The Dirichlet Markov Ensemble

Djalil Chafai

The Dirichlet Markov Ensemble

Djalil Chafaï

Preprint, September 2007

Abstract

We equip the polytope of $n \times n$ Markov matrices with the normalized trace of the Lebesgue measure of $\mathbb{R}^{n^2}$. This probability space provides random Markov matrices, with i.i.d. rows following the Dirichlet distribution of mean $(1/n, \ldots, 1/n)$. We show that if $\mathbf{M}$ is such a random matrix, then the empirical spectral distribution of $n\mathbf{M}\mathbf{M}^\top$ tends as $n \to \infty$ to a Marchenko-Pastur distribution. This phenomenon complements an already known result on the sub-dominant eigenvalue of certain random matrices with independent rows, which suggests that the typical spectral gap of a uniform random Markov matrix is of order $1 - 1/\sqrt{n}$ when n is large. However, some computer simulations reveal striking asymptotic spectral properties of such random matrices, still waiting for a rigorous mathematical analysis. In particular, we conjecture that the empirical distribution of the complex spectrum of $\sqrt{n}\mathbf{M}$ tends as $n \to \infty$ to the uniform distribution on the unit disc of the complex plane.

AMS 2000 Mathematical Subject Classification: 60F15; 15A52; 62H99.

Keywords: Random matrices; Markov matrices, Dirichlet distributions; Spectral gap; singular values.

1 Introduction

An $n \times n$ square real matrix $\mathbf{M}$ is Markov if and only if its entries are non-negative and each line sum up to 1, i.e. if and only if each row of $\mathbf{M}$ belongs to the simplex

$$\Lambda_n = \{(x_1, \ldots, x_n) \in [0,1]^n \text{ such that } x_1 + \cdots + x_n = 1\} \tag{1}$$

which is the portion of the unit $\|\cdot\|_1$-sphere of $\mathbb{R}^n$ with non-negative coordinates. The spectrum of a Markov matrix lies in the unit disc $D(0,1) = \{z \in \mathbb{C}; |z| \leqslant 1\}$, contains 1, and is symmetric with respect to the real axis in the complex plane.

Uniform distribution on Markov matrices

Let $\mathcal{M}_n$ be the set of $n \times n$ Markov matrices. We need to give a precise meaning to the notion of uniform distribution on $\mathcal{M}_n$. This set is a convex compact polytope with $n(n-1)$ degrees of freedom if $n > 1$. It has zero Lebesgue measure in $\mathbb{R}^{n^2}$.

Since $\mathcal{M}_n$ is a polytope of $\mathbb{R}^{n^2}$, the trace of the Lebesgue measure on it makes sense[1], despite its zero Lebesgue measure in $\mathbb{R}^{n^2}$. Since $\mathcal{M}_n$ is additionally compact, the trace of the Lebesgue measure can be normalized into a probability distribution. We thus define *the uniform distribution* $\mathcal{U}(\mathcal{M}_n)$ on $\mathcal{M}_n$ as the normalized trace of the Lebesgue measure of $\mathbb{R}^{n^2}$. The following theorem relates $\mathcal{U}(\mathcal{M}_n)$ to the Dirichlet distribution.

Theorem 1.1 (Dirichlet Markov Ensemble). *We have $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$ if and only if the rows of $\mathbf{M}$ are i.i.d. and follow the Dirichlet law of mean $(\frac{1}{n}, \ldots, \frac{1}{n})$. The probability distribution $\mathcal{U}(\mathcal{M}_n)$ is invariant by permutations of rows and columns.*

The set $\mathcal{M}_n$ is also a compact semi-group for the matrix product. The following two theorems concern the translation invariance of $\mathcal{U}(\mathcal{M}_n)$ and the question of the existence of an idempotent probability distribution on $\mathcal{M}_n$.

Theorem 1.2 (Translation invariance). *For every $\mathbf{T} \in \mathcal{M}_n$, the law $\mathcal{U}(\mathcal{M}_n)$ is invariant by the left translation $\mathbf{M} \mapsto \mathbf{TM}$ if and only if $\mathbf{T}$ is a permutation matrix. The same holds true for the right translation $\mathbf{M} \mapsto \mathbf{MT}$.*

Theorem 1.3 (Idempotent distributions). *There is no probability distribution on $\mathcal{M}_n$, absolutely continuous with respect to $\mathcal{U}(\mathcal{M}_n)$, with full support, and which is invariant by every left translations $\mathbf{M} \mapsto \mathbf{TM}$ where $\mathbf{T}$ runs over $\mathcal{M}_n$. The same holds true for right translations.*

The proofs of theorems 1.1, 1.2, and 1.3 are given in section 2.

