+1}}{M^M} < \infty. \tag{80}$$

Again the bounds can certainly be improved but they give a good picture of the relation between the case of degenerate $\hat{\nu}$ [cf. Eq. (72)] and the general case, $\hat{\nu}_j \neq \hat{\nu}_i$ for $j \neq i$.

V. MUTUAL INFORMATION FOR PROGRESSIVE SCATTERING

We now turn to a brief discussion of the mutual information, which is an important quantity in wireless telecommunication. We look at a MIMO communication channel with multifold scattering as mentioned in Sec. I. The communication link is described by a channel matrix given by a product of complex ($\beta = 2$) matrices from the Wishart ensemble as in Eq. (1). The mutual information is defined as

$$\begin{aligned}\mathcal{I}(\gamma, s) &= \log_2 \det\left[\mathbf{1}_{N_0} + \gamma\frac{\mathbb{Y}_M\mathbb{Y}_M^\dagger}{\mathcal{N}_M}\right] \\ &= \sum_{a=1}^{N_0}\log_2\left(1+\gamma\frac{s_a}{\mathcal{N}_M}\right),\end{aligned} \tag{81}$$

where γ is the constant signal-to-noise ratio at the transmitter and s_a are the singular values distributed according to the density (52). The mutual information measures an upper bound for the spectral efficiency in bits per time per bandwidth (bit/s/Hz).

In order to evaluate the expectation value of the mutual information, the so-called ergodic mutual information, we rewrite the logarithm as a Meijer G-function, see Eq. (A10). We use the expression (47) for the functions $\varphi_n^M(s)$, while we write $p_n^M(s)$ in polynomial form (43). The integration over the product of two Meijer G-functions can be performed using Eq. (A6), which finally yields

$$\begin{aligned}\hat{\mathbb{E}}\{\mathcal{I}(\gamma,\hat{s})\} = \frac{1}{\ln 2}\sum_{n=0}^{N_0-1}\sum_{k,\ell=0}^{n}\frac{(-1)^{k+\ell}}{(n-k)!(n-\ell)!}\frac{n!}{k!\ell!}\frac{(n+\nu_1)!}{(\ell+\nu_1)!}\prod_{m=1}^{M}\frac{1}{(k+\nu_m)!} \\ \times G^{M+2,\,1}_{2,\,M+2}\left(\begin{matrix}0,\,1\\ k+1+\nu_M,\,\ldots,\,k+\ell+1+\nu_1,\,0,\,0\end{matrix}\middle|\,\gamma^{-1}\right).\end{aligned} \tag{82}$$

For square matrices, i.e. $\nu_i = 0$ for all $i = 1 \ldots M$, this triple sum was derived in [23]. Although it is not obvious from this formulation, the mutual information is also independent of the ordering of ν_m. This is reflected after simplifying the expression (82) with help of a combination of the Eqs. (40), (48), (A5), and (A8) to

$$\hat{\mathbb{E}}\{\mathcal{I}(\gamma,\hat{s})\} = \frac{1}{\ln 2}\sum_{n=0}^{N_0-1}\sum_{k=0}^{n}\frac{(-1)^k}{(n-k)!k!}\prod_{m=1}^{M}\frac{1}{(k+\nu_m)!}G^{M+2,\,2}_{3,\,M+3}\left(\begin{matrix}k-n+1,\,0,\,1\\ k+1+\nu_M,\,\ldots,\,k+1+\nu_1,\,0,\,0,\,k+1\end{matrix}\middle|\,\gamma^{-1}\right). \tag{83}$$

Hence, the channel matrix does not depend on the ordering of the scattering objects as long as the signal passes through all scatterers.

