

ON THE LIMITING MOMENTS OF VANDERMONDE RANDOM MATRICES

Oyvind Ryan, Merouane Debbah

▶ **To cite this version:**

Oyvind Ryan, Merouane Debbah. ON THE LIMITING MOMENTS OF VANDERMONDE RANDOM MATRICES. WORKSHOP ON PHYSICS-INSPIRED PARADIGMS IN WIRELESS COMMUNICATIONS AND NETWORKS, Apr 2008, Germany. 10 p. hal-00328165

HAL Id: hal-00328165

https://centralesupelec.hal.science/hal-00328165

Submitted on 9 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON THE LIMITING MOMENTS OF VANDERMONDE RANDOM MATRICES

Øyvind Ryan⋆, and Mérouane Debbah†

⋆ University of Oslo, Oslo, Norway
† SUPELEC, Gif-sur-Yvette, France
E-mail: oyvindry@ifi.uio.no,merouane.debbah@supelec.fr

ABSTRACT

In this paper, 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, sampling theory or precoding. 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. 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.

1. INTRODUCTION

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

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

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 $\frac{L}{N} \to c$. 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] 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 $\mathbf{V}$. While we provide the full computation of lower order moments,

This work was supported by Alcatel-Lucent within the Alcatel-Lucent chair on flexible radio.

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 2 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 3. 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 4 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. Section 5 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.

2. 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

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

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*

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

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

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

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

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

The proof of theorem 1 can be found in [21]. Although the limit of $K_{\rho,\omega,N}$ as $N \to \infty$ may not exist, it will be clear from section 4 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

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

(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

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

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

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

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

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

then, after performing the substitutions

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

(7) can be written

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

For the first 5 moments this becomes

$$\begin{aligned} m_1 &= K_1 d_1 \\ m_2 &= K_2 d_2 + K_{1,1} d_1^2 \\ m_3 &= K_3 d_3 + K_{2,1} d_2 d_1^2 + K_{1,1,1} d_1^3 \\ m_4 &= K_4 d_4 + K_{3,1} d_3 d_1 + K_{2,2} d_2^2 + K_{2,1,1} d_2 d_1^2 + \\ &\quad K_{1,1,1,1} d_1^4 \\ m_5 &= K_5 d_5 + K_{4,1} d_4 d_1 + + K_{3,2} d_3 d_2 + \\ &\quad K_{3,1,1} d_3 d_1^2 + K_{2,2,1} d_2^2 d_1 + K_{2,1,1,1} d_2 d_1^3 + \\ &\quad K_{1,1,1,1,1} d_1^5 \\ &\vdots \qquad \vdots \end{aligned} \tag{11}$$

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]. 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. $D_1 D_2$ should for instance be replaced with

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

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).

3. UNIFORMLY DISTRIBUTED ω

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

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

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 [22]. 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 [21]. 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 [21]. 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 [21]. To compute higher moments m_k, $K_{\rho,u}$ must be computed for partitions of higher order.

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:

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

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 [21]. 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 [21], 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.

4. 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*

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

The proof is given in [21].

Besides providing us with a deconvolution method for finding the mixed moments of the $\{\mathbf{D}_r(N)\}_{1\leq r\leq n}$, 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

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

for $0 \leq x \leq \infty$, 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

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

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

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

for any density p_ω. In [23], several examples are provided where the integrals (14) are computed.

5. 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.

5.1. 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 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. $\hat{M}_j = \frac{1}{K}\sum_{i=1}^K \hat{m}_{ji}$.

