Random Vandermonde Matrices-Part I: Fundamental results

Øyvind Ryan, *Member, IEEE*
Centre of Mathematics for Applications,
University of Oslo, P.O. Box 1053 Blindern, 0316 Oslo, Norway
Phone: +47 93 24 83 21
Fax: +47 22 85 43 49
Email: oyvindry@ifi.uio.no
Mérouane Debbah, *Member, IEEE*
SUPELEC,
Alcatel-Lucent Chair on Flexible Radio, Plateau de Moulon,
3 rue Joliot-Curie,
91192 GIF SUR YVETTE CEDEX, France
Phone: +33 1 69 85 20 07
Fax: +33 1 69 85 12 59
Email: merouane.debbah@supelec.fr

***Abstract*—In this first part, analytical methods for finding moments of random Vandermonde matrices are developed. Vandermonde Matrices play an important role in signal processing and communication applications such as direction of arrival estimation, precoding or sparse sampling theory for example. Within this framework, we extend classical freeness results on random matrices with i.i.d entries and show that Vandermonde structured matrices can be treated in the same vein with different tools. We focus on various types of Vandermonde matrices, namely Vandermonde matrices with or without uniformly distributed phases, as well as generalized Vandermonde matrices (with non-uniform distribution of powers). In each case, we provide explicit expressions of the moments of the associated Gram matrix, as well as more advanced models involving the Vandermonde matrix. Comparisons with classical i.i.d. random matrix theory are provided and free deconvolution results are also discussed.**

***Index Terms*—Vandermonde matrices, Random Matrices, deconvolution, limiting eigenvalue distribution, MIMO.**

I. Introduction

We will consider Vandermonde matrices $\mathbf{V}$ of dimension $N \times L$ of the form

$$\mathbf{V} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & \cdots & 1 \\ e^{-j\omega_1} & \cdots & e^{-j\omega_L} \\ \vdots & \ddots & \vdots \\ e^{-j(N-1)\omega_1} & \cdots & e^{-j(N-1)\omega_L} \end{pmatrix} \tag{1}$$

where ω_1,...,ω_L are independent and identically distributed (phases) taking values on $[0, 2\pi)$. Such matrices occur frequently in many applications, such as finance [1], signal array processing [2], [3], [4], [5], [6], ARMA processes [7], cognitive radio [8], security [9], wireless communications [10] and biology [11] and have been much studied. The main results are related to the distribution of the determinant of (1) [12]. The large majority of known results on the eigenvalues of the associated Gram matrix concern Gaussian matrices [13] or matrices with independent entries. None have dealt with the Vandermonde case. For the Vandermonde case, the results depend heavily on the distribution of the entries, and do not give any hint on the asymptotic behaviour as the matrices become large. In the realm of wireless channel modelling, [14] has provided some insight on the behaviour of the eigenvalues of random Vandermonde matrices for a specific case, without any formal proof. We prove here that the case is in fact more involved than what was claimed.

In many applications, N and L are quite large, and we may be interested in studying the case where both go to ∞ at a given ratio, with $\frac{L}{N} \rightarrow c$. Results in the literature say very little on the asymptotic behaviour of (1) under this growth condition. The results, however, are well known for other models. The factor $\frac{1}{\sqrt{N}}$, as well as the assumption that the Vandermonde entries $e^{-j\omega_i}$ lie on the unit circle, are included in (1) to ensure that our analysis will give limiting asymptotic behaviour. Without this assumption, the problem at hand is more involved, since the rows of the Vandermonde matrix with the highest powers would dominate in the calculations of the moments when the matrices grow large, and also grow faster to infinity than the $\frac{1}{\sqrt{N}}$ factor in (1), making asymptotic analysis difficult. In general, often the moments, not the moments of the determinants, are the quantities we seek. Results in the literature also say very little on the moments of Vandermonde matrices. The literature says very little on the mixed moments of Vandermonde matrices and matrices independent from them. This is in contrast to Gaussian matrices, where exact expressions [15] and their asymptotic behaviour [16]

This project is partially sponsored by the project BIONET (INRIA).
This work was supported by Alcatel-Lucent within the Alcatel-Lucent Chair on flexible radio at SUPELEC

are known using the concept of freeness [16] which is central for describing the mixed moments.

The derivation of the moments are a useful basis for performing deconvolution. For Gaussian matrices, deconvolution has been handled in the literature [17], [18], [15], [19]. Similar flavored results will here be proved for Vandermonde matrices. Concerning the moments, it will be the asymptotic moments of random matrices of the form $\mathbf{V}^H\mathbf{V}$ which will be studied, where $(.)^H$ denotes hermitian transpose. We will also consider mixed moments of the form $\mathbf{D}\mathbf{V}^H\mathbf{V}$, where $\mathbf{D}$ is a square diagonal matrix independent from $\mathbf{V}$.

We will also extend our results to what are called generalized Vandermonde matrices, i.e. matrices where the columns do not consist of uniformly distributed powers. These are important for applications to finance [1]. The tools used for standard Vandermonde matrices in this paper will allow us to find the asymptotic behaviour of many generalized Vandermonde matrices.

While we provide the full computation of lower order moments, we also describe how the higher order moments can be computed. Tedious evaluation of many integrals is needed for this, but numerical methods can also be applied. Surprisingly, it turns out that the first three limit moments can be expressed in terms of the Marčhenko Pastur law [16], [20]. For higher order moments this is not the case, although we state an interesting inequality involving the Vandermonde limit moments and the moments of the classical Poisson distribution and the Marčhenko Pastur law, also known as the free Poisson distribution [16].

This paper is organized as follows: Section II contains a general result for the mixed moments of Vandermonde matrices and matrices independent from them. We will differ between the case where the phase ω in (1) are uniformly distributed on $[0.2\pi)$, and the more general cases. The case of uniformly distributed phases is handled in section III. In this case it turns out that one can have very nice expressions, for both the asymptotic moments, as well as for the lower order moments. Section IV considers the more general case when ω has a continous density, and shows how the asymptotics can be described in terms of the case when ω is uniformly distributed. The case where the density of ω has singularities displays different asymptotic behaviour, and is handled in section V. Section VI states results on generalized Vandermonde matrices. The case when the powers also have some random distributions is also handled here. Section VII handles mixed moments of independent Vandermonde matrices. Section VIII discusses our results and puts them in a general deconvolution perspective, comparing with other deconvolution results, such as those for Gaussian deconvolution.

In the following, upper (lower boldface) symbols will be used for matrices (column vectors) whereas lower symbols will represent scalar values, $(.)^T$ will denote transpose operator, $(.)^\star$ conjugation and $(.)^H = \left((.)^T\right)^\star$ hermitian transpose. $\mathbf{I}_n$ will represent the $n \times n$ identity matrix. We let tr_n be the normalized trace for matrices of order $n \times n$, and Tr the non-normalized trace. $\mathbf{V}$ will be used only to denote Vandermonde matrices with a given phase distribution. The dimensions of the Vandermonde matrices will always be $N \times L$ unless otherwise stated, and the phase distribution of the Vandermonde matrices will always be denoted by ω.

II. A GENERAL RESULT FOR THE MIXED MOMENTS OF VANDERMONDE MATRICES

We first state a general theorem applicable to Vandermonde matrices with any phase distribution. The proof for this theorem, as well as for theorems succeeding it, are based on calculations where partitions are highly involved. We denote by $\mathcal{P}(n)$ the set of all partitions of $\{1, ..., n\}$, and we will use ρ as notation for a partition in $\mathcal{P}(n)$. The set of partitions will be equipped with the refinement order $\leq$, i.e. $\rho_1 \leq \rho_2$ if and only if any block of ρ_1 is contained within a block of ρ_2. Also, we will write $\rho = \{\rho_1, ..., \rho_k\}$, where ρ_j are the blocks of ρ, and let $|\rho|$ denote the number of blocks in ρ. We denote by 0_n the partition with n blocks, and by 1_n the partition with 1 block.

In the following $\mathbf{D}_r(N), 1 \leq r \leq n$ are diagonal $L \times L$ matrices, and $\mathbf{V}$ is of the form (1). We will attempt to find

$$M_n = \lim_{N\to\infty} E[tr_L(\quad \begin{array}{l} \mathbf{D}_1(N)\mathbf{V}^H\mathbf{V}\mathbf{D}_2(N)\mathbf{V}^H\mathbf{V} \\ \cdots \times \mathbf{D}_n(N)\mathbf{V}^H\mathbf{V})] \end{array} \tag{2}$$

for many types of Vandermonde matrices, under the assumption that $\frac{L}{N} \to c$, and under the assumption that the $\mathbf{D}_r(N)$ have a joint limit distribution as $N \to \infty$ in the following sense:

Definition 1: We will say that the $\{\mathbf{D}_r(N)\}_{1\leq r\leq n}$ have a joint limit distribution as $N \to \infty$ if the limit

$$D_{i_1,...,i_s} = \lim_{N\to\infty} tr_L\left(\mathbf{D}_{i_1}(N)\cdots\mathbf{D}_{i_s}(N)\right) \tag{3}$$

exists for all choices of $i_1, ..., i_s$. For $\rho = \{\rho_1, ..., \rho_k\}$, with $\rho_i = \{\rho_{i1}, ..., \rho_{i|\rho_i|}\}$, we also define $D_{\rho_i} = D_{i_{\rho_{i1}},...,i_{\rho_{i|\rho_i|}}}$, and $D_\rho = \prod_{i=1}^k D_{\rho_i}$.

Had we replaced Vandermonde matrices with Gaussian matrices, free deconvolution results [19] could help us compute the quantities $D_{i_1,...,i_s}$ from M_n. For this, the cumulants of the Gaussian matrices are needed, which asymptotically have a very nice form. For Vandermonde matrices, the role of cumulants is taken by the following quantites

Definition 2: Define

$$\begin{aligned} K_{\rho,\omega,N} = \quad & \frac{1}{N^{n+1-|\rho|}} \times \\ & \int_{(0,2\pi)^{|\rho|}} \prod_{k=1}^n \frac{1-e^{jN(\omega_{b(k-1)}-\omega_{b(k)})}}{1-e^{j(\omega_{b(k-1)}-\omega_{b(k)})}}, \\ & d\omega_1 \cdots d\omega_{|\rho|}, \end{aligned} \tag{4}$$

where $\omega_{\rho_1}, ..., \omega_{\rho_{|\rho|}}$ are i.i.d. (indexed by the blocks of ρ), all with the same distribution as ω, and where $b(k)$ is the block of ρ which contains k (where notation is cyclic, i.e. $b(-1) = b(n)$). If the limit

$$K_{\rho,\omega} = \lim_{N\to\infty} K_{\rho,\omega,N}$$

exists, then $K_{\rho,\omega}$ is called a *Vandermonde mixed moment expansion coefficient.*

These coefficients will for Vandermonde matrices play the same role as the cumulants do for large Gaussian matrices. We will not call them cumulants, however, since they don't share the same multiplicative properties (embodied in what is called the moment cumulant formula).

The following is the main result of the paper. Different versions of it adapted to different Vandermonde matrices will be stated in the succeeding sections.

Theorem 1: Assume that the $\{\mathbf{D}_r(N)\}_{1\leq r\leq n}$ have a joint limit distribution as $N\to\infty$. Assume also that all Vandermonde mixed moment expansion coefficients $K_{\rho,\omega}$ exist. Then the limit

$$M_n = \lim_{N\to\infty} E[tr_L(\mathbf{D}_1(N)\mathbf{V}^H\mathbf{V}\mathbf{D}_2(N)\mathbf{V}^H\mathbf{V} \cdots \times \mathbf{D}_n(N)\mathbf{V}^H\mathbf{V})] \quad (5)$$

also exists when $\frac{L}{N}\to c$, and equals

$$\sum_{\rho\in\mathcal{P}(n)} K_{\rho,\omega}c^{|\rho|-1}D_\rho. \quad (6)$$

The proof of theorem 1 can be found in appendix A. Although the limit of $K_{\rho,\omega,N}$ as $N\to\infty$ may not exist, it will be clear from section IV that it exists when the density of ω is continous. Theorem 1 explains how convolution with Vandermonde matrices can be performed, and also provides us an extension of the concept of free convolution to Vandermonde matrices. Note that when $\mathbf{D}_1(N)=\cdots=\mathbf{D}_n(N)=I_L$, we have that

$$M_n = \lim_{N\to\infty} E\left[tr_L\left(\left(\mathbf{V}^H\mathbf{V}\right)^n\right)\right],$$

so that our our results also include the limit moments of the Vandermonde matrices themselves. M_n corresponds also to the limit moments of the empirical eigenvalue distribution $F^N_{\mathbf{V}^H\mathbf{V}}$ defined by

$$F^N_{\mathbf{V}^H\mathbf{V}}(\lambda) = \frac{\#\{i|\lambda_i\leq\lambda\}}{N},$$

(where λ_i are the (random) eigenvalues of $\mathbf{V}^H\mathbf{V}$), i.e.

$$M_n = \lim_{N\to\infty} E\left[\int \lambda^n dF^N(\lambda)\right].$$

(6) will also be useful on the scaled form

$$cM_n = \sum_{\rho\in\mathcal{P}(n)} K_{\rho,\omega}(cD)_\rho. \quad (7)$$

When $\mathbf{D}_1(N)=\mathbf{D}_2(N)=\cdots=\mathbf{D}_n(N)$, we denote their common value $\mathbf{D}(N)$, and define the sequence $D=(D_1,D_2,...)$ with $D_n=\lim_{N\to\infty} tr_L\left((\mathbf{D}(N))^n\right)$. In this case D_ρ does only depend on the block cardinalities $|\rho_j|$, so that we can group together the $K_{\rho,\omega}$ for ρ with equal block cardinalities. If we group the blocks of ρ so that their cardinalities are in descending order, and set

$$\mathcal{P}(n)_{r_1,r_2,\ldots,r_k} = \{\rho=\{\rho_1,...,\rho_k\}\in\mathcal{P}(n)||\rho_i|=r_i\forall i\},$$

where $r_1\geq r_2\geq\cdots\geq r_k$, and also write

$$K_{r_1,r_2,\ldots,r_k} = \sum_{\rho\in\mathcal{P}(n)_{r_1,r_2,\ldots,r_k}} K_{\rho,\omega}, \quad (8)$$

then, after performing the substitutions

$$\begin{aligned} m_n &= (cM)_n = c\lim_{N\to\infty} E\left[tr_L\left(\left(\mathbf{D}(N)\mathbf{V}^H\mathbf{V}\right)^n\right)\right], \\ d_n &= (cD)_n = c\lim_{N\to\infty} tr_L\left(\mathbf{D}^n(N)\right), \end{aligned} \quad (9)$$

(7) can be written

$$m_n = \sum_{\substack{r_1,\ldots,r_k \\ r_1+\cdots+r_k=n}} K_{r_1,r_2,\ldots,r_k}\prod_{j=1}^k d_{r_j}. \quad (10)$$

For the first 5 moments this becomes

$$\begin{aligned} m_1 &= K_1d_1 \\ m_2 &= K_2d_2+K_{1,1}d_1^2 \\ m_3 &= K_3d_3+K_{2,1}d_2d_1^2+K_{1,1,1}d_1^3 \\ m_4 &= K_4d_4+K_{3,1}d_3d_1+K_{2,2}d_2^2+K_{2,1,1}d_2d_1^2+ \\ &\quad K_{1,1,1,1}d_1^4 \\ m_5 &= K_5d_5+K_{4,1}d_4d_1++K_{3,2}d_3d_2+ \\ &\quad K_{3,1,1}d_3d_1^2+K_{2,2,1}d_2^2d_1+K_{2,1,1,1}d_2d_1^3+ \\ &\quad K_{1,1,1,1,1}d_1^5 \\ \vdots & \qquad \vdots \end{aligned} \quad (11)$$

Thus, the algorithm for computing the asymptotic mixed moments of Vandermonde matrices with matrices independent from them can be split in two:

- (9), which scales with the matrix aspect ratio c, and
- (11), which performs computations independent of the matrix aspect ratio c.

Similar splitting of the algorithm for computing the asymptotic mixed moments of Wishart matrices and matrices independent from them was derived in [19].