VI. CONCLUSIONS AND OUTLOOK

In this paper we have studied the correlations of the singular values of the product of M rectangular complex matrices from independent Wishart ensembles. This generalizes the classical result for the so-called Wishart-Laguerre unitary ensemble (or chiral unitary ensemble) at $M = 1$ and is a direct extension of a recent result for the product of square matrices [23]. We have seen that the problem of determining the statistical properties of the product of rectangular matrices can be equivalently formulated as a problem with the product of quadratic matrices and a modified, also called "induced," measure, see [24] for a general derivation. The expense of this reformulation of the problem is the introduction of additional determinants in the partition function.

We have shown that the joint probability density function for the singular values can be expressed in terms of Meijer G-functions. The approach we have used relies on an integration formula for the Meijer G-function as well as

on the Harish-Chandra-Itzykson-Zuber integration formula. Due to the latter this method is limited to the complex case ($\beta = 2$). Furthermore, it has been shown, using a two-matrix model and the method of biorthogonal polynomials, that all correlation functions can be expressed as a determinantal point process containing Meijer G-functions. From the explicit expressions we derived it follows that all correlation functions are independent of the ordering of the matrix dimensions.

The level density (or one-point correlation function) was discussed in detail. We used the spectral density to calculate all moments and derived its macroscopic limit. In particular, we analyzed the location of the end points of the spectrum in the macroscopic limit for arbitrary M and derived some narrow bounds for the location of these edges.

As an application we briefly discussed the ergodic mutual information, and how the singular values of products of random matrices are related to progressive scattering in MIMO communication channels.

The results presented in this work concern matrices of finite size, while previous results for the product of rectangular random matrices were only derived in the macroscopic large-N_0 limit. The explicit expressions for all correlation functions at finite size make it possible to also discuss microscopic properties, such as the local correlations in the bulk and at the edges. Due to known universality results for random matrices it is expected that such an analysis should reproduce the universal sine and Airy kernel in the bulk and at the soft edge(s), respectively, after an appropriate unfolding. Close to the origin the level statistics will crucially depend on whether the difference of the individual matrix dimensions to the smallest one, $\nu_m = N_m - N_0$, scales with N_0. If it does, this will lead to a soft edge. If not, it is expected that the microscopic behavior at the origin will be sensitive to M and ν_m. For a single matrix with $M = 1$ (the Wishart-Laguerre ensemble), it is already known that this limit yields different Bessel universality classes labeled by ν_1.

Furthermore, the determinantal structure of the correlation functions make it possible to study the distribution of individual singular values, which is an intriguing problem in its own right.

It has been pointed out in [37] that for the product of two square matrices, $M = 2$ and $\nu_1 = 0$, the biorthogonal polynomials in question are special cases of multiple orthogonal polynomials associated with the modified Bessel function of the second kind. It is an intriguing task to see whether this approach can be extended to the more general case with $M \geqslant 2$ and rectangular matrices. Progress in this direction has already been made [38].

ACKNOWLEDGMENTS

We acknowledge partial support by SFB|TR12 "Symmetries and Universality in Mesoscopic Systems" of the German Science Foundation DFG (G.A.) and by the International Graduate College IRTG 1132 "Stochastic and Real World Models" of the German Science Foundation DFG (J.R.I). Moreover, we thank Arno Kuijlaars and Lun Zhang [38] as well as Eugene Strahov [36] for sharing their private communications with us.

APPENDIX: SPECIAL FUNCTIONS AND SOME OF THEIR IDENTITIES

In this appendix we collect some definitions and identities for the generalized hypergeometric function and for the Meijer G-function which are used in this paper.

The generalized hypergeometric function is defined by a power series in its region of convergence [41],

$$ {}_pF_q\left(\begin{matrix}a_1,\ \dots\ ,a_p\\ b_1,\ \dots\ ,b_q\end{matrix}\,\middle|\, z\right) \equiv \sum_{k=0}^{\infty} \frac{\prod_{i=1}^{p}(a_i)_k}{\prod_{i=1}^{q}(b_i)_k}\frac{z^k}{k!}, \tag{A1}$$