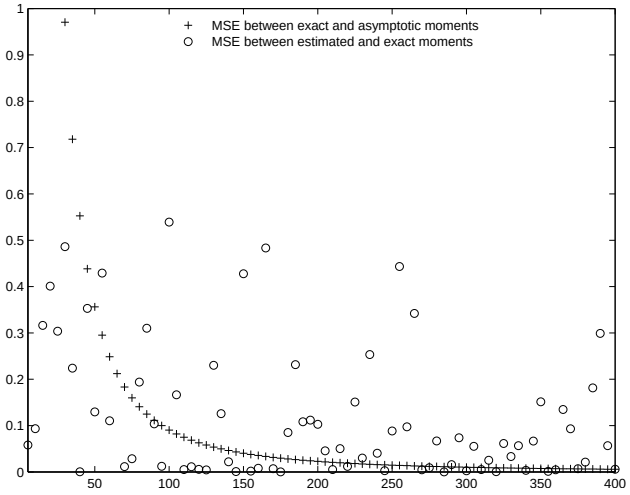


Fig. 1. MSE of the first 4 estimated moments from the exact moments for 80 samples for varying matrix sizes, with $N = L$. Matrices are on the form $\mathbf{V}^H\mathbf{V}$ with $\mathbf{V}$ a Vandermonde matrix with uniformly distributed phases. The MSE of the first 4 exact moments from the asymptotic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

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 $\sum_{j=1}^{4}\left(E_j - A_j\right)^2$.

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 [24]), 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.

5.2. 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

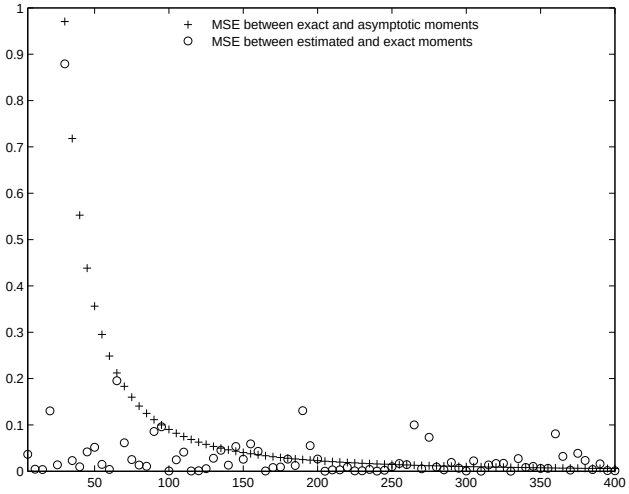


Fig. 2. MSE of the first 4 moments from the actual moments for 320 samples for varying matrix sizes, with $N = L$. Matrices are on the form $\mathbf{V}^H\mathbf{V}$ with $\mathbf{V}$ a Vandermonde matrix with uniformly distributed phases. The MSE of the moments and the asymptotic moments is also shown.

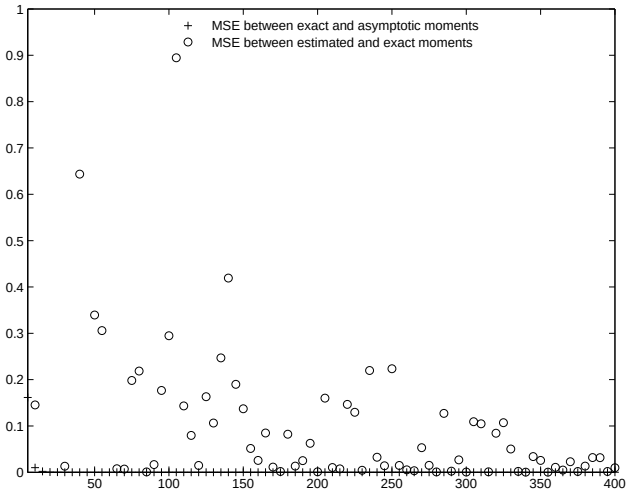


Fig. 3. MSE of the first 4 moments from the actual moments for 5 samples for varying matrix sizes, with $N = L$. Matrices are on the form $\frac{1}{N}\mathbf{X}\mathbf{X}^H$ with $\mathbf{X}$ a complex standard Gaussian matrix. The MSE of the moments and the asymptotic moments is also shown.