Alternatively, (11) gives us means of performing deconvolution. Indeed, suppose that one knows all the moments of $\mathbf{D}\mathbf{V}^H\mathbf{V}$, i.e. the m_k, and would like to infer on the moments of $\mathbf{D}$, i.e. the d_k. By solving recursively the equations (11), one is able to retrieve the d_i: For example,

$$\begin{aligned} d_1 &= \frac{m_1}{K_1} \\ d_2 &= \frac{m_2-K_{1,1}\left(\frac{m_1}{K_1}\right)^2}{K_2}, \end{aligned}$$

and so on. Although the matrices $\mathbf{D}_i(N)$ are assumed to be determinstic matrices throughout the paper, all formulas extend naturally to the case when $\mathbf{D}_i(N)$ are random matrices independent from $\mathbf{V}$. The only difference when the $\mathbf{D}_i(N)$ are random is that certain quantities are replaced with fluctuations. D_1D_2 should for instance be replaced with

$$\lim_{N\to\infty} E\left[tr_L\left(\mathbf{D}(N)\right)tr_L\left((\mathbf{D}(N))^2\right)\right]$$

when $\mathbf{D}_i(N)$ is random.

In the next sections, we will derive and analyze the Vandermonde mixed moment expansion coefficients $K_{\rho,\omega}$ for various cases, which is essential for the the algorithm (11).

III. UNIFORMLY DISTRIBUTED ω

We will let u denote the uniform distribution on $[0,2\pi)$. We can write

$$\begin{aligned} K_{\rho,u,N} = & \frac{1}{(2\pi)^{|\rho|}N^{n+1-|\rho|}}\times \\ & \int_{(0,2\pi)^{|\rho|}}\prod_{k=1}^n \frac{1-e^{jN(x_{b(k-1)}-x_{b(k)})}}{1-e^{j(x_{b(k-1)}-x_{b(k)})}} \\ & dx_1\cdots dx_{|\rho|}, \end{aligned} \quad (12)$$

where integration is w.r.t. Lebesgue measure. In this case one particular class of partitions will be useful to us, the noncrossing partitions:

Definition 3: A partition is said to be noncrossing if, whenever $i<j<k<l$, i and k are in the same block, and also j and l are in the same block, then all i,j,k,l are in the same block. The set of noncrossing partitions is denoted by $NC(n)$.

The noncrossing partitions have already shown their usefulness in expressing the freeness relation in a particularly nice way [21]. Their appearance here is somewhat different than in the case for the relation to freeness:

Theorem 2: Assume that the $\{\mathbf{D}_r(N)\}_{1\leq r\leq n}$ have a joint limit distribution as $N\to\infty$, Then the Vandermonde mixed moment expansion coefficient

$$K_{\rho,u} = \lim_{N\to\infty} K_{\rho,u,N}$$

exists for all ρ. Moreover, $0<K_{\rho,u}\leq 1$, the $K_{\rho,u}$ are rational numbers for all ρ, and $K_{\rho,u}=1$ if and only if ρ is noncrossing.

The proof of theorem 2 can be found in appendix B. Due to theorem 1, theorem 2 guarantees that the asymptotic mixed moments (5) exist when $\frac{L}{N}\to c$ for uniform phase distribution, and are given by (6). The values $K_{\rho,u}$ are in general hard to compute for higher order ρ with crossings. We have performed some of these computations. It turns out that the following computations suffice to obtain the 7 first moments.

Lemma 1: The following holds:

$$\begin{aligned}
K_{\{\{1,3\},\{2,4\}\},u} &= \frac{2}{3}\\
K_{\{\{1,4\},\{2,5\},\{3,6\}\},u} &= \frac{1}{2}\\
K_{\{\{1,4\},\{2,6\},\{3,5\}\},u} &= \frac{1}{2}\\
K_{\{\{1,3,5\},\{2,4,6\}\},u} &= \frac{11}{20}\\
K_{\{\{1,5\},\{3,7\},\{2,4,6\}\},u} &= \frac{9}{20}\\
K_{\{\{1,6\},\{2,4\},\{3,5,7\}\},u} &= \frac{9}{20}.
\end{aligned}$$

The proof of lemma 1 is given in appendix C. Combining theorem 2 and lemma 1 into this form, we will prove the following:

Theorem 3: Assume $\mathbf{D}_1(N)=\mathbf{D}_2(N)=\cdots=\mathbf{D}_n(N)$. When $\omega=u$, (11) takes the form

$$\begin{aligned}
m_1 &= d_1\\
m_2 &= d_2+d_1^2\\
m_3 &= d_3+3d_2d_1+d_1^3\\
m_4 &= d_4+4d_3d_1+\frac{8}{3}d_2^2+6d_2d_1^2+d_1^4\\
m_5 &= d_5+5d_4d_1+\frac{25}{3}d_3d_2+10d_3d_1^2+\\
&\quad \frac{40}{3}d_2^2d_1+10d_2d_1^3+d_1^5\\
m_6 &= d_6+6d_5d_1+12d_4d_2+15d_4d_1^2+\\
&\quad \frac{151}{20}d_3^2+50d_3d_2d_1+20d_3d_1^3+\\
&\quad 11d_2^3+40d_2^2d_1^2+15d_2d_1^4+d_1^6\\
m_7 &= d_7+7d_6d_1+\frac{49}{3}d_5d_2+21d_5d_1^2+\\
&\quad \frac{497}{20}d_4d_3+84d_4d_2d_1+35d_4d_1^3+\\
&\quad \frac{1057}{20}d_3^2d_1+\frac{693}{10}d_3d_2^2+175d_3d_2d_1^2+\\
&\quad 35d_3d_1^4+77d_2^3d_1+\frac{280}{3}d_2^2d_1^3+\\
&\quad 21d_2d_1^5+d_1^7.
\end{aligned}$$

Theorem 2 and lemma 1 reduces the proof of theorem 3 to a simple count of partitions. Theorem 3 is proved in appendix D. To compute higher moments m_k, $K_{\rho,u}$ must be computed for partitions of higher order. The computations performed in appendix C and D should convince the reader that this can be done, but is very tedious.

Following the proof of theorem 2, we can also obtain formulas for the fluctuations of mixed moments of Vandermonde matrices. We will not go into details on this, but only state the following equations without proof:

$$\begin{aligned}
&\lim_{N\to\infty} E\left[tr_L\left(\left(\mathbf{D}(N)\mathbf{V}^H\mathbf{V}\right)^n\right)\left(tr_L\left(\mathbf{D}(N)\mathbf{V}^H\mathbf{V}\right)\right)^m\right]\\
&= E\left[tr_L\left(\left(\mathbf{D}(N)\mathbf{V}^H\mathbf{V}\right)^n\right)\right]D_1^m\\
&c\lim_{N\to\infty} E\left[Tr\left(\left(\mathbf{D}(N)\mathbf{V}^H\mathbf{V}\right)^2\right)tr_L\left(\left(\mathbf{D}(N)\mathbf{V}^H\mathbf{V}\right)^2\right)\right]\\
&= \frac{4}{3}d_2^2+4d_2d_1^2+4d_3d_1+d_4.
\end{aligned} \tag{13}$$

Following the proof of theorem 2 again, we can also obtain exact expressions for moments of lower order random Vandermonde matrices with uniformly distributed phases, not only the limit. We state these only for the first four moments.

Theorem 4: Assume $\mathbf{D}_1(N)=\mathbf{D}_2(N)=\cdots=\mathbf{D}_n(N)$.

When $\omega = u$, (11) takes the exact form

$$
\begin{aligned}
m_1 &= d_1 \\
m_2 &= \left(1 - N^{-1}\right) d_2 + d_1^2 \\
m_3 &= \left(1 - 3N^{-1} + 2N^{-2}\right) d_3 \\
&\quad +3\left(1 - N^{-1}\right) d_1 d_2 + d_1^3 \\
m_4 &= \left(1 - \frac{20}{3}N^{-1} + 11N^{-2} - \frac{37}{6}N^{-3}\right) d_4 \\
&\quad + \left(4 - 12N^{-1} + 8N^{-2}\right) d_3 d_1 \\
&\quad + \left(\frac{8}{3} - 5N^{-1} + \frac{19}{6}N^{-2}\right) d_2^2 \\
&\quad +6\left(1 - N^{-1}\right) d_2 d_1^2 + d_1^4.
\end{aligned}
$$

Theorem 4 is proved in appendix E. Exact formulas for the higher order moments also exist, but they become increasingly complex, as entries for higher order terms L^{-k} also enter the picture. These formulas are also harder to prove for higher order moments. In many cases, exact expressions are not what we need: First order approximations (i.e. expressions where only the L^{-1}-terms are included) can suffice for many purposes. In appendix E, we explain how the simpler case of these first order approximations can be computed. It seems much harder to prove a similar result when the phases are not uniformly distributed.

IV. ω WITH CONTINOUS DENSITY

The following result tells us that the limit $K_{\rho,\omega}$ exists for many ω, and also gives a useful expression for them in terms of the density of ω, and $K_{\rho,u}$.

Theorem 5: The Vandermonde mixed moment expansion coefficients $K_{\rho,\omega} = \lim_{N\to\infty} K_{\rho,\omega,N}$ exist whenever the density p_ω of ω is continous on $[0, 2\pi)$. If this is fulfilled, then

$$K_{\rho,\omega} = K_{\rho,u}(2\pi)^{|\rho|-1}\left(\int_0^{2\pi} p_\omega(x)^{|\rho|}dx\right). \tag{14}$$

The proof is given in appendix F.

Besides providing us with a deconvolution method for finding the mixed moments of the $\{\mathbf{D}_r(N)\}_{1\le r\le n}$, theorem 5 also provides us with a way of inspecting the phase distribution ω, by first finding the moments of the density, i.e. $\int_0^{2\pi} p_\omega(x)^k dx$. However, note that we can not expect to find the density of ω itself, only the density of the density of ω. To see this, define

$$Q_\omega(x) = \mu\left(\{x | p_\omega \le x\}\right)$$

for $0 \le x \le \infty$, where μ is uniform measure on the unit circle. Write also $q_\omega(x)$ as the corresponding density, so that $q_\omega(x)$ is the density of the density of ω. Then it is clear that

$$\int_0^{2\pi} p_\omega(x)^{|\rho|}dx = \int_0^{\infty} x^n q_\omega(x)dx. \tag{15}$$

These quantities correspond to the moments of the measure with density q_ω, which can help us obtain the density q_ω itself (i.e. the density of the density of ω). However, the density p_ω can not be obtained, since we see that any reorganization of its values which do not change its density q_ω will provide the same values in (15).

Note also that theorem 5 gives a very special role to the uniform phase distribution, in the sense that it minimizes the moments of the Vandermonde matrices $\mathbf{V}^H\mathbf{V}$. This follows from (14), since

$$\int_0^{2\pi} p_u(x)^{|\rho|}dx \le \int_0^{2\pi} p_\omega(x)^{|\rho|}dx$$

for any density p_ω. In [22], several examples are provided where the integrals (14) are computed.

V. ω WITH DENSITY SINGULARITIES

The asymptotics of Vandermonde matrices are different when the density of ω has singularities, and depends on the density growth rates near the singular points. It will be clear from these results that one can not perform deconvolution for such ω to obtain the higher order moments of the $\{\mathbf{D}_r(N)\}_{1\le r\le n}$ (only their first moment can be obtained). The asymptotics are first described for ω with atomic density singularities, as this is the simplest case to prove. After this, densities with polynomial growth rates near the singularities are addressed.

Theorem 6: Assume that $p_\omega = \sum_{i=1}^r p_i\delta_{\alpha_i}$ is atomic (where $\delta_{\alpha_i}(x)$ is dirac measure (point mass) at α_i), and denote by $p^{(n)} = \sum_{i=1}^r p_i^n$ the corresponding moments. Then

$$
\begin{aligned}
\lim_{N\to\infty} \quad E[Tr(\quad & \mathbf{D}_1(N)\frac{1}{N}\mathbf{V}^H\mathbf{V}\mathbf{D}_2(N)\frac{1}{N}\mathbf{V}^H\mathbf{V} \\
& \cdots \times \mathbf{D}_n(N)\frac{1}{N}\mathbf{V}^H\mathbf{V})] \\
= \quad & c^{n-1}p^{(n)}\lim_{N\to\infty}\prod_{i=1}^n tr_L\left(\mathbf{D}_i(N)\right).
\end{aligned}
$$

Note here that the non-normalized trace is used.

The proof can be found in appendix G. In particular, theorem 6 states that the asymptotic moments of $\frac{1}{N}\mathbf{V}^H\mathbf{V}$ coincide with the moments of p_ω, up to the scaling factor c^{n-1}. The theorem is of great importance for the estimation of the angles α_i and the point masses p_i in our Vandermonde deconvolution framework. In blind seismic and telecommunication applications, one would like to detect the angles α_i through deconvolution. Unfortunately, theorem 6 tells us that this is impossible, since the $p^{(n)}$ (which are moments which we can find through deconvolution), do not depend on them (this parallels theorem 5, since also there we could not recover the density p_ω itself). Having found the $p^{(n)}$ through deconvolution, one can, however, find the point masses p_i, by solving for $p_1, p_2, ...$ in the Vandermonde equation

$$\begin{pmatrix} p_1 & p_2 & \cdots & p_r \\ p_1^2 & p_2^2 & \cdots & p_r^2 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ \vdots \end{pmatrix} = \begin{pmatrix} p^{(1)} \\ p^{(2)} \\ \vdots \end{pmatrix},$$

even if the number of atoms may be unknown.

The case when the density has non-atomic singularities is more complicated. We provide only the following result, which addresses the case when the density has polynomic growth rate near the singularities.

Theorem 7: Assume that

$$\lim_{x\to\alpha_i}|x-\alpha_i|^s p_\omega(x) = p_i \text{ for some } 0 < s < 1$$

for a set of points $\alpha_1, ..., \alpha_r$, with p_ω continous for $\omega \neq \alpha_1, ..., \alpha_r$. Then

$$\begin{aligned}\lim_{N\to\infty} \quad E[Tr(\quad & \mathbf{D}_1(N)\frac{1}{N^s}\mathbf{V}^H\mathbf{V}\mathbf{D}_2(N)\frac{1}{N^s}\mathbf{V}^H\mathbf{V} \\ & \cdots \times \mathbf{D}_n(N)\frac{1}{N^s}\mathbf{V}^H\mathbf{V})] \\ = \quad & c^{n-1}q^{(n)}\lim_{N\to\infty}\prod_{i=1}^n tr_L\left(\mathbf{D}_i(N)\right)\end{aligned}$$

where

$$\begin{aligned} q^{(n)} \;=\; & \left(2\Gamma(1-s)\cos\left(\tfrac{(1-s)\pi}{2}\right)\right)^n p^{(n)}\times \\ & \int_{[0,1]^n}\prod_{k=1}^n \tfrac{1}{|x_{k-1}-x_k|^{1-s}}, \end{aligned} \qquad (16)$$

and $p^{(n)} = \sum_i p_i^n$. Note here that the non-normalized trace is used.

The proof can be found in appendix H. Also in this case it is only the point masses p_i which can be found through deconvolution, not the singularity locations α_i. Note that the integral in (16) can also be written as an m-fold convolution. Similarly, the definition of $K_{\rho,\omega,N}$ given by (4) can also be viewed as a 2-fold convolution when ρ has two blocks, and as a 3-fold convolution when ρ has three blocks (but not for ρ with more than 3 blocks).

A very useful application of theorem 7 is the case when $\omega = \sin(x)$, with x uniformly distributed. The density will then be of the form $\frac{d\arcsin(\omega)}{d\omega} = \frac{1}{\sqrt{1-\omega^2}}$, which goes to infinity near $\omega = \pm 1$ (which correspond to $x = \pm\pi/2$) at rate $x^{-1/2}$. Theorem 7 thus applies with $s = 1/2$. For this case, however, the "edges" at $\pm\pi/2$ are never reached in practice [22], i.e. we can restrict ω in our analysis to clusters of intervals $U_i[\alpha_i, \beta_i]$ not containing ± 1, for which the results of section IV suffice. In this way, we also avoid the computation of the cumbersome integral (16).