where the Pochhammer symbol is defined by $(a)_0 = 1$ and $(a)_n \equiv (a+n-1)(a)_{n-1} = a(a+1)\cdots(a+n-1)$ for $n \geqslant 1$. It is clear that the hypergeometric series (A1) terminates if any of the a_i's is a negative integer. In particular, if n is a positive integer, then

$$ {}_{p+1}F_q\left(\begin{matrix}-n,a_1,\ \dots\ ,a_p\\ b_1,\ \dots\ ,b_q\end{matrix}\,\middle|\, z\right) = \sum_{k=0}^{n} \frac{(-1)^k n!}{(n-k)!}\frac{\prod_{i=1}^{p}(a_i)_k}{\prod_{i=1}^{q}(b_i)_k}\frac{z^k}{k!}, \tag{A2}$$

which is a polynomial of degree n or less.

The Meijer G-function can be considered as a generalization of the generalized hypergeometric function. It is usually defined by a contour integral in the complex plane [41],

$$ G^{m,n}_{p,q}\left(\begin{matrix}a_1,\ \dots\ ,a_p\\ b_1,\ \dots\ ,b_q\end{matrix}\,\middle|\, z\right) \equiv \frac{1}{2\pi\imath}\int_L du\, z^u \frac{\prod_{i=1}^{m}\Gamma[b_i-u]\prod_{i=1}^{n}\Gamma[1-a_i+u]}{\prod_{i=n+1}^{p}\Gamma[a_i-u]\prod_{i=m+1}^{q}\Gamma[1-b_i+u]}. \tag{A3}$$

The contour runs from $-\imath\infty$ to $+\imath\infty$ and is chosen such that it separates the poles stemming from $\Gamma[b_i-u]$ and the poles stemming from $\Gamma[1-a_i+u]$. Furthermore, this contour can be considered as an inverse Mellin transform. For an extensive discussion of the integration path L and the requirements for convergence, see [42].

It follows that the Mellin transform of a Meijer G-function is given by [41]

$$ \int_0^{\infty} ds\, s^{u-1} G^{m,n}_{p,q}\left(\begin{matrix}a_1,\ \dots\ ,a_p\\ b_1,\ \dots\ ,b_q\end{matrix}\,\middle|\, sz\right) = z^{-u}\frac{\prod_{i=1}^{m}\Gamma[b_i+u]\prod_{i=1}^{n}\Gamma[1-a_i-u]}{\prod_{i=n+1}^{p}\Gamma[a_i+u]\prod_{i=m+1}^{q}\Gamma[1-b_i-u]}, \tag{A4}$$

which results from the definition of the Meijer G-function (A3). In combination with the definition of the gamma function we have another identity,

$$ \int_0^{\infty} dt\, e^{-t} t^{b_0-1} G^{m,n}_{p,q}\left(\begin{matrix}a_1,\ \dots\ ,a_p\\ b_1,\ \dots\ ,b_q\end{matrix}\,\middle|\, \frac{s}{t}\right) = G^{m+1,n}_{p,q+1}\left(\begin{matrix}a_1,\ \dots\ ,a_p\\ b_0,\ \dots\ ,b_q\end{matrix}\,\middle|\, s\right). \tag{A5}$$

Both of these integral identities are used throughout this paper. Another integral identity, which is used in Sec. V, allows us to integrate over the product of two Meijer G-functions [43],

$$\int_0^\infty ds\, G^{m,n}_{p,q}\left(\begin{matrix}a_1,\ \dots\ ,a_p\\ b_1,\ \dots\ ,b_q\end{matrix}\,\middle|\,\eta s\right) G^{\mu,\nu}_{\sigma,\tau}\left(\begin{matrix}c_1,\ \dots\ ,c_\sigma\\ d_1,\ \dots\ ,d_\tau\end{matrix}\,\middle|\,\omega s\right) = \frac{1}{\omega} G^{m+\nu,n+\mu}_{p+\tau,q+\sigma}\left(\begin{matrix}a_1,\ \dots\ ,a_n,\ -d_1,\ \dots\ ,-d_\tau,a_{n+1},\ \dots\ ,a_p\\ b_1,\ \dots\ ,b_m,\ -c_1,\ \dots\ ,-c_\sigma,b_{m+1},\ \dots\ ,b_q\end{matrix}\,\middle|\,\frac{\eta}{\omega}\right). \tag{A6}$$

The full set of restrictions on the indices for this integration formula can be found in [43].