of

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

as $n\to\infty$ [22]. 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 [22, 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]

$$f^{\mu_c}(x)=(1-\frac{1}{c})^+\delta_0(x)+\frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi cx}, \quad (16)$$

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 $\lambda\alpha^n$ [22], 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

$$\begin{array}{rcl}\phi(a_1^n) & = & \sum_{\rho\in NC(n)} c^{n-|\rho|}=\sum_{\rho\in NC(n)} c^{|K(\rho)|-1} \\ & = & \sum_{\rho\in NC(n)} c^{|\rho|-1}.\end{array} \quad (17)$$

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

$$E(a_2^n)=\sum_{\rho\in\mathcal{P}(n)} c^{|\rho|}. \quad (18)$$

We immediately recognize the $c^{|\rho|-1}$-entry of theorem 1 in (17) and (18) (except for an additional power of c in (18)). Combining theorem 2 with $\mathbf{D}_1(N)=\cdots=\mathbf{D}_n(N)=\mathbf{I}_N$, (17), and (18), we thus get the following corollary to theorem 2:

Corollary 1 *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

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

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 1 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 2 *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

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

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 $C_n=\frac{1}{n+1}\binom{2n}{n}$ [22], so they are easily computed. The number of partitions of $\mathcal{P}(n)$ are also known as the Bell numbers B_n [22]. 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 2 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

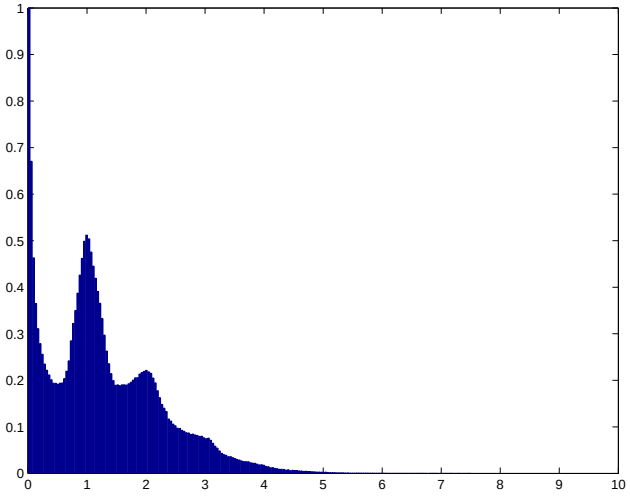


Fig. 4. Histogram of the mean eigenvalue distribution of 640 samples of $\mathbf{V}^H\mathbf{V}$, with $\mathbf{V}$ a 1600×1600 Vandermonde matrix with uniformly distributed phases.

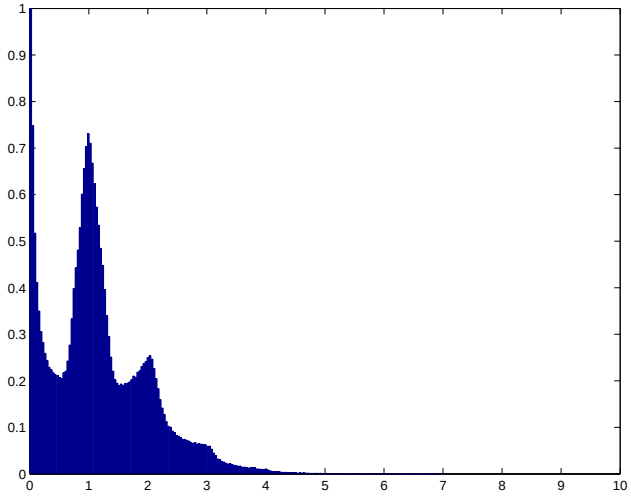


Fig. 5. Histogram of the mean eigenvalue distribution of 640 samples of $\mathbf{V}^H\mathbf{V}$, with $\mathbf{V}$ a 1600×1200 Vandermonde matrix with uniformly distributed phases.

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 1) 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.