VI. Generalized Vandermonde matrices

Until now, we have been considering Vandermonde matrices where the columns have a uniform distribution of powers. In this section we will look at matrices where this is not the case. Such matrices are called generalized Vandermonde matrices, and are of the form

$$\mathbf{V} = \frac{1}{\sqrt{N}}\begin{pmatrix} e^{-jf(1)\omega_1} & \cdots & e^{-jf(1)\omega_L} \\ e^{-jf(2)\omega_1} & \cdots & e^{-jf(2)\omega_L} \\ \vdots & \ddots & \vdots \\ e^{-jf(N)\omega_1} & \cdots & e^{-jf(N)\omega_L}\end{pmatrix}, \qquad (17)$$

where f is a discrete function taking values in $\{0, ..., N-1\}$, and whose empirical distribution function converges to a function P_f, i.e.

$$\lim_{N\to\infty}\frac{|\{k|f(k)\leq Nx\}|}{N} = P_f(x)$$

for $0 \leq x \leq 1$. We will denote by p_f the density of P_f. We will also consider a second type of generalized Vandermonde matrices, where f in (17) is replaced by a random variable λ taking values in $[0, N)$ (uniformly distributed or not), i.e.

$$\mathbf{V} = \frac{1}{\sqrt{N}}\begin{pmatrix} e^{-j\lambda_1\omega_1} & \cdots & e^{-j\lambda_1\omega_L} \\ e^{-j\lambda_2\omega_1} & \cdots & e^{-j\lambda_2\omega_L} \\ \vdots & \ddots & \vdots \\ e^{-j\lambda_N\omega_1} & \cdots & e^{-j\lambda_N\omega_L}\end{pmatrix}, \qquad (18)$$

with the λ_i mutually independent, and also independent from the ω_j. The integrals $K_{\rho,\omega}$ and $K_{\rho,\omega,N}$ can be defined as in (4) here also. They will, however, additionally depend on f or λ, so they will be denoted by $K_{\rho,\omega,f}$, $K_{\rho,\omega,f,N}$, $K_{\rho,\omega,\lambda}$, and $K_{\rho,\omega,\lambda,N}$.

We first look at the case when ω is uniformly distributed. We explain how to compute the limit distributions based on the results for non-generalized Vandermonde matrices. The equations (36) of appendix B are now replaced by

$$\sum_{k\in\rho_j} f(i_{k-1}) = \sum_{k\in\rho_j} f(i_k). \qquad (19)$$

Since the distribution of f converges to a probability measure with density p_f, we can prove the following:

Theorem 8: The Vandermonde mixed moment expansion coefficients $K_{\rho,u,f}$ can be computed by evaluating integrals over the same volumes as those in the proof of lemma 1 in appendix C, with additional insertions of the density p_f in the integrands. The same applies for $K(\rho, u, \lambda)$.

Proof: We only explain how the proof of this goes for certain ρ, in particular when ρ is noncrossing. The equations (37) are the same also for generalized Vandermonde matrices with uniformly distributed phases, with the difference that the variables $x_1, ..., x_\rho$ now all have the density p_f. When ρ is noncrossing $K_{\rho,u,f}$ becomes

$$\prod_{i=1}^{n+1-|\rho|}\int_0^1 p_f(x)^{|K(\rho)_i|}dx, \qquad (20)$$

where we have used the observation from appendix B that the free variables in the equation system (37) are given by the block structure in the Kreweras complement $K(\rho)$ [21]. In (20) we have also used that $|K(\rho)| = n+1-|\rho|$, and have denoted the blocks of $K(\rho)$ by $K(\rho)_i$.

As another example, $K_{\{\{1,3\},\{2,4\}\},u,f}$ becomes the sum of

$$\begin{aligned}&\int_0^1\int_0^{1-x_1}\int_0^{x_1+x_3} \\ &\quad p_f(x_1)p_f(x_2)p_f(x_3)p_f(x_1+x_3-x_2)dx_2dx_3dx_1\end{aligned} \qquad (21)$$

and

$$\begin{aligned}&\int_0^1\int_{1-x_1}^1\int_{x_1+x_3-1}^1 \\ &\quad p_f(x_1)p_f(x_2)p_f(x_3)p_f(x_1+x_3-x_2)dx_2dx_3dx_1,\end{aligned} \qquad (22)$$

according to the integrals computed in appendix C. The other $K_{\rho,u,f}$ are computed by inserting densities in the integrand similarly: For each ρ we compute the reduced row echelon form of the equation system (37), and insert the dependence equations from the reduced form (such as $x_4 = x_1 + x_3 - x_2$ in the above) into the integrand variables as above. ■

That the same result applies when matrices of the form (18) is used, is apparent from the law of large numbers.

These "generalized" integrals are easily computed based on the evaluation of the integrals in appendix C, for cases when p_f is a polynomial.

Similar reasoning applies when ω has a continous density: Theorem 5 can be used in this case also, with the change that the integrals for $K_{\rho,u}$ are replaced with integrals with additional insertions of the density p_ω, as explained in theorem 5.

We will not consider generalized Vandermonde matrices with density singularities.

VII. The joint distribution of independent Vandermonde matrices

In the case when many independent random Vandermonde matrices are involved, the following holds:

Theorem 9: Assume that the $\{\mathbf{D}_r(N)\}_{1\leq r\leq n}$ have a joint limit distribution as $N \to \infty$. Assume also that $\mathbf{V}_1, \mathbf{V}_2, ...$ are independent Vandermonde matrices with the same phase distribution ω, and that the density of ω is continous. Then the limit

$$\begin{array}{ll}\lim_{N\to\infty} E[tr_L(& \mathbf{D}_1(N)\mathbf{V}_{i_1}^H\mathbf{V}_{i_2}\mathbf{D}_2(N)\mathbf{V}_{i_2}^H\mathbf{V}_{i_3} \\ & \cdots \times \mathbf{D}_n(N)\mathbf{V}_{i_n}^H\mathbf{V}_{i_1})]\end{array}$$

also exists when $\frac{L}{N} \to c$, and equals

$$\sum_{\rho\leq\sigma\in\mathcal{P}(n)} K_{\rho,\omega} c^{|\rho|} D_\rho, \tag{23}$$

where σ is the partition where k and j are in the same block if and only if $i_k = i_j$.

For the proof of theorem 9 and the next results, we define σ_j to be the blocks of σ, i.e.

$$\sigma_j = \{k | i_k = j\}.$$

Proof: Note that theorem 5 guarantees that the limit $K_{\rho,\omega} = \lim_{N\to\infty} K_{\rho,\omega,N}$ exists. The partition ρ simply is a grouping of random variables into independent groups. It is therefore impossible for a block in ρ to contain elements from both σ_1 and σ_2, so that any block is contained in either σ_1 or σ_2. As a consequence, $\rho \leq \sigma$. ■

Corollary 1: The first three mixed moments

$$M_n = \lim_{N\to\infty} E\left[tr_L\left(\left(\mathbf{V}_1^H\mathbf{V}_2\mathbf{V}_2^H\mathbf{V}_1\right)^n\right)\right]$$

of independent Vandermonde matrices $\mathbf{V}_1, \mathbf{V}_2$ are given by

$$\begin{aligned} M_1 &= I_2 \\ M_2 &= \frac{2}{3}I_2 + 2I_3 + I_4 \\ M_3 &= \frac{11}{20}I_2 + 4I_3 + 9I_4 + 6I_5 + I_6, \end{aligned}$$

where

$$I_k = (2\pi)^{|\rho|-1}\left(\int_0^{2\pi} p_\omega(x)^{|\rho|}dx\right).$$

In particular, when the phases are uniformly distributed, the first three mixed moments are given by

$$\begin{aligned} M_1 &= 1 \\ M_2 &= \frac{11}{3} \\ M_3 &= \frac{411}{20} \end{aligned}$$

Proof: This follows in the same way as theorem 3 is proved from lemma 1, by only considering ρ which are less than σ, and also by using theorem 5. σ are for the listed moments $\{\{1\},\{2\}\}$, $\{\{1,3\},\{2,4\}\}$, and $\{\{1,3,5\},\{2,4,6\}\}$, respectively. ■

The results here can also be extended to the case with independent Vandermonde matrices with different phase distributions:

Theorem 10: Assume that $\{\mathbf{V}_i\}_{1\leq i\leq s}$ are independent Vandermonde matrices, where $\mathbf{V}_i$ has continous phase distribution ω_i. Denote by p_{ω_i} the density of ω_i. Then equation (23) still holds, with $K_{\rho,\omega}$ replaced by

$$K_{\rho,u}(2\pi)^{|\rho|-1}\int_0^{2\pi}\prod_{i=1}^s p_{\omega_i}(x)^{|\rho_i|}dx,$$

where ρ_i is the partition of σ_i consisting of the blocks of ρ contained in σ_i,

The proof is omitted, as it is a straightforward extension of the proofs of theorems 5 and 9. Until now, we have not treated mixed moments of the form

$$\mathbf{D}_1(N)\mathbf{V}_{i_2}\mathbf{V}_{i_2}^H\mathbf{D}_2(N)\mathbf{V}_{i_3}\mathbf{V}_{i_3}^H\cdots\times\mathbf{D}_n(N)\mathbf{V}_{i_1}\mathbf{V}_{i_1}^H,$$

which are the same as the mixed moments of theorem 9 except for the position of the $\mathbf{D}_i(N)$. We will not go into depths on this, but only remark that this case can be treated in the same vein as generalized Vandermonde matrices by replacing the density p_f (or p_λ in case of continous generalized Vandermonde matrices) with functions $p_{D_i}(x)$ defined by $p_{D_i}(x) = \mathbf{D}_i(N)(\lfloor Lx\rfloor, \lfloor Lx\rfloor)$ for $0 \leq x \leq 1$. This also covers the case of mixed moments of independent, generalized Vandermonde matrices (and, in fact, there are no restrictions on the horizontal and vertical phase densities p_{ω_i} and p_{λ_j} for each matrix. They may all be different). The proof for this is straightforward.

VIII. Discussion

We have already explained that one can perform deconvolution with Vandermonde matrices in a similar way to how one can perform deconvolution for Gaussian matrices. We have, however, also seen that there are many differences.

A. Convergence rates

In [15], almost sure convergence of Gaussian matrices was shown by proving exact formulas for the distribution of lower order Gaussian matrices. These deviated from their limits by terms of the form $1/L^2$. In theorem 4, we see that terms of the form $1/L$ are involved, which indicates that we can not hope for almost sure convergence of Vandermonde matrices. There is no reason why Vandermonde matrices should have the almost sure convergence property, due to their very different degree of randomness when compared to Gaussian matrices. Figures 1, 2 show the speed of convergence of the moments of Vandermonde matrices (with uniformly distributed phases) towards the asymptotic moments as the matrix dimensions grow, and as the number of samples grow. The differences between the asymptotic moments and the exact moments are

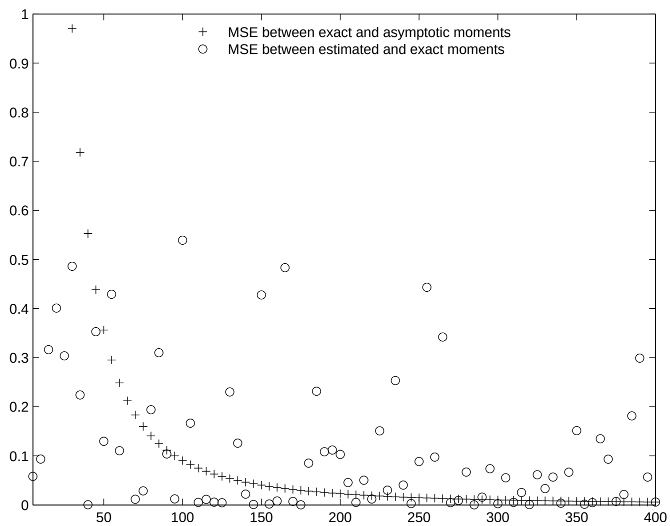

Fig. 1. MSE of the first 4 estimated moments from the exact moments for 80 samples for varying matrix sizes, with $N = L$. Matrices are on the form $\mathbf{V}^H\mathbf{V}$ with $\mathbf{V}$ a Vandermonde matrix with uniformly distributed phases. The MSE of the first 4 exact moments from the asymptotic moments is also shown.

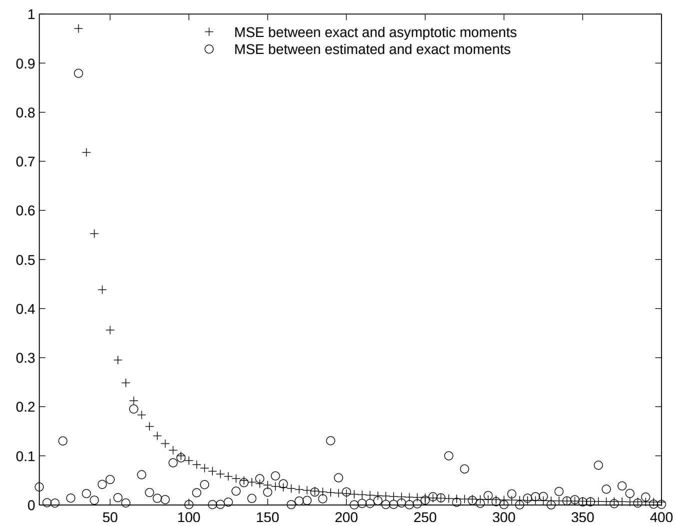

Fig. 2. MSE of the first 4 moments from the actual moments for 320 samples for varying matrix sizes, with $N = L$. Matrices are on the form $\mathbf{V}^H\mathbf{V}$ with $\mathbf{V}$ a Vandermonde matrix with uniformly distributed phases. The MSE of the moments and the asymptotic moments is also shown.

also shown. To be more precise, the MSE of figures 1 and 2 is computed as follows:

1) K samples $\mathbf{V}_i$ are independently generated using (1).
2) The 4 first sample moments $\hat{m}_{ji} = \frac{1}{L}tr_n\left(\left(\mathbf{V}_i^H\mathbf{V}_i\right)^j\right)$ $(1 \leq j \leq 4)$ are computed from the samples.
3) The 4 first estimated moments $\hat{M}_j$ are computed as the mean of the sample moments, i.e. $\hat{M}_j = \frac{1}{K}\sum_{i=1}^K \hat{m}_{ji}$.
4) The 4 first exact moments E_j are computed using theorem 4.
5) The 4 first asymptotic moments A_j are computed using theorem 3.
6) The mean squared error (MSE) of the first 4 estimated moments from the exact moments is computed as $\sum_{j=1}^4 \left(\hat{M}_j - E_j\right)^2$.
7) The MSE of the first 4 exact moments from the asymptotic moments is computed as $\sum_{j=1}^4 (E_j - A_j)^2$.

Figures 1 and 2 are in sharp contrast with Gaussian matrices, as shown in figure 3. First of all, it is seen that the asymptotic moments can be used just as well instead of the exact moments (for which expressions can be found in [23]), due to the $O(1/N^2)$ convergence of the moments. Secondly, it is seen that only 5 samples were needed to get a reliable estimate for the moments.

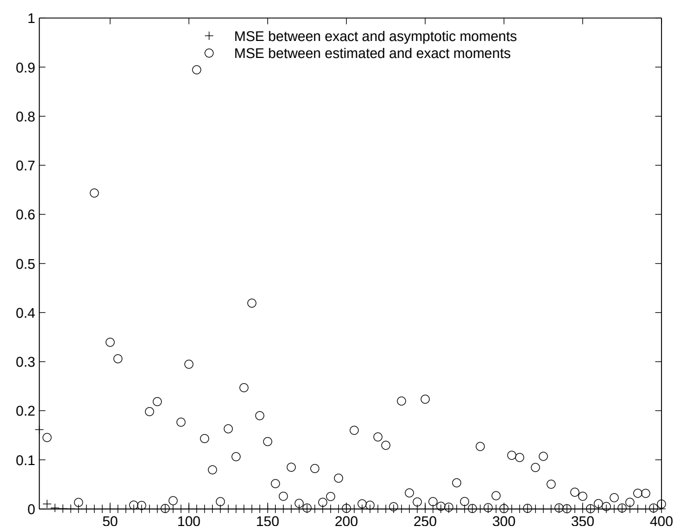

Fig. 3. MSE of the first 4 moments from the actual moments for 5 samples for varying matrix sizes, with $N = L$. Matrices are on the form $\frac{1}{N}\mathbf{X}\mathbf{X}^H$ with $\mathbf{X}$ a complex standard Gaussian matrix. The MSE of the moments and the asymptotic moments is also shown.