In addition to the integral identities given above, we need some other identities for the Meijer G-function. We employ several times that it is possible to absorb powers of the argument into the Meijer G-function by making a shift in the arguments [41],

$$z^\rho G^{m,n}_{p,q}\left(\begin{matrix}a_1,\ \dots\ ,a_p\\ b_1,\ \dots\ ,b_q\end{matrix}\,\middle|\,z\right) = G^{m,n}_{p,q}\left(\begin{matrix}a_1+\rho,\ \dots\ ,a_p+\rho\\ b_1+\rho,\ \dots\ ,b_q+\rho\end{matrix}\,\middle|\,z\right). \tag{A7}$$

For computing the function $\varphi_n^M(s)$ in Sec. III, we need the differential identity [43]

$$z^n \frac{d^n}{dz^n} G^{m,n}_{p,q}\left(\begin{matrix}a_1,\ \dots\ ,a_p\\ b_1,\ \dots\ ,b_q\end{matrix}\,\middle|\,\frac{1}{z}\right) = (-1)^n G^{m,n+1}_{p+1,q+1}\left(\begin{matrix}1-n,a_1,\ \dots\ ,a_p\\ b_1,\ \dots\ ,b_q,1\end{matrix}\,\middle|\,\frac{1}{z}\right). \tag{A8}$$

We also use that the generalized hypergeometric polynomial is related to the Meijer G-function by

$${}_1F_q\left(\begin{matrix}-n\\ b_1,\ \dots\ ,b_q\end{matrix}\,\middle|\,z\right) = n!\prod_{i=1}^{q}\Gamma[b_i]\, G^{1,0}_{1,M+1}\left(\begin{matrix}n+1\\ 0,1-b_1,\ \dots\ ,1-b_q\end{matrix}\,\middle|\,z\right), \tag{A9}$$

in order to write the polynomial $p_n^M(s)$ as a Meijer G-function in Sec. III.

As a last remark of this appendix, it should be mentioned that the Meijer G-function contains a vast number of elementary and special functions as special cases (e.g. see [44]). We mention that

$$G^{1,0}_{0,1}\left(\begin{matrix}-\\ b\end{matrix}\,\middle|\,z\right) = z^b e^{-z} \quad \text{and} \quad G^{1,2}_{2,2}\left(\begin{matrix}1,1\\ 1,0\end{matrix}\,\middle|\,z\right) = \ln(1+z), \tag{A10}$$

which becomes useful in Secs. II and V, respectively.

[1] G. Akemann, J. Baik, and P. Di Francesco, *The Oxford Handbook of Random Matrix Theory* (Oxford University Press, Oxford, 2011).

[2] A. Crisanti, G. Paladin, and A. Vulpiani, *Products of Random Matrices in Statistical Physics* (Springer, Heidelberg, 1993).

[3] A. D. Jackson, B. Lautrup, P. Johansen, and M. Nielsen, Phys. Rev. E **66**, 066124 (2002).

[4] E. Gudowska-Nowak, R. A. Janik, J. Jurkiewicz, and M. A. Nowak, Nucl. Phys. B **670**, 479 (2003).

[5] J. C. Osborn, Phys. Rev. Lett. **93**, 222001 (2004).

[6] G. Akemann, Int. J. Mod. Phys. A **22**, 1077 (2007).

[7] A. M. Brzoska, F. Lenz, J. W. Negele, and M. Thies, Phys. Rev. D **71**, 034008 (2005).

[8] R. Narayanan and H. Neuberger, J. High Energy Phys. 12 (2007) 066.