5.3. Deconvolution

Deconvolution with Vandermonde matrices (as stated in (6) in theorem 1) differs from the Gaussian deconvolution counterpart [22] in the sense that there is no multiplicative [22] 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

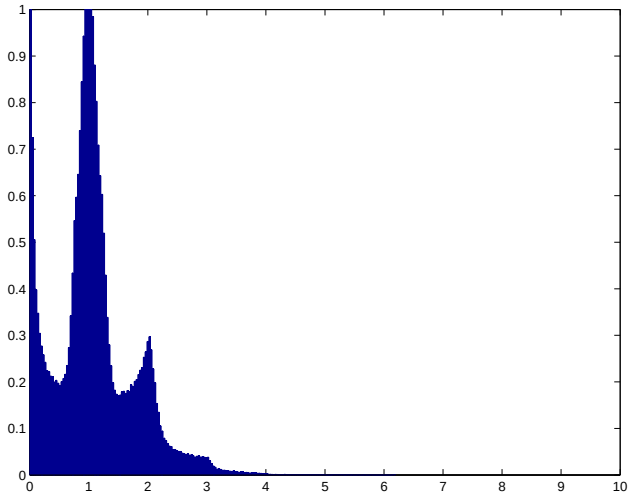


Fig. 6. Histogram of the mean eigenvalue distribution of 640 samples of $\mathbf{V}^H\mathbf{V}$, with $\mathbf{V}$ a 1600×800 Vandermonde matrix with uniformly distributed phases.

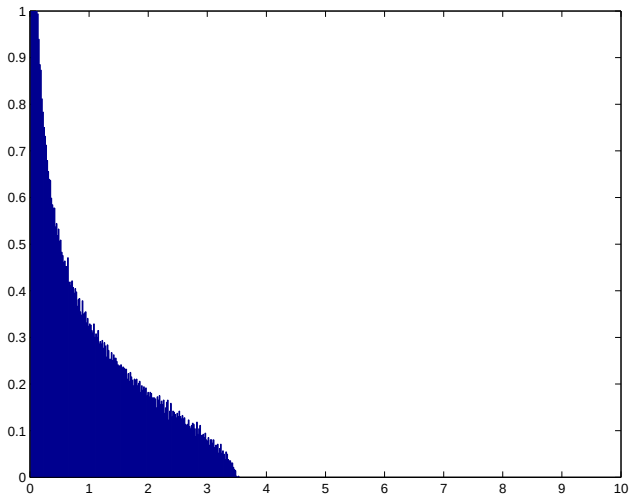


Fig. 7. Histogram of the mean eigenvalue distribution of 20 samples of $\frac{1}{N}\mathbf{X}\mathbf{X}^H$, with $\mathbf{X}$ an $L \times N = 1200 \times 1600$ complex, standard, Gaussian matrix.

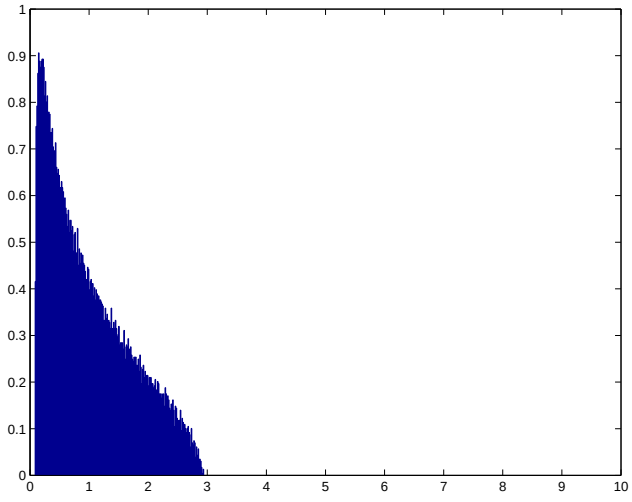


Fig. 8. Histogram of the mean eigenvalue distribution of 20 samples of $\frac{1}{N}\mathbf{X}\mathbf{X}^H$, with $\mathbf{X}$ an $L \times N = 800 \times 1600$ complex, standard, Gaussian matrix.