B. Inequalities between moments of Vandermonde matrices and moments of known distributions

We will state an inequality involving the moments of Vandermonde matrices, and the moments of known distributions from probability theory. The classical Poisson distribution with rate λ and jump size α is defined as the limit of

$$\left(\left(1-\frac{\lambda}{n}\right)\delta_0 + \frac{\lambda}{n}\delta_\alpha\right)^{*N}$$

as $n \to \infty$ [21]. For our analysis, we will only need the classical Poisson distribution with rate c and jump size 1. We will denote this quantity by ν_c. The free Poisson distribution with rate λ and jump size α is defined similarly as the limit of

$$\left(\left(1-\frac{\lambda}{n}\right)\delta_0 + \frac{\lambda}{n}\delta_\alpha\right)^{\boxplus N}$$

as $n \to \infty$, where $\boxplus$ is the free probability counterpart of classical additive convolution [21], [16]. For our analysis, we will only need the free Poisson distribution with rate $\frac{1}{c}$ and jump size c. We will denote this quantity by μ_c. μ_c is the same as the better known Marčhenko Pastur law, i.e. it has the density [16]

$$f^{\mu_c}(x) = (1-\frac{1}{c})^+\delta_0(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi c x}, \quad (24)$$

where $(z)^+ = \max(0,z)$, $a=(1-\sqrt{c})^2$, $b=(1+\sqrt{c})^2$. Since the classical (free) cumulants of the classical (free) Poisson distribution are $\lambda\alpha^n$ [21], we see that the (classical) cumulants of ν_c are $c,c,c,c,...$, and that the (free) cumulants of μ_c are $1,c,c^2,c^3,....$ In other words, if a_1 has the distribution μ_c, then

$$\begin{aligned}\phi(a_1^n) &= \sum_{\rho\in NC(n)} c^{n-|\rho|} = \sum_{\rho\in NC(n)} c^{|K(\rho)|-1}\\ &= \sum_{\rho\in NC(n)} c^{|\rho|-1}.\end{aligned} \quad (25)$$

Here we have used the Kreweras complementation map, which is an order-reversing isomorphism of $NC(n)$ which satisfies $|\rho|+|K(\rho)| = n+1$ (here ϕ is the expectation in a non-commutative probability space). Also, if a_2 has the distribution ν_c, then

$$E(a_2^n) = \sum_{\rho\in\mathcal{P}(n)} c^{|\rho|}. \quad (26)$$

We immediately recognize the $c^{|\rho|-1}$-entry of theorem 1 in (25) and (26) (except for an additional power of c in (26)). Combining theorem 2 with $\mathbf{D}_1(N) = \cdots = \mathbf{D}_n(N) = \mathbf{I}_N$, (25), and (26), we thus get the following corollary to theorem 2:

Corollary 2: Assume that $\mathbf{V}$ has uniformly distributed phases. Then the limit moment

$$M_n = \lim_{N\to\infty} E\left[tr_L\left(\left(\mathbf{V}^H\mathbf{V}\right)^n\right)\right]$$

satsifies the inequality

$$\phi(a_1^n) \leq M_n \leq \frac{1}{c}E(a_2^n),$$

where a_1 has the distribution μ_c of the Marčhenko Pastur law, and a_2 has the Poisson distribution ν_c. In particular, equality occurs for $m=1,2,3$ and $c=1$ (since all partitions are noncrossing for $m=1,2,3$).

Corollary 2 thus states that the moments of Vandermonde matrices with uniformly distributed phases are bounded above and below by the moments of the classical and free Poisson distributions, respectively. The different Poisson distributions enter here because their (free and classical) cumulants resemble the $c^{|\rho|-1}$-entry in theorem 1, where we also can use that $K_{\rho,u} = 1$ if and only if ρ is noncrossing to get a connection with the Marčhenko Pastur law. To see how close the asymptotic Vandermonde moments are to these upper and lower bounds, the following corollary to theorem 3 contains the first moments:

Corollary 3: When $c=1$, the limit moments

$$M_n = \lim_{N\to\infty} E\left[tr_L\left(\left(\mathbf{V}^H\mathbf{V}\right)^n\right)\right],$$

the moments fp_n of the Marčhenko Pastur law μ_1, and the moments p_n of the Poisson distribution ν_1 satisfy

$$\begin{array}{lllll} fp_4 = 14 & \leq & M_4 = \frac{44}{3} \approx 14.67 & \leq & p_4 = 15\\ fp_5 = 42 & \leq & M_5 = \frac{146}{3} \approx 48.67 & \leq & p_5 = 52\\ fp_6 = 132 & \leq & M_6 = \frac{3571}{20} \approx 178.55 & \leq & p_6 = 203\\ fp_7 = 429 & \leq & M_7 = \frac{2141}{3} \approx 713.67 & \leq & p_7 = 877.\end{array}$$

The first three moments coincide for the three distributions, and are $1,2$, and 5, respectively.

The numbers fp_n and p_n are simply the number of partitions in $NC(n)$ and $\mathcal{P}(n)$, respectively. The number of partitions in $NC(n)$ equals the Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$ [21], so they are easily computed. The number of partitions of $\mathcal{P}(n)$ are also known as the Bell numbers B_n [21]. They can easily be computed from the recurrence relation

$$B_{n+1} = \sum_{k=0}^n B_k\binom{n}{k}.$$

It is not known whether the limiting distribution of our Vandermonde matrices has compact support. Corollary 3 does not help us in this respect, since the Marčhenko Pastur law has compact support, and the classical Poisson distribution has not. In figure 4, the mean eigenvalue distribution of 640 samples of a 1600×1600 Vandermonde matrix with uniformly distributed phases is shown. While the Poisson distribution ν_1 is purely atomic and has masses at 0, 1, 2, and 3 which are e^{-1}, e^{-1}, $e^{-1}/2$, and $e^{-1}/6$ (the atoms consist of all integer multiples), the Vandermonde histogram shows a more continous eigenvalue ditribution, with the peaks which the Poisson distribution has at integer multiples clearly visible here as well (the peaks are not as sharp though). We remark that the support of $\mathbf{V}^H\mathbf{V}$ goes all the way up to N, but lies within $[0,N]$. It is also unknown whether the peaks at integer multiples in the Vandermonde histogram grow to infinity as we let $N\to\infty$. From the histogram, only the peak at 0 seems to be of atomic nature. In figures 5 and 6, the same histogram is shown for 1600×1200 (i.e. $c=0.75$) and 1600×800 (i.e. $c=0.5$) Vandermonde matrices, respectively. It should come as no surprise that the effect of decreasing c is stretching the eigenvalue density vertically, and compressing it horizontally. just as the case for the different Marčhenko Pastur laws. Eigenvalue histograms for Gaussian matrices which in the limit give the corresponding (in the sense of corollary 2) Marčhenko Pastur laws for figures 5 (i.e. $\mu_{0.75}$) and 6 (i.e. $\mu_{0.5}$), are shown in figures 7 and 8.

C. *Deconvolution*

Deconvolution with Vandermonde matrices (as stated in (6) in theorem 1) differs from the Gaussian deconvolution counterpart [21] in the sense that there is no multiplicative [21] structure involved, since $K_{\rho,\omega}$ is not multiplicative in ρ. The Gaussian equivalent of theorem 3 (i.e. $\mathbf{V}^H\mathbf{V}$ replaced with $\frac{1}{N}\mathbf{X}\mathbf{X}^H$, with $\mathbf{X}$ an $L\times N$ complex, standard, Gaussian

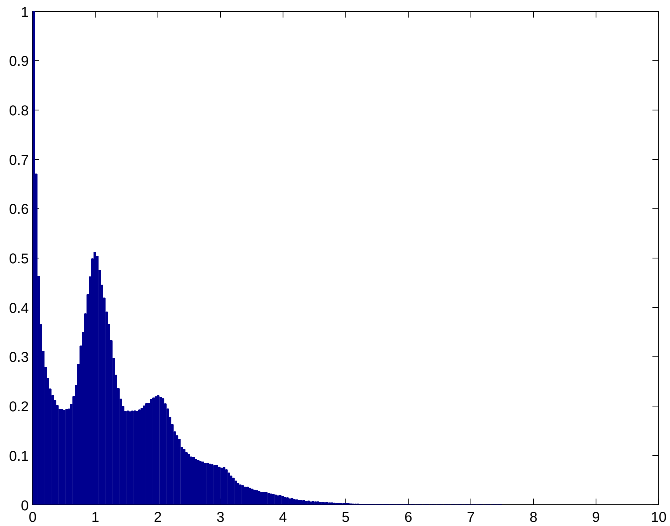

Fig. 4. Histogram of the mean eigenvalue distribution of 640 samples of $\mathbf{V}^H\mathbf{V}$, with $\mathbf{V}$ a 1600×1600 Vandermonde matrix with uniformly distributed phases.

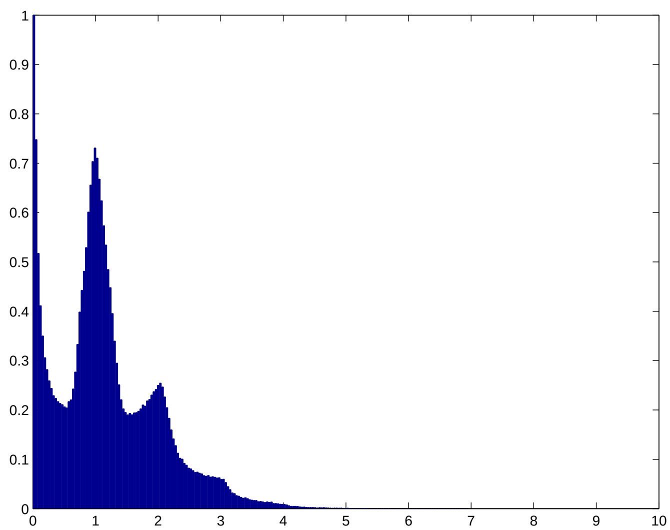

Fig. 5. Histogram of the mean eigenvalue distribution of 640 samples of $\mathbf{V}^H\mathbf{V}$, with $\mathbf{V}$ a 1600×1200 Vandermonde matrix with uniformly distributed phases.

matrix) is

$$\begin{aligned}
m_1 &= d_1 \\
m_2 &= d_2 + d_1^2 \\
m_3 &= d_3 + 3d_2d_1 + d_1^3 \\
m_4 &= d_4 + 4d_3d_1 + 3d_2^2 + 6d_2d_1^2 + d_1^4 \\
m_5 &= d_5 + 5d_4d_1 + 5d_3d_2 + 10d_3d_1^2 + \\
& \quad 10d_2^2d_1 + 10d_2d_1^3 + d_1^5 \\
m_6 &= d_6 + 6d_5d_1 + 6d_4d_2 + 15d_4d_1^2 + \\
& \quad 3d_3^2 + 30d_3d_2d_1 + 20d_3d_1^3 + \\
& \quad 5d_2^3 + 10d_2^2d_1^2 + 15d_2d_1^4 + d_1^6 \\
m_7 &= d_7 + 7d_6d_1 + 7d_5d_2 + 21d_5d_1^2 + \\
& \quad 7d_4d_3 + 42d_4d_2d_1 + 35d_4d_1^3 + \\
& \quad 21d_3^2d_1 + 21d_3d_2^2 + 105d_3d_2d_1^2 + \\
& \quad 35d_3d_1^4 + 35d_2^3d_1 + 70d_2^2d_1^3 + \\
& \quad 21d_2d_1^5 + d_1^7,
\end{aligned} \tag{27}$$

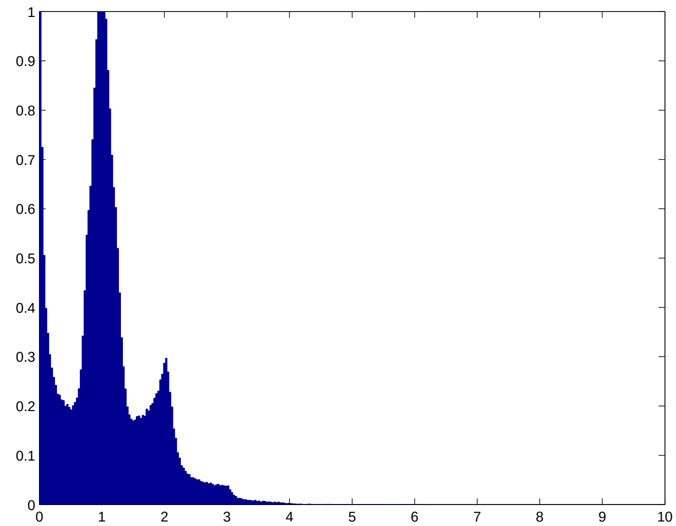

Fig. 6. Histogram of the mean eigenvalue distribution of 640 samples of $\mathbf{V}^H\mathbf{V}$, with $\mathbf{V}$ a 1600×800 Vandermonde matrix with uniformly distributed phases.

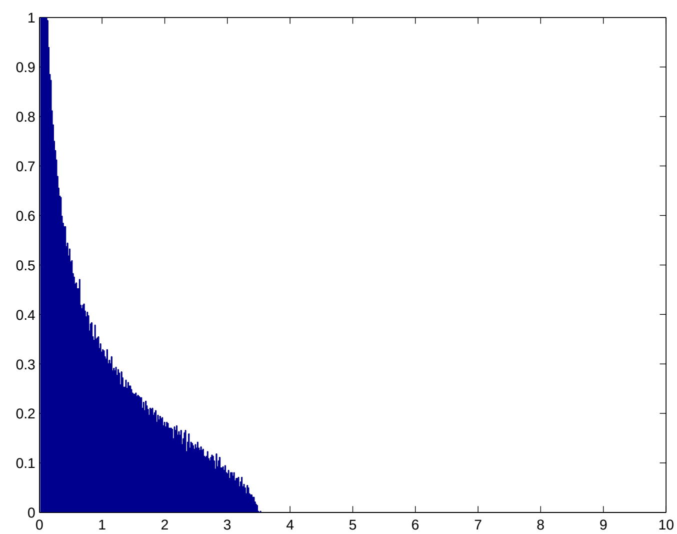

Fig. 7. Histogram of the mean eigenvalue distribution of 20 samples of $\frac{1}{N}\mathbf{X}\mathbf{X}^H$, with $\mathbf{X}$ an $L \times N = 1200 \times 1600$ complex, standard, Gaussian matrix.

(where the m_i and the d_i are computed as in (9) by scaling the respective moments by c). This follows immediately from asymptotic freeness, and from the fact that $\frac{1}{N}\mathbf{X}\mathbf{X}^H$ converges to the Marčhenko Pastur law μ_c. In particular, when all $\mathbf{D}_i(N) = I_L$ and $c = 1$, we obtain the limit moments: 1,2,5,14,42,132,429, which also were listed in corollary 3. One can also write down a Gaussian equivalent to the fluctuations of Vandermonde matrices (13) (fluctuations of Gaussian ma-

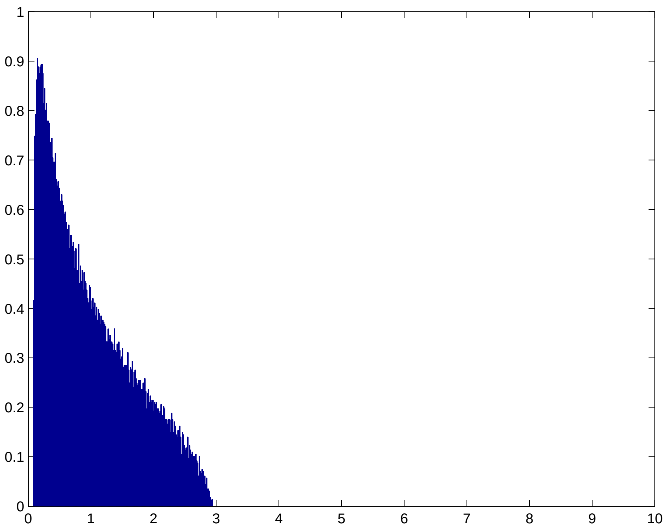

Fig. 8. Histogram of the mean eigenvalue distribution of 20 samples of $\frac{1}{N}\mathbf{X}\mathbf{X}^H$, with $\mathbf{X}$ an $L \times N = 800 \times 1600$ complex, standard, Gaussian matrix.

trices are handled more thoroughly in [24]). These are

$$\begin{array}{l} E\left[\left(tr_n\left(\mathbf{D}(N)\frac{1}{N}\mathbf{X}\mathbf{X}^H\right)\right)^2\right] \\ = (tr_n(\mathbf{D}(N))^2 + \frac{1}{nN}tr_n(\mathbf{D}(N)^2) \\ E\left[\left(tr_n\left(\mathbf{D}(N)\frac{1}{N}\mathbf{X}\mathbf{X}^H\right)\right)^n\right] \\ = (tr_n(\mathbf{D}(N))^n + O(N^{-2}) \\ E\left[tr_n\left(\mathbf{D}(N)\frac{1}{N}\mathbf{X}\mathbf{X}^H\right)tr_n\left(\left(\mathbf{D}(N)\frac{1}{N}\mathbf{X}\mathbf{X}^H\right)^2\right)\right] \\ = tr_n(\mathbf{D}(N))tr_n(\mathbf{D}(N)^2) + O(N^{-2}). \end{array} \tag{28}$$

These equations can be proved using the same combinatorical methods as in [23]. Only the first equation is here stated as an exact expression. The second and third equations also have exact counterparts, but their computations are more involved. Similarly, one can write down a Gaussian equivalent to theorem 4 for the exact moments. For the first three moments (the fourth moment is dropped, since this is more involved), these are

$$\begin{array}{rcl} m_1 & = & d_1 \\ m_2 & = & d_2 + d_1^2 \\ m_3 & = & \left(1 + N^{-2}\right)d_3 + 3d_1d_2 + d_1^3. \end{array}$$

This follows from a careful count of all possibilities after the matrices have been multiplied together (for this, see also [23], where one can see that the restriction that the matrices $\mathbf{D}_i(N)$ are diagonal can be dropped in the Gaussian case). It is seen, contrary to theorem 4 for Vandermonde matrices, that the second exact moment equals the second asymptotic moment from (27), and also that the convergence is faster (i.e. $O(n^{-2})$) for the third moment (this will also be the case for higher moments).