[9] J. P. Blaizot and M. A. Nowak, Phys. Rev. Lett. **101**, 102001 (2008).

[10] J.-P. Bouchaud, L. Laloux, M. A. Miceli, and M. Potters, Eur. Phys. J. B **55**, 201 (2007).

[11] A. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications* (Now Publishers, Hanover, MA, 2004).

[12] M. L. Mehta, *Random Matrices*, 3rd ed. (Academic Press, New York, 2004).

[13] Z. Burda, R. A. Janik, and B. Waclaw, Phys. Rev. E **81**, 041132 (2010).

[14] Z. Burda, A. Jarosz, G. Livan, M. A. Nowak, and A. Swiech, Phys. Rev. E **82**, 061114 (2010).

[15] F. Götze and A. Tikhomirov, arXiv:1012.2710.

[16] S. O'Rourke and A. Soshnikov, Electron. J. Probab. **16**, 2219 (2011).

[17] R. R. Müller, IEEE Trans. Inf. Theor. **48**, 2086 (2002).

[18] T. Banica, S. Belinschi, M. Capitaine, and B. Collins, Can. J. Math. **63**, 3 (2011).

[19] F. Benaych-Georges, Ann. Inst. Henri Poincaré Probab. Stat. **46**, 644 (2010).

[20] G. Akemann and Z. Burda, J. Phys. A **45**, 465201 (2012).

[21] G. Akemann and E. Strahov, J. Stat. Phys. **151**, 987 (2013).

[22] J. R. Ipsen, J. Phys. A **46**, 265201 (2013).

[23] G. Akemann, M. Kieburg, and L. Wei, J. Phys. A **46**, 275205 (2013).

[24] J. R. Ipsen and M. Kieburg, arXiv:1310.4154.

[25] B. Collins, I. Nechita, and K. Życzkowski, J. Phys. A **43**, 275303 (2010).

[26] K. Życzkowski, K. A. Penson, I. Nechita, and B. Collins, J. Math. Phys. **52**, 062201 (2011).

[27] G. J. Foschini, Bell Labs Tech. J. **1**, 41 (1996).

[28] G. J. Foschini and M. J. Gans, Wireless Pers. Commun. **6**, 311 (1998).

[29] E. Telatar, Eur. Trans. Telecommun. **10**, 585 (1999).

[30] Harish-Chandra, Am. J. Math. **79**, 87 (1957).

[31] C. Itzykson and J. B. Zuber, J. Math. Phys. **21**, 411 (1980).
[32] J. Fischmann, W. Bruzda, B. A. Khoruzhenko, H.-J. Sommers, and K. Życzkowski, J. Phys. A **45**, 075203 (2012).
[33] B. Eynard and M. L. Mehta, J. Phys. A **31**, 4449 (1998).
[34] M. Bertola, M. Gekhtman, and J. Szmigielski, Commun. Math. Phys. **287**, 983 (2009).
[35] M. Bertola, M. Gekhtman, and J. Szmigielski, J. Approx. Theory **162**, 832 (2010).
[36] E. Strahov (private communication, 2013).
[37] L. Zhang, J. Math. Phys. **54**, 083303 (2013).
[38] A. Kuijlaars and L. Zhang, arXiv:1308.1003.
[39] K. A. Penson and K. Życzkowski, Phys. Rev. E **83**, 061118 (2011).
[40] Vinayak, Phys. Rev. E **88**, 042130 (2013).
[41] I. I. S. Gradshtein, I. I. M. Ryzhik, and A. Jeffrey, *Table on Integrals, Series, and Products* (Academic Press, San Diego, 2000).
[42] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions* (Cambridge University Press, New York, 2010).
[43] A. A. P. Prudnikov, Y. A. Brychkov, I. U. A. Brychkov, and O. I. Maričev, *Integrals and Series: More Special Functions*, Vol. 3 (Gordon and Breach Science Publishers, London, 1990).
[44] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions* (McGraw-Hill, London, 1953).