Singular values and the Marchenko-Pastur distribution

The study of the spectral properties of large dimensional random matrices is a very active topic, connected to many areas of mathematics, see for instance the books [26], [24], [6] and the survey [5]. If $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$, then almost surely, the real matrix $\mathbf{M}$ is invertible, non-normal, with neither independent nor centered entries. The singular values of certain large dimensional centered random matrices with independent rows is considered for instance in [2] and [27].

[1] Actually, one can define the trace of the Lebesgue measure and then the uniform distribution on many compact subsets of the Euclidean space, by using the notion of Hausdorff measure [19]. See also [15] for an approximate simulation method based on billiards and random reflexions.

For any square $n \times n$ matrix $\mathbf{A}$ with real or complex entries, let the complex eigenvalues $\lambda_1(\mathbf{A}), \dots, \lambda_n(\mathbf{A})$ of $\mathbf{A}$ be labeled so that $|\lambda_1(\mathbf{A})| \geqslant \cdots \geqslant |\lambda_n(\mathbf{A})|$. The *empirical spectral distribution* (ESD) of $\mathbf{A}$ is the discrete probability distribution on $\mathbb{C}$ with at most n atoms defined by

$$\frac{1}{n}\sum_{k=1}^{n} \delta_{\lambda_k(\mathbf{A})}.$$

The singular values $s_1(\mathbf{A}) \geqslant \cdots \geqslant s_n(\mathbf{A}) \geqslant 0$ of $\mathbf{A}$ are the eigenvalues of the positive semi-definite Hermitian matrix $(\mathbf{A}\mathbf{A}^*)^{1/2}$ where

$$\mathbf{A}^* = \overline{\mathbf{A}}^\top$$

denotes the conjugate transpose of $\mathbf{A}$. Namely, for every $1 \leqslant k \leqslant n$,

$$s_k(\mathbf{A}) = \lambda_k(\sqrt{\mathbf{A}\mathbf{A}^*}) = \sqrt{\lambda_k(\mathbf{A}\mathbf{A}^*)}.$$

In particular, the atoms of the ESD of $(\mathbf{A}\mathbf{A}^*)^{1/2}$ are $s_1(\mathbf{A}), \dots, s_n(\mathbf{A})$. If $\mathbf{A}$ is normal, i.e. $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$, then $s_k(\mathbf{A}) = |\lambda_k(\mathbf{A})|$ for every $1 \leqslant k \leqslant n$. Back to our Dirichlet Markov Ensemble, if $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$ then $\mathbf{M}$ is almost surely a non-normal matrix, and thus one cannot express the singular values of $\mathbf{M}$ in terms of the eigenvalues of $\mathbf{M}$. The following theorem gives the asymptotic behavior of distribution built from the singular values of $\mathbf{M}$.

Theorem 1.4 (Singular values for Dirichlet Markov Ensemble). *Let $(X_{i,j})_{1 \leqslant i,j < \infty}$ be an infinite array of i.i.d. exponential random variables of unit mean. For every n, let $\mathbf{M}$ be the $n \times n$ random matrix defined for every $1 \leqslant i,j \leqslant n$ by*

$$\mathbf{M}_{i,j} = \frac{X_{i,j}}{\sum_{k=1}^n X_{i,k}}.$$

Then $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$, and almost surely, the ESD of $n\mathbf{M}\mathbf{M}^\top$ tends as $n \to \infty$ to the Marchenko-Pastur distribution of density

$$x \mapsto \frac{1}{2\pi x}\sqrt{(4-x)x}\ \mathrm{I}_{[0,4]}(x).$$

As a consequence, almost surely, the ESD of $\sqrt{n\mathbf{M}\mathbf{M}^\top}$ tends as $n \to \infty$ to the quarter-circle Wigner distribution of density

$$x \mapsto \frac{1}{\pi}\sqrt{4-x^2}\ \mathrm{I}_{[0,2]}(x).$$

The proof of theorem 1.4 is given in section 3. Since $|\lambda_1(\mathbf{A})| \leqslant s_1(\mathbf{A})$ for any square matrix $\mathbf{A}$, and since $\lambda_1(\mathbf{M}) = 1$, we have

$$\lambda_1(n\mathbf{M}\mathbf{M}^\top) = ns_1(\mathbf{M})^2 \geqslant n|\lambda_1(\mathbf{M})|^2 = n \underset{n\to\infty}{\longrightarrow} +\infty.$$

However, theorem 1.4 implies in particular that almost surely

$$\frac{1}{n}\,\mathrm{Card}\{1 \leqslant k \leqslant n \text{ such that } \lambda_k(n\mathbf{M}\mathbf{M}^\top) > 4\} \underset{n\to\infty}{\longrightarrow} 0.$$