The two types of (de)convolution also differ in how they can be computed in practice. In [19], an algorithm for free convolution with the Marčhenko Pastur law was sketched. A similar algorithm may not exist for Vandermonde convolution. However, Vandermonde convolution can be subject to numerical approximation: To see this, note first that theorem 5 splits the numerics into two parts: The approximation of the integrals $\int p_\omega(x)^{|\rho|}dx$, and the approximation of the $K_{\rho,u}$. A strategy for obtaining the latter quantities could be to randomly generate many numbers between 0 and 1 and estimate the volume as the ratio of the solutions which satisfy (37) in appendix B. Implementations of the various Vandermonde convolution variants given in this paper can be found in [25].

In practice, one often has a random matrix model where independent Gaussian and Vandermonde matrices are both present. In such cases, it should be possible to combine the individual results for both of them. In [22], examples on how this can be done are presented.

IX. Conclusion and further directions

We have shown how asymptotic moments of random Vandermonde matrices can be computed analytically, and treated many different cases. Vandermonde matrices with uniformly distributed phases proved to be the easiest case and was given separate treatment, and it was shown how the case with more general phases could be expressed in terms of the case of uniformly distributed phases. The case where the phase distribution has singularities was also given separate treatment, as this case displayed different asymptotic behaviour. Also mixed moments of independent Vandermonde matrices were computed, as well as the moments of generalized Vandermonde matrices. In addition to the general asymptotic expressions stated, exact expressions for the first moments of Vandermonde matrices with uniformly distributed phases were also stated.

Throughout the paper, we assumed that only diagonal matrices were involved in mixed moments of Vandermonde matrices. The case of non-diagonal matrices is harder to address, and should be addressed in future research. The analysis of the support of the eigenvalues is also of importance, as well as the behavior of the maximum and minimum eigenvalue. The methods presented in this paper can not be used directly to obtain explicit expressions for the asymptotic mean eigenvalue distribution, so this is also a case for future research. A way of attacking this problem could be to develop for Vandermonde matrices analytic counterparts to what one has in free pobability (such as the R- and S-transform and their connection with the Stieltjes transform).

Finally, another case for future research is the asymptotic behaviour of Vandermonde matrices when the matrix entries lie outside the unit circle. The asymptotics are very different in this case. The choice of Vandermonde matrix entries on the unit circle was applied for this paper since the asymptotic behaviour is more easily addressed in this case.

Appendix A
The proof of theorem 1

We can write

$$E\left[tr_L\left(\mathbf{D}_1(N)\mathbf{V}^H\mathbf{V}\mathbf{D}_2(N)\mathbf{V}^H\mathbf{V}\cdots\mathbf{D}_n(N)\mathbf{V}^H\mathbf{V}\right)\right] \tag{29}$$

as

$$L^{-1} \sum_{\substack{i_1,\ldots,i_n \\ j_1,\ldots,j_n}} E(\begin{array}{l} \mathbf{D}_1(N)(j_1,j_1)\mathbf{V}^H(j_1,i_2)\mathbf{V}(i_2,j_2) \\ \mathbf{D}_2(N)(j_2,j_2)\mathbf{V}^H(j_2,i_3)\mathbf{V}(i_3,j_3) \\ \vdots \\ \mathbf{D}_2(N)(j_n,j_n)\mathbf{V}^H(j_n,i_1)\mathbf{V}(i_1,j_1)) \end{array} \tag{30}$$

The $(j_1,...,j_n)$ give rise to a partition ρ of $\{1,...,n\}$, where each block ρ_j consists of equal values, i.e.

$$\rho_j = \{k | j_k = j\}.$$

Write

$$\rho_j = \{\rho_{j1}, \rho_{j2}, ..., \rho_{j|\rho_j|}\}.$$

When $(j_1,...,j_n)$ give rise to ρ, we see that since

$$j_{\rho_{j1}} = j_{\rho_{j2}} = \cdots = j_{\rho_{j|\rho_j|}},$$

we also have that

$$\omega_{j_{\rho_{j1}}} = \omega_{j_{\rho_{j2}}} = \cdots = \omega_{j_{\rho_{j|\rho_j|}}},$$

and we will denote their common value by ω_{ρ_j} as in definition 2. With this in mind, it is straightforward to verify that (30) can be written as

$$\begin{array}{l} \sum_{\rho\in\mathcal{P}(n)} \\ \sum_{(i_1,\ldots,i_n)} \\ \sum_{\substack{(j_1,\ldots,j_n) \\ \text{giving rise to } \rho}} \\ \quad N^{-n}L^{-1} \\ \quad \prod_{k=1}^{|\rho|} E\left(e^{j\left(\sum_{k\in\rho_j} i_{k-1} - \sum_{k\in\rho_j} i_k\right)\omega_{\rho_k}}\right) \\ \quad \mathbf{D}_1(N)(j_1,j_1)\mathbf{D}_2(N)(j_2,j_2) \\ \quad \cdots \times \mathbf{D}_n(N)(j_n,j_n) \end{array} \tag{31}$$

We will in the following switch between the form (31) and the form

$$\begin{array}{l} \sum_{\rho\in\mathcal{P}(n)} \\ \sum_{\substack{(j_1,\ldots,j_n) \\ \text{giving rise to } \rho}} \\ \sum_{(i_1,\ldots,i_n)} \\ \quad N^{|\rho|-n-1}c^{|\rho|-1}L^{-|\rho|} \\ \quad \prod_{k=1}^{n} E\left(e^{j(\omega_{b(k-1)}-\omega_{b(k)})i_k}\right) \\ \quad \mathbf{D}_1(N)(j_1,j_1)\mathbf{D}_2(N)(j_2,j_2) \\ \quad \cdots \times \mathbf{D}_n(N)(j_n,j_n), \end{array} \tag{32}$$

where we also have reorganized the powers of N and L in (31), and changed the order of summation (i.e. summed over the different $i_1,...,i_n$ first). (32) will also be written

$$\begin{array}{l} \sum_{\rho\in\mathcal{P}(n)} \\ \sum_{\substack{(j_1,\ldots,j_n) \\ \text{giving rise to } \rho}} \\ \quad c^{|\rho|-1}L^{-|\rho|}K_{\rho,\omega,N} \\ \quad \mathbf{D}_1(N)(j_1,j_1)\mathbf{D}_2(N)(j_2,j_2) \\ \quad \cdots \times \mathbf{D}_n(N)(j_n,j_n), \end{array} \tag{33}$$

where $K_{\rho,\omega,N}$ is defined in theorem 1. This form is obtained from (32) by using the geometric sum formula.

The notation for a joint limit distribution simplifies (32). Indeed, add to (32) for each ρ the terms

$$\begin{array}{l} \sum_{\rho'\in\mathcal{P}(n),\rho'>\rho} \\ \sum_{\substack{(j_1,\ldots,j_n) \\ \text{giving rise to } \rho'}} \\ \quad c^{|\rho|-1}L^{-|\rho|}K_{\rho,\omega,N} \\ \quad \mathbf{D}_1(N)(j_1,j_1)\mathbf{D}_2(N)(j_2,j_2) \\ \quad \cdots \times \mathbf{D}_n(N)(j_n,j_n) \end{array} \tag{34}$$

These go to 0 as $N\to\infty$, since they are bounded by

$$c^{|\rho|-1}L^{-|\rho|}K_{\rho,\omega,N}L^{|\rho'|} = K_{\rho,\omega,N}c^{|\rho|-1}L^{|\rho'|-|\rho|} = O(L^{-1}).$$

After this addition, the limit of (33) can be written

$$\sum_{\rho\in\mathcal{P}(n)} c^{|\rho|-1}K_{\rho,\omega}D_\rho, \tag{35}$$

which is what we had to show.

APPENDIX B
THE PROOF OF THEOREM 2

Note that

$$E\left(e^{j\left(\sum_{k\in\rho_j} i_{k-1} - \sum_{k\in\rho_j} i_k\right)\omega_{\rho_j}}\right) = 0$$

when

$$\sum_{k\in\rho_j} i_{k-1} \neq \sum_{k\in\rho_j} i_k,$$

and 1 if

$$\sum_{k\in\rho_j} i_{k-1} = \sum_{k\in\rho_j} i_k. \tag{36}$$

We thus define

$$S_{\rho,N} = \{i_1,...,i_n\} | \sum_{k\in\rho_j} i_{k-1} = \sum_{k\in\rho_j} i_k \forall j\in\{1,...,|\rho|\},$$

and $|S_{\rho,N}|$ to be the cardinality of $S_{\rho,N}$. With this definition in place, it is obvious that

$$K_{\rho,u} = \lim_{N\to\infty} K_{\rho,u,N} = \lim_{N\to\infty} \frac{1}{N^{n+1-|\rho|}}|S(\rho,N)|$$

Finding the limit distribution thus boils down to finding $|S_{\rho,N}|$, which is equivalent to finding the number of solutions to equations of the form (36), where the variables are integers constrained to lie between 1 and N. For lemma 1 we will compute $|S_{\rho,N}|$ for certain ρ of lower order. To prove theorem 2, we need not compute specific $|S_{\rho,N}|$.

First we explain why $K_{\rho,u}\leq 1$. It is clear that $|S_{\rho,N}|$ is the number of integer solutions $(i_1,...,i_n)$ between 1 and N to a system of the form $\mathbf{Ai}=\mathbf{0}$, where $\mathbf{i}=(i_1,...,i_n)$, and $\mathbf{A}$ is $|\rho|\times n$, with all entries being -1, 0 or 1. Also, it is clear from (36) that each column of A contains exactly one -1 and one 1, or contains just zeroes. Such a matrix has rank $|\rho|-1$, as can be found through elementary row reduction. Hence, there are $|\rho|-1$ pivot columns in $\mathbf{A}$, so that there are $n+1-|\rho|$ free variables among $(i_1,...,i_n)$ in the solution set of $\mathbf{Ai}=\mathbf{0}$. Therefore, $|S_{\rho,N}|\leq N^{n+1-|\rho|}$, which proves that $K_{\rho,u}\leq 1$.

Also, by dividing the equations (36) by N, and letting N go to infinity, we see that $K_{\rho,u}$ can alternatively be expressed as the volume of the solution set of

$$\sum_{k\in\rho_j} x_{k-1} = \sum_{k\in\rho_j} x_k. \tag{37}$$

as a volume in $\mathbb{R}^{n+1-|\rho|}$ (i.e. the volume is computed after expressing the remaining $|\rho|-1$ variables in the $n+1-|\rho|$ free variables). Since $1 \leq i_k \leq N$, we have that $0 \leq x_k \leq 1$, so that the volume lies within $[0,1]^{n+1-|\rho|}$, and is bounded by a finite set of hyperplanes due to (37). The integral for such a volume can be expressed for any given ρ (however complex). Although we will only compute a few of these integrals directly, it is clear that the integral computes to a rational number greater than 0 but less than 1, since only polynomials are involved in the integration procedure, and since only 0 and 1 may be constant upper or lower bounds in the integrals. From these integrals it is also clear that the integral is equal to 1 if and only if the reduced row echelon form of (37) only contains rows with 2 nonzero entries (these 2 entries will then be 1 and -1 respectively), after removing the rows which have only 0's. This corresponds to solutions where each constrained variable is equal to one of the free variables. For the rest of the proof it therefore suffices to show that such a solution set occurs if and only if the partition ρ is noncrossing.

If ρ is noncrossing, there exists a block ρ_1 (after renumbering the blocks if necessary) which consists of a single interval of numbers, say $\{r, r+1, ..., r+|\rho_1|\}$. This block's equation in (36) is easily seen to imply that $i_{r-1} = i_{r+|\rho_1|}$. Also, $i_r, ..., i_{r+|\rho_1|-1}$ can be chosen arbitrarily. Therefore, this block gives rise to $|\rho_1|-1$ free variables. We now add together the equation for the block ρ_1, and the equation for the block ρ_2 which contains $r+|\rho_1|+1$ (again after renumbering the blocks if necessary), and replaces the two rows with this sum. Columns $r, ..., r+|\rho_1|$ are easily seen to contain only 0's, so that these can be removed from our equation system (since we are just interested in counting the number of free variables in the solution set. These removed variables gave rise to $|\rho_1|-1$ free variables). The new equation system corresponds to the equation system for another noncrossing partition of $\{1, ..., n-|\rho_1|\}$ (created by merging the blocks ρ_1 and ρ_2), with $|\rho|-1$ blocks. The step where we find a block which is an interval can now be repeated to combine two more blocks to merge, and this process can be repeated until we remain with 1 block with $|\rho_{|\rho|}|$ elements after $|\rho|-1$ block merges. It is clear that this last block gves rise to $|\rho_{|\rho|}|$ free variables. If we sum up the total number of free variables we get

$$|\rho_{|\rho|}| + \sum_{i=1}^{|\rho|-1} (|\rho_i|-1) = n - (|\rho|-1) = n+1-|\rho|.$$

All in all we see that the solution set is as described as above (i.e. each constrained variable is equal to one of the free variables), so that $N^{n+1-|\rho|}$ choices of $i_1, ..., i_n$ satisfy (36), which shows that $K_{\rho,u} = 1$ when ρ is noncrossing. It is easy to see that, when ρ has crossings, the procedure followed above will fail, so that at least one of the constrained variables is not equals to a free variable. But then $K_{\rho,u} < 1$ for such ρ, which proves the theorem.

We remark that it is the form (37) which will be used in the other appendices to compute $K_{\rho,u}$ for certain lower order ρ. From the proof, we see that when ρ is noncrossing, there exists a partition of $\{1, ..., n\}$ into $n+1-|\rho|$ blocks, where two elements are defined to be in the same block if and only if their corresponding variables are equal. It is obvious from the construction above that this partition is the Kreweras complement of ρ, denoted $K(\rho)$ [21]. This fact is used elsewhere in this paper.

APPENDIX C
THE PROOF FOR LEMMA 1

We will in the following compute the volume of the solution set of (37), as a volume in $[0,1]^{n+1-|\rho|} \subset \mathbb{R}^{n+1-|\rho|}$, as explained in the proof of theorem 2. These integrals are very tedious to compute. The formula

$$\frac{r!s!}{(r+s+1)!} = \int_0^1 x^r (1-x)^s dx$$

can be used to simplify some of the calculations for higher values of n.

A. Computation of $K_{\{\{1,3\},\{2,4\}\},u}$

This is equivalent to finding the volume of the solution set of

$$x_1 + x_3 = x_2 + x_4$$

in $\mathbb{R}^3$. Since this means that

$$x_4 = x_1 + x_3 - x_2 \text{ lies between } 0 \text{ and } 1,$$

we can set up the following integral bounds: When $x_1 + x_3 \leq 1$, we must have that $0 \leq x_2 \leq x_1 + x_3$, so that we get the contribution

$$\begin{aligned} &\int_0^1 \int_0^{1-x_1} \int_0^{x_1+x_3} dx_2 dx_3 dx_1 \\ &= \int_0^1 \left(\frac{1}{2} - \frac{1}{2}x_1^2\right) dx_1 \\ &= \left[\frac{1}{2}x_1 - \frac{1}{6}x_1^3\right]_0^1 \\ &= \frac{1}{2} - \frac{1}{6} = \frac{1}{3}. \end{aligned}$$

When $1 \leq x_1 + x_3$, we must have that $x_1 + x_3 - 1 \leq x_2 \leq 1$, so that we get the contribution

$$\begin{aligned} &\int_0^1 \int_{1-x_1}^1 \int_{x_1+x_3-1}^1 dx_2 dx_3 dx_1 \\ &= \int_0^1 \left(-\frac{1}{2}(1-x_1)^2 + \frac{1}{2}\right) dx_1 \\ &= \left[\frac{1}{6}(1-x_1)^3 + \frac{1}{2}x_1\right]_0^1 \\ &= -\frac{1}{6} + \frac{1}{2} = \frac{1}{3}. \end{aligned}$$

Adding the contributions together we get $\frac{2}{3}$, which is the stated expression for $K_{\{\{1,3\},\{2,4\}\},u}$.