matrix) is

$$\begin{aligned}
m_1 &= d_1 \\
m_2 &= d_2 + d_1^2 \\
m_3 &= d_3 + 3d_2d_1 + d_1^3 \\
m_4 &= d_4 + 4d_3d_1 + 3d_2^2 + 6d_2d_1^2 + d_1^4 \\
m_5 &= d_5 + 5d_4d_1 + 5d_3d_2 + 10d_3d_1^2 + \\
&\quad 10d_2^2d_1 + 10d_2d_1^3 + d_1^5 \\
m_6 &= d_6 + 6d_5d_1 + 6d_4d_2 + 15d_4d_1^2 + \\
&\quad 3d_3^2 + 30d_3d_2d_1 + 20d_3d_1^3 + \\
&\quad 5d_2^3 + 10d_2^2d_1^2 + 15d_2d_1^4 + d_1^6 \\
m_7 &= d_7 + 7d_6d_1 + 7d_5d_2 + 21d_5d_1^2 + \\
&\quad 7d_4d_3 + 42d_4d_2d_1 + 35d_4d_1^3 + \\
&\quad 21d_3^2d_1 + 21d_3d_2^2 + 105d_3d_2d_1^2 + \\
&\quad 35d_3d_1^4 + 35d_2^3d_1 + 70d_2^2d_1^3 + \\
&\quad 21d_2d_1^5 + d_1^7,
\end{aligned} \tag{19}$$

(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 2. One can also write down a Gaussian equivalent to the fluctuations of Vandermonde matrices (13) (fluctuations of Gaussian matrices are handled more thoroughly in [25]). These are

$$\begin{aligned}
&E\left[\left(tr_n\left(\mathbf{D}(N)\tfrac{1}{N}\mathbf{X}\mathbf{X}^H\right)\right)^2\right] \\
&= (tr_n(\mathbf{D}(N))^2 + \tfrac{1}{nN}tr_n(\mathbf{D}(N)^2) \\
&E\left[\left(tr_n\left(\mathbf{D}(N)\tfrac{1}{N}\mathbf{X}\mathbf{X}^H\right)\right)^n\right] \\
&= (tr_n(\mathbf{D}(N))^n + O(N^{-2}) \\
&E\left[tr_n\left(\mathbf{D}(N)\tfrac{1}{N}\mathbf{X}\mathbf{X}^H\right) tr_n\left(\left(\mathbf{D}(N)\tfrac{1}{N}\mathbf{X}\mathbf{X}^H\right)^2\right)\right] \\
&= tr_n(\mathbf{D}(N))tr_n(\mathbf{D}(N)^2) + O(N^{-2}).
\end{aligned} \tag{20}$$

These equations can be proved using the same combinatorical methods as in [24]. 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

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

This follows from a careful count of all possibilities after the matrices have been multiplied together (for this, see also [24], 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 (19), 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).

6. 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. 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 probability (such as the R- and S-transform and their connection with the Stieltjes transform).

7. 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*, Prentice Hall, Englewood Cliffs, NJ, 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] L. Sampaio, M. Kobayashi, Ø. Ryan, and M. Debbah, "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.L. Girko, *Theory of Random Determinants*, Kluwer Academic Publishers, 1990.

[13] M. L. Mehta, *Random Matrices*, Academic Press, New York, 2nd edition, 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] Ø. Ryan and M. Debbah, "Random Vandermonde matrices-part I: Fundamental results," *Submitted to IEEE Trans. on Information Theory*, 2008.

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

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

[24] Ø. Ryan and M. Debbah, "Channel capacity estimation using free probability theory," *Submitted to IEEE Trans. Signal Process.*, 2007, http://arxiv.org/abs/0707.3095.

[25] 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.