Computation of certain $K_{\rho,u}$ can be simplified by the following: Let $a_l^{(m)}(x)$ be the polynomial which gives the volume in $\mathbb{R}^{m-1}$ of the solutions set to $x_1 + \cdots + x_m = x$ (constrained to $0 \leq x_i \leq 1$) for $l \leq x \leq l+1$. It is clear that these satisfy the integral equations

$$a_l^{(m+1)}(x) = \int_{x-1}^{l} a_{l-1}^{(m)}(t)dt + \int_{l}^{x} a_l^{(m)}(t)dt, \qquad (38)$$

which can be used to compute the $a_l^m(x)$ recursively. Note first that $a_0^{(1)}(x) = 1$. For $m = 2$ we have

$$\begin{aligned} a_0^{(2)}(x) &= \int_0^x a_0^{(1)}(t)dt = x \\ a_1^{(2)}(x) &= \int_{x-1}^1 a_0^{(1)}(t)dt = 2 - x. \end{aligned}$$

For $m = 3$ we have

$$\begin{aligned} a_0^{(3)}(x) &= \int_0^x a_0^{(2)}(t)dt = \frac{1}{2}x^2 \\ a_1^{(3)}(x) &= \int_{x-1}^1 a_0^{(2)}(t)dt + \int_1^x a_1^{(2)}(t)dt \\ &= 1 - \frac{1}{2}(x-1)^2 - \frac{1}{2}(2-x)^2 \\ a_2^{(3)}(x) &= \int_{x-1}^2 a_1^{(2)}(t)dt = \frac{1}{2}(3-x)^2. \end{aligned}$$

By integrating the $a_0^{(2)}(x)$, we can double-check our computation of $K_{\{\{1,3\},\{2,4\}\},u}$ above:

$$\begin{aligned} &\int_0^1 (a_0^{(2)})^2(t)dt + \int_1^2 (a_1^{(2)})^2(t)dt \\ &= \left[\frac{1}{3}t^3\right]_0^1 + \left[-\frac{1}{3}(2-t)^3\right]_1^2 \\ &= \frac{2}{3}. \end{aligned}$$

B. Computation of $K_{\{\{1,3,5\},\{2,4,6\}\},u}$

For $m = 3$, integration gives

$$\begin{aligned} &\int_0^1 (a_0^{(3)})^2(t)dt + \int_1^2 (a_1^{(3)})^2(t)dt + \int_2^3 (a_2^{(3)})^2(t)dt \\ &= \left[\frac{1}{20}t^5\right]_0^1 + \\ &\quad [t + \frac{1}{20}(t-1)^5 - \frac{1}{20}(2-t)^5 - \frac{1}{3}(t-1)^3 + \\ &\quad \frac{1}{3}(2-t)^3 + \frac{1}{60}(t-1)^5]_1^2 + \\ &\quad \left[-\frac{1}{20}(3-t)^5\right]_2^3 \\ &= \frac{1}{20} + 1 + \frac{1}{20} + \frac{1}{20} - \frac{1}{3} - \frac{1}{3} + \frac{1}{60} + \frac{1}{20} \\ &= \frac{11}{20}, \end{aligned}$$

which is the stated expression for $K_{\{\{1,3,5\},\{2,4,6\}\},u}$.

C. Computation of $K_{\{\{1,4\},\{2,5\},\{3,6\}\},u}$

This is equivalent to finding the volume of the solution set of

$$x_1 + x_4 = x_2 + x_5 = x_3 + x_6$$

in $\mathbb{R}^4$, which is computed as

$$\begin{aligned} &\int_0^1 (a_0^{(2)})^3(t)dt + \int_1^2 (a_1^{(2)})^3(t)dt \\ &= \left[\frac{1}{4}t^4\right]_0^1 + \left[-\frac{1}{4}(2-t)^4\right]_1^2 \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \end{aligned}$$

which is the stated expression for $K_{\{\{1,4\},\{2,5\},\{3,6\}\},u}$.

D. Computation of $K_{\{\{1,4\},\{2,6\},\{3,5\}\},u}$

This is equivalent to finding the volume of the solution set of

$$\begin{aligned} x_1 + x_4 &= x_2 + x_5 \\ x_2 + x_6 &= x_3 + x_1 \end{aligned}$$

in $\mathbb{R}^4$. Since this means that

$$\begin{aligned} x_5 &= x_1 - x_2 + x_4 \text{ lies between } 0 \text{ and } 1, \\ x_6 &= x_1 - x_2 + x_3 \text{ lies between } 0 \text{ and } 1, \end{aligned}$$

we can set up the following integral bounds:

For $x_2 \geq x_1$ we must have $x_2 - x_1 \leq x_3, x_4 \leq 1$, so that we get the contribution

$$\begin{aligned} &\int_0^1 \int_{x_1}^1 \int_{x_2-x_1}^1 \int_{x_2-x_1}^1 dx_4 dx_3 dx_2 dx_1 \\ &= \int_0^1 \int_{x_1}^1 (1 - x_2 + x_1)^2 dx_2 dx_1 \\ &= \int_0^1 (-\frac{1}{3}x_1^3 + \frac{1}{3})dx_1 \\ &= \left[-\frac{1}{12}x_1^4 + \frac{1}{3}x_1\right]_0^1 \\ &= \frac{1}{3} - \frac{1}{12} = \frac{1}{4}. \end{aligned}$$

It is clear that for $x_1 \geq x_2$ we get the same result by symmetry, so that the total contribution is $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$, which proves the claim.

E. Computation of $K_{\{\{1,5\},\{3,7\},\{2,4,6\}\},u}$

This is equivalent to finding the volume of the solution set of

$$\begin{aligned} x_1 + x_5 &= x_2 + x_6 \\ x_3 + x_7 &= x_4 + x_1 \end{aligned}$$

in $\mathbb{R}^5$, or

$$\begin{aligned} x_6 &= x_5 + x_1 - x_2 \text{ lies between } 0 \text{ and } 1, \\ x_7 &= x_4 + x_1 - x_3 \text{ lies between } 0 \text{ and } 1\ . \end{aligned} \qquad (39)$$

Assume first that $x_1 \leq x_2 \leq x_3$. Then $x_2 - x_1 \leq x_5 \leq 1$ and $x_3 - x_1 \leq x_4 \leq 1$, so that we get the contribution

$$\begin{aligned}
&\int_0^1 \int_{x_1}^1 \int_{x_2}^1 \int_{x_2-x_1}^1 \int_{x_3-x_1}^1 dx_4 dx_5 dx_3 dx_2 dx_1 \\
&= \int_0^1 \int_{x_1}^1 \int_{x_2}^1 (1 - x_2 + x_1)(1 - x_3 + x_1) dx_3 dx_2 dx_1 \\
&= \int_0^1 \int_{x_1}^1 (\frac{1}{2}(1 - x_2 + x_1)^3 \\
&\quad -\frac{1}{2}x_1^2(1 - x_2 + x_1)) dx_2 dx_1 \\
&= \int_0^1 \left[-\frac{1}{8}(1 - x_2 + x_1)^4 + \frac{1}{4}x_1^2(1 - x_2 + x_1)^2 \right]_{x_1}^1 dx_1 \\
&= \int_0^1 (\frac{1}{8} - \frac{1}{8}x_1^4 + \frac{1}{4}x_1^4 - \frac{1}{4}x_1^2) dx_1 \\
&= \frac{1}{8} - \frac{1}{40} + \frac{1}{20} - \frac{1}{12} \\
&= \frac{15 - 3 + 6 - 10}{120} = \frac{8}{120} = \frac{1}{15}.
\end{aligned}$$

We get the same contribution for $x_1 \leq x_3 \leq x_2$ by symmetry.

Assume that $x_3 \leq x_2 \leq x_1$. Then $0 \leq x_5 \leq 1 + x_2 - x_1$ and $0 \leq x_4 \leq 1 + x_3 - x_1$, so that we get the contribution

$$\begin{aligned}
&\int_0^1 \int_0^{x_1} \int_0^{x_2} \int_0^{1+x_2-x_1} \int_0^{1+x_3-x_1} dx_4 dx_5 dx_3 dx_2 dx_1 \\
&= \int_0^1 \int_0^{x_1} \int_0^{x_2} (1 + x_2 - x_1)(1 + x_3 - x_1) dx_3 dx_2 dx_1 \\
&= \int_0^1 \int_0^{x_1} (\frac{1}{2}(1 + x_2 - x_1)^3 \\
&\quad -\frac{1}{2}(1 + x_2 - x_1)(1 - x_1)^2) dx_2 dx_1 \\
&= \int_0^1 (\frac{1}{8} - \frac{1}{8}(1 - x_1)^4 \\
&\quad -\frac{1}{4}(1 - x_1)^2 + \frac{1}{4}(1 - x_1)^4) dx_1 \\
&= \frac{1}{8} - \frac{1}{40} - \frac{1}{12} + \frac{1}{20} \\
&= \frac{15 - 3 - 10 + 6}{120} = \frac{8}{120} = \frac{1}{15}.
\end{aligned}$$

We get the same contribution for $x_2 \leq x_3 \leq x_1$ by symmetry.

Assume that $x_2 \leq x_1 \leq x_3$. Then $0 \leq x_5 \leq x_2 - x_1 + 1$ and $x_3 - x_1 \leq x_4 \leq 1$, so that we get the contribution

$$\begin{aligned}
&\int_0^1 \int_0^{x_1} \int_{x_1}^1 \int_0^{x_2-x_1+1} \int_{x_3-x_1}^1 dx_4 dx_5 dx_3 dx_2 dx_1 \\
&= \int_0^1 \int_0^{x_1} \int_{x_1}^1 (1 + x_2 - x_1)(1 - x_3 + x_1) dx_3 dx_2 dx_1 \\
&= \int_0^1 \int_0^{x_1} (-\frac{1}{2}x_1^2(1 + x_2 - x_1) \\
&\quad +\frac{1}{2}(1 + x_2 - x_1)) dx_2 dx_1 \\
&= \int_0^1 \left(-\frac{1}{4}x_1^2 + \frac{1}{4}x_1^2(1 - x_1)^2 + \frac{1}{4} - \frac{1}{4}(1 - x_1)^2 \right) dx_1 \\
&= -\frac{1}{12} + \frac{1}{120} + \frac{1}{4} - \frac{1}{12} \\
&= \frac{-10 + 1 + 30 - 10}{120} \\
&= \frac{11}{120}.
\end{aligned}$$

We get the same contribution for $x_3 \leq x_1 \leq x_2$ by symmetry.

Adding the six contributions together, we get

$$\frac{4}{15} + \frac{11}{60} = \frac{27}{60} = \frac{9}{20},$$

which proves the claim.

F. *The computation of* $K_{\{\{1,6\},\{2,4\},\{3,5,7\}\},u}$

This is equivalent to finding the volume of the solution set of

$$\begin{aligned}
x_1 + x_6 &= x_2 + x_7 \\
x_2 + x_4 &= x_3 + x_5
\end{aligned}$$

in $\mathbb{R}^5$, or

$$\begin{aligned}
x_6 &= x_7 + x_2 - x_1 \text{ lies between } 0 \text{ and } 1, \\
x_5 &= x_4 + x_2 - x_3 \text{ lies between } 0 \text{ and } 1, .
\end{aligned}$$

This can be obtained from (39) by a permutation of the variables, so the contribution from $K_{\{\{1,6\},\{2,4\},\{3,5,7\}\},u}$ must also be $\frac{9}{20}$, which proves the claim.

Appendix D
The proof for theorem 3

We will have use for the following result, taken from [21]:

Lemma 2: The number of noncrossing partitions in $NC(n)$ with r_1 blocks of length 1, r_2 blocks of length 2 and so on (so that $r_1 + 2r_2 + 3r_3 + \cdots nr_n = n$) is

$$\frac{n!}{r_1!r_2!\cdots r_n!(n + 1 - r_1 - r_2 \cdots r_n)!}.$$

Using this and a similar formula for the number of partitions with prescribed block sizes, we obtain the following list of cardinalities for noncrossing partitions in $NC(7)$ with prescribed block sizes. The cardinalities of all partitions in $\mathcal{P}(7)$ with these prescribed block sizes is also shown in parenthesis:

- (7): 1 (of 1)
- $(6, 1)$: 7 (of 7)
- $(5, 2)$: 7 (of 21)

- $(5,1,1)$: 21 (of 21)
- $(4,3)$: 7 (of 35)
- $(4,2,1)$: 42 (of 105)
- $(4,1,1,1)$: 35 (of 35)
- $(3,3,1)$: 21 (of 70)
- $(3,2,2)$: 21 (of 105)
- $(3,2,1,1)$: 105 (of 210)
- $(3,1,1,1,1)$: 35 (of 35)
- $(2,2,2,1)$: 35 of (of 105)
- $(2,2,1,1,1)$: 70 (of 105)
- $(2,1,1,1,1,1)$: 21 (of 21)
- $(1,1,1,1,1,1,1)$: 1 (of 1)

This totals 429 noncrossing partitions, and 877 partitions. A similar listing can be written down for partitions of order 4, 5, and 6 also.

For the proof, we need to compute (8) for all possible block cardinalities $(r_1,...,r_k)$, and insert these in (11). The formulas for the three first moments are obvious, since all partitions of length ≤ 3 are noncrossing. For the remaining computations, the following two observations save a lot of work:

- If $\rho_1 \in \mathcal{P}(n_1)$, $\rho_2 \in \mathcal{P}(n_2)$ with $n_1 < n_2$, and ρ_1 can be otained from ρ_2 by omitting elements k in $\{1,...,n_2\}$ such that k and $k+1$ are in the same block, then we must have that $K_{\rho_1,u} = K_{\rho_2,u}$. This is straightforward to prove since it folows from the proof of theorem 2 that i_{k+1} can be chosen arbitrarily between 1 and N in such a case.
- $K_{\rho_1,u} = K_{\rho_2,u}$ if the set of equations (37) for ρ_1 can be obtained by a permutation of the variables in the set of equations for ρ_2. Since the rank of the matrix for (37) equals the number of equations -1, we actually need only have that $|\rho_1|-1$ of the $|\rho_1|$ equations can be obtained from permutation of $|\rho_2|-|$ equations of the $|\rho_2|$ equations in the equation system for ρ_2

A. The moment of fourth order

The result is here obvious except for the case for the three partitions with block cardinalities $(2,2)$ (for all other block cardinalities, all partitions are noncrossing, so that $K_{r_1,r_2,...,r_k}$ is simply the number of noncrossing partitions with block cardinalities $(r_1,...,r_k)$. this number can be computed from lemma 2). Two of the partitions with blocks of cardinality $(2,2)$ are noncrossing, the third one is not. We see from lemma 1 that the total contribution is

$$\begin{aligned} K_{2,2} &= 2 + K_{\{\{1,3\},\{2,4\}\},u} \\ &= 2 + \tfrac{2}{3} = \tfrac{8}{3}. \end{aligned}$$

The formula for the fourth moment follows.

B. The moment of fifth order

Here two cases require extra attention:

1) $\rho = \{\rho_1,\rho_2\}$ with $|\rho_1| = 3$, $|\rho_2| = 2$: There are 10 such partitions, and 5 of them have crossings and constribute with $K_{\{\{1,3\},\{2,4\}\},u}$. The total contribution is therefore

$$\begin{aligned} & 5 + 5 \times K_{\{\{1,3\},\{2,4\}\},u} \\ = \; & 5 + 5 \times \tfrac{2}{3} = \tfrac{25}{3}. \end{aligned}$$

2) $\rho = \{\rho_1,\rho_2,\rho_3\}$ with $|\rho_1| = |\rho_2| = 2$, $|\rho_3| = 1$: There are 15 such partitions, of which 5 have crossings. The total contribution is therefore

$$\begin{aligned} & 10 + 5 \times K_{\{\{1,3\},\{2,4\}\},u} \\ = \; & 10 + 5 \times \tfrac{2}{3} = \tfrac{40}{3}. \end{aligned}$$

C. The moment of sixth order

Five cases require extra attention:

1) $\rho = \{\rho_1,\rho_2\}$ with $|\rho_1| = 4$, $|\rho_2| = 2$: There are 15 such partitions, and 6 of them are noncrossing. The crossing ones contribute with $K_{\{\{1,3\},\{2,4\}\},u}$, so the total contribution is

$$\begin{aligned} & 6 + 9K_{\{\{1,3\},\{2,4\}\},u} \\ = \; & 6 + 9 \times \tfrac{2}{3} = 12. \end{aligned}$$

2) $\rho = \{\rho_1,\rho_2\}$ with $|\rho_1| = |\rho_2| = 3$: There are 10 such partitions. 3 of these are noncrossing. One of the crossing partitions contribute with $K_{\{\{1,3,5\},\{2,4,6\}\},u}$, the others contribute with $K_{\{\{1,3\},\{2,4\}\},u}$. The total contribution is therefore

$$\begin{aligned} & 3 + 6 \times K_{\{\{1,3\},\{2,4\}\},u} + K_{\{\{1,3,5\},\{2,4,6\}\},u} \\ = \; & 3 + 6 \times \tfrac{2}{3} + \tfrac{11}{20} = \tfrac{151}{20}. \end{aligned}$$

3) $\rho = \{\rho_1,\rho_2,\rho_3\}$ with $|\rho_1| = 3, |\rho_2| = 2, |\rho_3| = 1$: There are 60 such partitions, of which 30 are noncrossing. The total contribution is

$$\begin{aligned} & 30 + 30 \times K_{\{\{1,3\},\{2,4\}\},u} \\ = \; & 30 + 30 \times \tfrac{2}{3} = 50. \end{aligned}$$

4) $\rho = \{\rho_1,\rho_2,\rho_3\}$ with $|\rho_1| = |\rho_2| = |\rho_3| = 2$: There a 15 such partitions. 5 of them are noncrossing. 4 of the partitions with crossings have no inner block, and each of these contributes with $K_{\{\{1,4\},\{2,5\},\{3,6\}\},u}$. The remaining 6 partitions with crossings have an inner block, and each contributes with $K_{\{\{1,3\},\{2,4\}\},u}$. The total contribution is therefore

$$\begin{aligned} & 5 + 4K_{\{\{1,4\},\{2,5\},\{3,6\}\},u} + 6K_{\{\{1,3\},\{2,4\}\},u} \\ = \; & 5 + 4 \times \tfrac{1}{2} + 6 \times \tfrac{2}{3} = 11. \end{aligned}$$

5) $\rho = \{\rho_1,\rho_2,\rho_3,\rho_4\}$ with $|\rho_1| = |\rho_2| = 2, |\rho_3| = |\rho_4| = 1$: There are 45 such partitions, of which 15 has crossings. The total contribution is

$$\begin{aligned} & 30 + 15K_{\{\{1,3\},\{2,4\}\},u} \\ = \; & 30 + 15 \times \tfrac{2}{3} = 40. \end{aligned}$$

D. The moment of seventh order

8 cases require extra attention:

1) $\rho = \{\rho_1,\rho_2\}$ with $|\rho_1| = 5$, $|\rho_2| = 2$: There are 21 such partitions, and 7 of them are noncrossing. The total contribution is

$$\begin{aligned} & 7 + 14 \times K_{\{\{1,3\},\{2,4\}\},u} \\ = \; & 7 + 14 \times \tfrac{2}{3} = \tfrac{49}{3}. \end{aligned}$$

2) $\rho = \{\rho_1,\rho_2\}$ with $|\rho_1| = 4$, $|\rho_2| = 3$: There are 35 such partitions, of which 7 are noncrossing. 7 of the partitions with crossings contribute with $K_{\{\{1,3,5\},\{2,4,6\}\},u}$, the rest contribute with $K_{\{\{1,3\},\{2,4\}\},u}$. The total contribution is

$$\begin{aligned} & 7 + 7 \times K_{\{\{1,3,5\},\{2,4,6\}\},u} + 21 \times K_{\{\{1,3\},\{2,4\}\},u} \\ = \; & 7 + 7 \times \tfrac{11}{20} + 21 \times \tfrac{2}{3} = \tfrac{497}{20}. \end{aligned}$$

3) $\rho = \{\rho_1, \rho_2, \rho_3\}$ with $|\rho_1| = 4$, $|\rho_2| = 2$, $|\rho_3| = 1$: The total contribution is

$$7 \times 12 = 84.$$

4) $\rho = \{\rho_1, \rho_2, \rho_3\}$ with $|\rho_1| = 3$, $|\rho_2| = 3$, $|\rho_3| = 1$: The total contribution is

$$7 \times \tfrac{151}{20} = \tfrac{1057}{20}.$$

5) $\rho = \{\rho_1, \rho_2, \rho_3\}$ with $|\rho_|| = 3$, $|\rho_2| = |\rho_3| = 2$: This is the hardest one to compute. A close inspection of all 105 such partitions in light of lemma 1 gives that 21 of them contribute with 1 (the noncrossing ones), 14 of them contribute with $\frac{9}{20}$, 42 of them contribute with $\frac{2}{3}$, and 28 of them contribute with $\frac{1}{2}$. The total contribution is therefore

$$21 + 14 \times \tfrac{9}{20} + 42 \times \tfrac{2}{3} + 28 \times \tfrac{1}{2} = 63 + \tfrac{63}{10} = \tfrac{693}{10}.$$

6) $\rho = \{\rho_1, \rho_2, \rho_3, \rho_4\}$ with $|\rho_1| = 3$, $|\rho_2| = 2$, $|\rho_3| = |\rho_4| = 1$: The total contribution is

$$21 \times \tfrac{25}{3} = 175.$$

7) $\rho = \{\rho_1, \rho_2, \rho_3, \rho_4\}$ with $|\rho_1| = |\rho_2| = |\rho_3| = 2$, $|\rho_4| = 1$: The total contribution is

$$7 \times 11 = 77.$$

8) $\rho = \{\rho_1, \rho_2, \rho_3, \rho_4, \rho_5\}$ with $|\rho_1| = |\rho_2| = 2$, $|\rho_3| = |\rho_4| = |\rho_5| = 1$: The total contribution is

$$35 \times \tfrac{8}{3} = \tfrac{280}{3}.$$

Appendix E
The proof of theorem 4

In order to get the exact expressions in theorem 4, we now need to keep track of the $K_{\rho,u,N}$ defined by (4), not only the limits $K_{\rho,u}$ (if we had not assumed $\omega = u$, the calculations for $K_{\rho,\omega,N}$ would be much more cumbersome). When ρ is a partition of $\{1, ..., n\}$ and $n \leq 4$, we have that $K_{\rho,u,N} = K_{\rho,u} = 1$ when $\rho \neq \{\{1,3\},\{2,4\}\}$. We also have that

$$K_{\{\{1,3\},\{2,4\}\},u,N} = \frac{2}{3} + \frac{1}{N} + \frac{1}{6N^2}, \tag{40}$$

where we have used that $\sum_{i=1}^N i^2 = \frac{N}{3}(N+1)(N+\frac{1}{2})$ [26]. We also need the exact expression for the quantity

$$\begin{aligned} T_\rho \;=\; & \sum_{\substack{(j_1,\ldots,j_n)\\ \text{giving rise to } \rho}} L^{-|\rho|} \\ & \mathbf{D}_1(N)(j_1,j_1)\mathbf{D}_2(N)(j_2,j_2) \\ & \cdots \times \mathbf{D}_n(N)(j_n,j_n) \end{aligned}$$

from (33) (i.e. we can not add (34) to obtain the approximation (35) here). We see that

$$T_\rho = D_\rho - \sum_{\rho' > \rho} L^{|\rho'|-|\rho|} T_{\rho'}, \tag{41}$$

(where D_ρ and D_n are defined as in section II, but without taking the limit) which can be used recursively to express the T_ρ in terms of the D_ρ. We obtain the following formulas for $n = 4$:

$$\begin{aligned} & T_{\{\{1,2,3,4\}\}} = D_4 \\ & T_{\{\{1,2,3\},\{4\}\}} = D_3 D_1 - L^{-1} D_4 \\ & T_{\{\{1,2\},\{3,4\}\}} = D_2^2 - L^{-1} D_4 \\ & T_{\{\{1,2\},\{3\},\{4\}\}} \\ & = D_2 D_1^2 - 2L^{-1}\left(D_3 D_1 - L^{-1} D_4\right) \\ & -L^{-1}\left(D_2^2 - L^{-1} D_4\right) - L^{-2} D_4 \\ & = D_2 D_1^2 - L^{-1}\left(D_2^2 + 2D_3 D_1\right) + 2L^{-2} D_4 \\ & T_{\{\{1\},\{2\},\{3\},\{4\}\}} \\ & = D_1^4 - 6L^{-1}\left(D_2 D_1^2 - L^{-1}\left(D_2^2 + 2D_3 D_1\right) + 2L^{-2} D_4\right) \\ & -3L^{-2}\left(D_2^2 - L^{-1} D_4\right) \\ & -4L^{-2}\left(D_3 D_1 - L^{-1} D_4\right) - L^{-3} D_4 \\ & = -6L^{-3} D_4 + L^{-2}\left(8 D_3 D_1 + 3 D_2^2\right) - 6L^{-1} D_2 D_1^2 + D_1^4. \end{aligned} \tag{42}$$

For $n = 3$ and $n = 2$ the formulas are

$$\begin{array}{lcl} T_{\{\{1,2,3\}\}} & = & D_3 \\ T_{\{\{1,2\},\{3\}\}} & = & D_1 D_2 - L^{-1} D_3 \\ T_{\{\{1\},\{2\},\{3\}\}} & = & D_1^3 - 3L^{-1} D_1 D_2 + 2L^{-2} D_3 \\ T_{\{\{1,2\}\}} & = & D_2 \\ T_{\{\{1\},\{2\}\}} & = & D_1^2 - L^{-1} D_2. \end{array} \tag{43}$$

It is clear that (42) and (43) cover all possibilities when it comes to partition block sizes. Using (9), and putting (40), (42), and (43) into (33) we get the expressions in theorem 4 after some calculations.

A. First order approximations to theorem 4

If we are only interested in first order approximations rather than exact expressions, (41) gives us

$$T_\rho \approx D_\rho - \sum_{\substack{\rho' > \rho \\ |\rho| - |\rho'| = 1}} L^{-1} D_{\rho'},$$

which is easier to compute. Also, we need only first order approximations to $K_{\rho,u,N}$, which is much easier to compute than the exact expression. For (40), this is

$$K_{\{\{1,3\},\{2,4\}\},u,N} \approx \frac{2}{3} + \frac{1}{N},$$

Inserting these two approximations in (33) gives a first order approximation of the moments.

Appendix F
The proof of theorem 5

For $\rho = 1_n$ theorem 5 is trivial. We will thus assume that $\rho \neq 1_n$ in the following. We first prove that $\lim_{N\to\infty} K_{\rho,\omega,N}$ exists whenever p_ω is continous. To simplify notation, define

$$\begin{aligned} F(\omega) \;&=\; \prod_{k=1}^n \frac{1 - e^{jN(\omega_{b(k-1)} - \omega_{b(k)})}}{1 - e^{j(\omega_{b(k-1)} - \omega_{b(k)})}} \\ &=\; \prod_{k=1}^n \frac{\sin\left(N(\omega_{b(k-1)} - \omega_{b(k)})/2\right)}{\sin\left((\omega_{b(k-1)} - \omega_{b(k)})/2\right)}, \end{aligned}$$

and set $\omega = (\omega_1, ..., \omega_{|\rho|})$ and $d\omega = d\omega_1 \cdots d\omega_{|\rho|}$. Since ω is continous, there exists a p_{max} such that $p_\omega(\omega_i) \leq p_{max}$ for all ω_i. Then we have that

$$\begin{aligned} |K_{\rho,\omega,N}| &\leq \frac{p_{max}^{|\rho|}}{N^{m+1-|\rho|}} \\ &\times \int_{[0,2\pi)^{|\rho|}} \prod_{k=1}^{n} \left| \frac{\sin\left(N(x_{b(k-1)}-x_{b(k)})/2\right)}{\sin\left((x_{b(k)}-x_{b(k+1)})/2\right)} \right| dx, \end{aligned}$$

where we have converted to Lebesgue measure. Consider first the set

$$U = \{\omega | |x_{b(k-1)} - x_{b(k)}| \leq \pi \forall k\}.$$

When $\frac{2\pi}{N} \leq |\omega_{b(k-1)} - \omega_{b(k)}| \leq \pi$, it is clear that

$$\left| \frac{\sin\left(N(x_{b(k-1)} - x_{b(k)})/2\right)}{\sin\left((x_{b(k-1)} - x_{b(k)})/2\right)} \right| \leq \left| \frac{4}{x_{b(k-1)} - x_{b(k)}} \right|, \quad (44)$$

since $\left|\sin\left(N(x_{b(k-1)} - x_{b(k)})/2\right)\right| \leq 1$, and since $|\sin(x)| \geq |\frac{x}{2}|$ when $|x| \leq \frac{\pi}{2}$. When $|x_{b(k-1)} - x_{b(k)}| \leq \frac{2\pi}{N}$ we have that

$$\left| \frac{\sin\left(N(x_{b(k-1)} - x_{b(k)})/2\right)}{\sin\left((x_{b(k-1)} - x_{b(k)})/2\right)} \right| \leq N. \quad (45)$$

Let $k_1, ..., k_{|\rho|} \in \mathbb{Z}$, and assume that $k_{|\rho|} = 0$. By using the triangle inequality, it is clear that on the set

$$D_{k_1,...,k_{|\rho|-1}} = \{\omega | \left| x_i - \frac{2k_i\pi}{N} \right| \leq \frac{\pi}{N} \forall 1 \leq i \leq |\rho|\},$$

when $|k_r - k_s| \geq 2$ for all r, s, the i'th factor in $F(x)$ is bounded by $\frac{4N}{\left(|k_{b(r-1)} - k_{b(r)}|-1\right)\pi}$ due to (44). Also, when $|k_r - k_s| < 2$ for some r, s, the corresponding factors in $F(x)$ are bounded by N on $D_{k_1,...,k_{|\rho|}}$ due to (45). Note also that the volume of $D_{k_1,...,k_{|\rho|-1}}$ is $(2\pi)^{|\rho|-1} N^{1-|\rho|}$. By adding some more terms (to compensate for the different behaviour for $|k_r - k_s| \geq 2$ and $|k_r - k_s| < 2$), we have that we can find a constant D that

$$\begin{aligned} &\frac{1}{N^{n+1-|\rho|}} \int_U |F(x)| dx \\ &\leq \frac{1}{N^{n+1-|\rho|}} N^n \\ &\times \sum_{\substack{0 \leq k_1,...,k_{|\rho|-1} < N \\ \text{all } k_i \text{ different}}} \left(\prod_{r=1}^{n} \frac{D}{|k_{b(r-1)} - k_{b(r)}|} \right) 2\pi (2\pi)^{|\rho|-1} N^{1-|\rho|} \\ &= (2\pi)^{|\rho|} D^n \sum_{\substack{0 \leq k_1,...,k_{|\rho|-1} < N \\ \text{all } k_i \text{ different}}} \prod_{r=1}^{n} \frac{1}{|k_{b(r-1)} - k_{b(r)}|}, \end{aligned} \quad (46)$$

where we have integrated w.r.t. $x_{|\rho|}$ also (i.e. $k_{|\rho|}$ is kept constant in (46)). A similar analysis as for U applies for the complement set

$$V = \{\omega | \pi \leq |x_{b(k-1)} - x_{b(k)}| \leq 2\pi \text{ for some } k\},$$

so that we can find a constant C such that

$$\begin{aligned} &\frac{1}{N^{n+1-|\rho|}} \int_{[0,2\pi)^{|\rho|}} |F(x)| dx \\ &\leq C \sum_{\substack{0 \leq k_1,...,k_{|\rho|-1} < N \\ \text{all } k_i \text{ different}}} \prod_{r=1}^{n} \frac{1}{|k_{b(r-1)} - k_{b(r)}|}, \end{aligned} \quad (47)$$

It is clear this sum converges: First of all, this is only needed to prove for $\rho = 0_n$, since the summands for $\rho \neq 0_n$ is only a subset of the summands for $\rho = 0_n$.

Secondly, for $\rho = 0_n$, (47) can be bounded by considering convolutions of the following function with itself:

$$f(x) = \begin{cases} \frac{1}{|x|} & \text{for } |x| > 1 \\ 0 & \text{for } |x| \leq 1 \end{cases} \quad (48)$$

The assumption that $f(x) = 0$ in a neighbourhood of zero is due to the fact that the k_i are all different. Note that $|f(x)| \leq \frac{1}{|x|^{1-\epsilon}}$ for any $0 < \epsilon < 1$. Also, the $n-2$-fold convolution (we wait with the $n-1$'th convolution till the end) of $\frac{1}{|x|^{1-\epsilon}}$ with itself exist outside 0 whenever $0 < (n-2)\epsilon < 1$, and is on the form $r\frac{1}{|x|^{1-(n-2)\epsilon}}$ for some constant r [26]. Therefore, (47) is bounded by

$$\begin{aligned} \int_{|x|>1} r \frac{1}{|x|^{1-(n-2)\epsilon}} \frac{1}{|x|} dx &= \int_{|x|>1} r \frac{1}{|x|^{2-(n-2)\epsilon}} dx \\ &= \frac{2r}{(n-2)\epsilon - 1}. \end{aligned}$$

This proves that the entire sum (47) is bounded, and thus also the statement on the existence of the limit $K(\rho, \omega)$ in theorem 5 when the density is continous.

For the rest of the proof of theorem 5 , we first record the following result:

Lemma 3: For any $\epsilon > 0$,

$$\lim_{N \to \infty} \frac{1}{N^{n+1-|\rho|}} \int_{B_{\epsilon,r}} F(\omega) d\omega = 0, \quad (49)$$

where

$$B_{\epsilon,r} = \{(\omega_1, ..., \omega_{|\rho|}) | |\omega_{b(r-1)} - \omega_{b(r)}| > \epsilon\}.$$

Proof: The set $B_{\epsilon,r}$ corresponds to those $k_1, ..., k_{|\rho|}$ in (47) for which $|k_{b(r-1)} - k_{b(r)}| > \frac{N}{2\pi}\epsilon$. Thus, for large N, we sum over $k_1, ..., k_{|\rho|}$ in (47) for which $|k_{b(r-1)} - k_{b(r)}|$ is arbitrarily large. By the convergence of the Fourier integral of $\frac{1}{|x|}$, it is clear that this converges to zero. ■

Define

$$B_\epsilon = \{(\omega_1, ..., \omega_{|\rho|}) | |\omega_i - \omega_j| > \epsilon \text{ for some } i, j\}.$$

If $\omega \in B_\epsilon$, there must exist an r so that $|\omega_{b(r-1)} - \omega_{b(r)}| > \frac{2\epsilon}{n}$, so that $\omega \in B_{r,2\epsilon/n}$. This means that

$$B_\epsilon \subset \cup_r B_{r,2\epsilon/n},$$

so that by lemma 3 also

$$\lim_{N \to \infty} \frac{1}{N^{n+1-|\rho|}} \int_{B_\epsilon} F(\omega) d\omega = 0.$$

This means that in the integral for $K_{\rho,\omega,N}$, we need only integrate over the ω which are arbitrarily close to the diagonal, (where $\omega_1 = \cdots = \omega_{|\rho|}$). We thus have

$$\begin{aligned} K_{\rho,\omega} &= \lim_{N\to\infty} \frac{1}{N^{n+1-|\rho|}} \int_{[0,2\pi)^{|\rho|}} F(x) \prod_{r=1}^{|\rho|} p_\omega(x_r) dx \\ &= \lim_{N\to\infty} \frac{1}{N^{n+1-|\rho|}} \int_{[0,2\pi)^{|\rho|}} F(x) p_\omega(x_{|\rho|})^{|\rho|} dx \\ &= \lim_{N\to\infty} \frac{1}{N^{n+1-|\rho|}} \int_0^{2\pi} p_\omega(x_{|\rho|})^{|\rho|} \\ &\quad \left(\int_{[0,2\pi)^{|\rho|-1}} F(x) dx_1 \cdots dx_{|\rho|-1} \right) \\ &\quad dx_{|\rho|}. \end{aligned}$$

We used here the fact that the density is continous. Using that

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N^{n+1-|\rho|}} \int_{[0,2\pi)^{|\rho|-1}} F(x) dx_1 \cdots dx_{|\rho|-1} \\ &= (2\pi)^{|\rho|-1} K_{\rho,u} \end{aligned} \quad (50)$$

when $x_{|\rho|}$ is kept fixed at an arbitrary value (this is straightforward by using the methods from the proof of theorem 2 and (12)), and again using the fact that the density is continous, we get that the above equals

$$K_{\rho,u} (2\pi)^{|\rho|-1} \int_0^{2\pi} p_\omega(x_{|\rho|})^{|\rho|} dx_{|\rho|},$$

which is what we had to show.

APPENDIX G
THE PROOF OF THEOREM 6

The contribution in the integral $K_{\rho,\omega,N}$ comes only from when the ω_i coincide with the atoms of p. Actually, we evaluate $\frac{1-e^{jN\omega}}{1-e^{j\omega}}$ in points on the form $\omega = \alpha_i - \alpha_j$. This evaluates to $N^n p_i^n$ when all ω_i are chosen equal to the same atom α_j. Since $\lim_{N\to\infty} \frac{1-e^{jN\omega}}{N(1-e^{j\omega})} = 0$ for any fixed $\omega \neq 0$, $\lim_{N\to\infty} K_{\rho,\omega,N} N^{-n} = 0$ when ω is chosen from nonequal atoms. (32) (with additional $1/N$-factors) thus becomes

$$\begin{array}{l} \sum_{\rho\in\mathcal{P}(n)} \\ \sum_{\substack{(j_1,\ldots,j_n) \\ \text{giving rise to } \rho}} \\ \sum_{(i_1,\ldots,i_n)} \\ \quad N^{|\rho|-2n-1} c^{|\rho|-1} L^{-|\rho|} \\ \quad \left(\sum_i N^n p_i^n + a_{\rho,N} N^n\right)) \\ \quad \mathbf{D}_1(N)(j_1,j_1)\mathbf{D}_2(N)(j_2,j_2) \\ \quad \cdots \times \mathbf{D}_n(N)(j_n,j_n), \end{array} \tag{51}$$

where $\lim_{N\to\infty} a_{\rho,N} = 0$. Multiplying both sides with N and letting N go to infinity gives

$$\lim_{N\to\infty} \sum_{\rho\in\mathcal{P}(n)} N^{|\rho|-n} c^{|\rho|-1} \left(\sum_i p_i^n + a_{\rho,N}\right) D_\rho.$$

It is clear that this converges to 0 when $\rho \neq 0_n$ (since $|\rho| < n$ in this case), so that the limit is

$$c^{n-1}\left(\sum_i p_i^n\right)\alpha_{0_n} = c^{n-1} p^{(n)} \lim_{N\to\infty} \prod_{i=1}^n tr_L\left(\mathbf{D}_i(N)\right),$$

which proves the claim

APPENDIX H
THE PROOF OF THEOREM 7

We need the following identity [26]:

$$\int_0^\infty x^{-s} e^{jnx} dx = \frac{\Gamma(1-s)}{|n|^{1-s}} e^{\frac{j sgn(n)(1-s)\pi}{2}},$$

where $sgn(x) = 1$ if $x > 0$, $sgn(x) = -1$ if $x < 0$, and 0 otherwise. From this it follows that

$$\begin{array}{l} \int_{-\infty}^\infty p_i |x-\alpha_i|^{-s} e^{jnx} dx = \\ 2p_i e^{jn\alpha_i} \frac{\Gamma(1-s)}{|n|^{1-s}} \cos\left(\frac{(1-s)\pi}{2}\right). \end{array} \tag{52}$$

Note that the measure with density p, has the same asymptotics near α_i as the measure with density $p_i|x-\alpha_i|^{-s}$ on

$$\left(-\left(\frac{1-s}{2p_i}\right)^{\frac{1}{1-s}}, \left(\frac{1-s}{2p_i}\right)^{\frac{1}{1-s}}\right).$$

As in the proof in appendix G, the integral for the expansion coefficients is dominated by the behaviour near the points $(\alpha_i, ..., \alpha_i)$. To see this, note that the behaviour near the singular points on the diagonal is $O\left(s(|\rho|-n)-1\right)$ when polynomic growth of order s of the density near the singular points is assumed. This is very much related to (47) in appendix F, since $K_{\rho,\omega}$ here in a similar way can be bounded by (taking into account new powers of N)

$$\begin{array}{l} C\frac{1}{N^{n+ns+1-|\rho|}} N^n N^{-|\rho|} N^{|\rho|s} \\ \times \sum_{\substack{0\le k_1,\ldots,k_{|\rho|}<N \\ \text{all } k_i \text{ different}}} \prod_{r=1}^n \frac{1}{|k_{b(r-1)}-k_{b(r)}|} \prod_{t=1}^{|\rho|} k_t^{-s}. \end{array} \tag{53}$$

In (53), the N^n-factor appears in exactly the same way as in the proof of theorem 5 in appendix F, $N^{-|\rho|}$ appears as a volume in $\mathbb{R}^{|\rho|}$, and $N^{|\rho|s}$ comes from evaluation of the density in the points $x_i = \frac{2k_i\pi}{N}$, $1 \le i \le |\rho|$). Since $\frac{1}{|x|^s}$ has a bounded integral around 0, and since the sum still converges (it is dominated by (47)), (53) is

$$O\left(s(|\rho|-n)-1\right).$$

This has it's highest order when $|\rho| = n$, so that we can restrict to looking at 0_n. Note also that we may just as well assume that $p_\omega(x)$ is identical to $p_i|x-\omega_i|^{-s}$ at an interval around ω_i, since $\lim_{x\to\alpha_i} |x-\alpha_i|^s p_\omega(x) = p_i$ implies that

$$p_\omega(x) = p_i|x-\omega_i|^{-s} + k(x)|x-\omega_i|^{-s} \tag{54}$$

where $\lim_{x\to\omega_i} k(x) = 0$. It is straightforward to see that the contribution of the second part in (54) to (53) vanishes as $N\to\infty$, so that we may just as well assume that $p_\omega(x)$ is identical to $p_i|x-\omega_i|^{-s}$ at an interval around ω_i, as claimed. Also, since

$$\lim_{n\to\infty} \int_{|x|>\epsilon} x^{-s} e^{jnx} dx = 0$$

for all $\epsilon > 0$, and since the contributions from large n dominate in (55) below (since $\sum_n |n|^{-s}$ diverges), it is clear that we can restrict to an interval around ω_i when computing the limit also (since p_ω is continous outside the singularity points, this follows from theorem 5, and due to the additional $\frac{1}{N^s}$-factor added to (1)). After restricting to 0_n, multiplying both sides with N, summing over all singularity points, and using (52), we obtain the approximation

$$\begin{array}{l} \sum_{(i_1,\ldots,i_n)} \\ \sum_a \\ \qquad N^{-ns} c^{n-1} \times \\ \qquad \left(2p_a\Gamma(1-s)\cos\left(\frac{(1-s)\pi}{2}\right)\right)^n \times \\ \qquad \prod_{k=1}^n \frac{e^{j(i_{k-1}-i_k)\alpha_a}}{|i_{k-1}-i_k|^{1-s}} \\ \qquad tr_L(\mathbf{D}_1(N)) tr_L(\mathbf{D}_2(N)) \cdots tr_L(\mathbf{D}_n(N)) \end{array} \tag{55}$$

to (32). Since $\prod_{k=1}^n e^{j(i_{k-1}-i_k)\alpha_a} = 1$, we recognize

$$\begin{array}{rcl} q^{(n,N)} & = & \left(2\Gamma(1-s)\cos\left(\frac{(1-s)\pi}{2}\right)\right)^n \left(\sum_a p_a^n\right) \times \\ & & \sum_{(i_1,\ldots,i_n)} N^{-ns} \prod_{k=1}^n \frac{1}{|i_{k-1}-i_k|^{1-s}}, \end{array}$$

as a factor in (55) such that the limit of (55) as $N\to\infty$ can be written

$$c^{n-1} \lim_{N\to\infty} q^{(n,N)} \lim_{N\to\infty} \prod_{i=1}^n tr_L\left(\mathbf{D}_i(N)\right).$$

It therefore suffices to prove that $\lim_{N\to\infty} q^{(n,N)} = q^{(n)}$. To see this, write

$$\begin{aligned} \frac{N^{-s}}{|i_{k-1}-i_k|^{1-s}} &= \frac{1}{N} \frac{1}{\left(\frac{1}{N}\right)^{1-s} |i_{k-1}-i_k|^{1-s}} \\ &= \frac{1}{N} \frac{1}{\left|\frac{i_{k-1}}{N} - \frac{i_k}{N}\right|^{1-s}}. \end{aligned}$$

Summing over all $1 \le i_1, ..., i_n \le N$, it is clear from this that $q^{(n,N)}$ can be viewed as a Riemann sum which converges to $q^{(n)}$ as $N\to\infty$.

REFERENCES

[1] R. Norberg, "On the Vandermonde matrix and its application in mathematical finance," *working paper no. 162, Laboratory of Actuarial Mathematics, Univ. of Copenhagen*, 1999.

[2] R. Schmidt, "Multiple emitter localization and signal parameter estimation," in *Proceedings of the RADC, Spectal Estimation Workshop, Rome*, 1979, pp. 243–258.

[3] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 33, pp. 387–392, 1985.

[4] D. H. Johnson and D. E. Dudgeon, *Array Signal processing: Concepts and Techniques*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[5] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, pp. 984–995, July 1989.

[6] B. Porat and B. Friedlander, "Analysis of the asymptotic relative efficiency of the MUSIC algorithm," *IEEE Transactions Acoustics Speech and Signal Processing*, vol. 36, pp. 532–544, apr. 1988.

[7] A. Klein and P. Spreij, "On Stein's equation, Vandermonde matrices and Fisher's information matrix of time series processes. part I: The autoregressive moving average process." *Universiteit van Amsterdam, AE-Report 7/99*, 1999.

[8] L. Sampaio, M. Kobayashi, Ø. Ryan, and M. Debbah, "Vandermonde matrices for security applications," *work in progress*, 2008.

[9] ——, "Vandermonde frequency division multiplexing," *9th IEEE Workshop on Signal Processing Advances for wireless applications, Recife, Brazil*, 2008.

[10] Z. Wang, A. Scaglione, G. Giannakis, and S. Barbarossa, "Vandermonde-Lagrange mutually orthogonal flexible transceivers for blind CDMA in unknown multipath," in *Proc. of IEEE-SP Workshop on Signal Proc. Advances in Wireless Comm.*, May 1999, pp. 42–45.

[11] J. J. Waterfall, J. Joshua, F. P. Casey, R. N. Gutenkunst, K. S. Brown, C. R. Myers, P. W. Brouwer, V. Elser, and J. P. Sethna, "Sloppy-model universality class and the Vandermonde matrix," *Physical Review Letters*, vol. 97, no. 15, 2006.

[12] V. Girko, *Theory of Random Determinants*. Kluwer Academic Publishers, 1990.

[13] M. L. Mehta, *Random Matrices*, 2nd ed. New York: Academic Press, 1991.

[14] R. R. Muller, "A random matrix model of communication via antenna arrays," *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2495–2506, 2002.

[15] S. Thorbjørnsen, "Mixed moments of Voiculescu's Gaussian random matrices," *J. Funct. Anal.*, vol. 176, no. 2, pp. 213–246, 2000.

[16] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*. American Mathematical Society, 2000.

[17] T. Anderson, "Asymptotic theory for principal component analysis," *Annals of Mathematical Statistics*, vol. 34, pp. 122–148, mar. 1963.

[18] K. Abed-Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," *IEEE Trans. on Information Theory*, vol. 43, pp. 499–511, mar. 1977.

[19] Ø. Ryan and M. Debbah, "Free deconvolution for signal processing applications," *Submitted to IEEE Trans. on Information Theory*, 2007, http://arxiv.org/abs/cs.IT/0701025.

[20] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. www.nowpublishers.com, 2004.

[21] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*. Cambridge University Press, 2006.

[22] Ø. Ryan and M. Debbah, "Random Vandermonde matrices-part II: Applications," *Submitted to IEEE Trans. on Information Theory*, 2008.

[23] ——, "Channel capacity estimation using free probability theory," *Submitted to IEEE Trans. Signal Process.*, 2007, http://arxiv.org/abs/0707.3095.

[24] J. A. Mingo and R. Speicher, "Second order freeness and fluctuations of random matrices: I. Gaussian and Wishart matrices and cyclic Fock spaces," pp. 1–46, 2005, arxiv.org/math.OA/0405191.

[25] Ø. Ryan, *Tools for convolution with Vandermonde matrices*, 2008, http://ifi.uio.no/˜oyvindry/vandermonde/.

[26] K. Rottmann, *Mathematische Formelsammlung*. B.I. Wissenschaftsverlag, 1991.