Random Q-matrices

Bryc, Dembo, and Jiang studied in [13] the limiting spectral distribution of *random Hankel, Markov, and Toeplitz matrices*. Let us explain briefly what they mean by "random Markov matrices". They proved the following theorem (see [13, Theorem 1.3] and also [33]) : let $(\mathbf{X}_{i,j})_{1<i<j<\infty}$ be an infinite triangular array of i.i.d. real random variables of mean 0 and variance 1. Let $\mathbf{Q}$ be the symmetric $n \times n$ random matrix defined for every $1 \leqslant i \leqslant j \leqslant n$ by $\mathbf{Q}_{i,j} = \mathbf{Q}_{j,i} = X_{i,j}$ if $i < j$, and

$$\mathbf{Q}_{i,i} = -\sum_{\substack{1\leqslant k\leqslant n\\ k\neq i}} \mathbf{Q}_{i,k} \quad \text{for every } 1 \leqslant i \leqslant n.$$

Then, almost surely, the ESD of $n^{-1/2}\mathbf{Q}$ converges as $n \to \infty$ to the free convolution[2] of a semi-circle law and a standard Gaussian law.

This result gives an answer to a precise question raised by Bai in his 1999 review article [5, Section 6.1.1 page 657]. The matrix $\mathbf{Q}$ is not Markov. However, it looks like a Markov generator, i.e. a Q-matrix, since its rows sum up to 0. Unfortunately, the assumptions do not allow the off-diagonal entries of $\mathbf{Q}$ to have non-negative support, and thus $\mathbf{Q}$ cannot be almost surely a Markov generator. In particular, if $\mathbf{I}$ stands for the identity matrix of size $n \times n$, the symmetric matrix $\mathbf{M} = \mathbf{Q} + \mathbf{I}$ cannot be almost surely Markov.

Eigenvalues and the circular law

If $\mathbf{M}$ is as in theorem 1.4, then $\lambda_1(\sqrt{n}\mathbf{M}) = \sqrt{n}$ goes to $+\infty$ as $n \to \infty$ while its weight in the ESD is $1/n$. Thus, it does not contribute to the limiting spectral distribution of $\sqrt{n}\mathbf{M}$. Numerical simulations (see figure 1) suggest that the empirical distribution of the rest of the spectrum tends as $n \to \infty$ to the uniform distribution on the unit disc. One can formulate this conjecture as follows.

[2] This limiting spectral distribution is a symmetric law on $\mathbb{R}$ with smooth bounded density of unbounded support. See [24] or [11] for Voiculescu's free convolution.

Conjecture 1.5 (Circle law for the Dirichlet Markov Ensemble). *If* $\mathbf{M}$ *is as is theorem 1.4, then almost surely, the ESD of* $\sqrt{n}\mathbf{M}$ *tends as* $n \to \infty$ *to the uniform distribution over the unit disc* $D(0,1) = \{z \in \mathbb{C}; |z| \leqslant 1\}$.

For $\mathbf{M}$ itself, apart $\lambda_1(\mathbf{M}) = 1$, the spectrum concentrates around 0 at speed $1/\sqrt{n}$, suggesting a typical spectral gap $1 - \lambda_2(\mathbf{M})$ of order $1 - 1/\sqrt{n}$ for large n.

The main difficulty in conjecture 1.5 lies in the fact that $\mathbf{M}$ is non-normal with non i.i.d. entries. The limiting spectral distributions of non-normal random matrices is an active and notoriously difficult subject, see for instance [4], [6, ch. 10], [22, 23], [28], and [37] for the centered case, and [14] for the non-centered case. The method used for the singular values for the proof of theorem 1.4 fails for the eigenvalues, due to the lack of variational formulas for the eigenvalues. In contrast to singular values, the eigenvalues of non-normal matrices are very sensitive to perturbations, a phenomenon captured by the notion of pseudo-spectrum [38].

The uniform distribution on the unit disc $D(0,1)$ is known as the *circle* or *circular law*. If $U = re^{\sqrt{-1}\theta}$ is a complex random variable distributed according to the uniform distribution on the disc $D(0,2\sigma)$ of radius $2\sigma > 0$, the module r and the argument θ or U are independent with joint law of density

$$(r,\theta) \mapsto \frac{1}{4\pi\sigma^2}\, r\, \mathrm{I}_{[0,2\pi]}(\theta)\mathrm{I}_{[0,2\sigma]}(r).$$

Both the real part $\mathfrak{Re}(U)$ and the imaginary part $\mathfrak{Im}(U)$ of U follow the Wigner semi-circle law on $\mathbb{R}$ with density

$$x \mapsto \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\, \mathrm{I}_{[-2\sigma,+2\sigma]}(x). \tag{2}$$

More generally, for any angle $\alpha \in [0, 2\pi)$, the random variables $\mathfrak{Re}(e^{\sqrt{-1}\alpha}U)$ and $\mathfrak{Im}(e^{\sqrt{-1}\alpha}U)$ follow the Wigner semi-circle law mentioned above. Additionally, if a real random variable S follows the Wigner semi-circle law mentioned above, then its square S^2 follows the Marchenko-Pastur law on $\mathbb{R}_+$ with density

$$x \mapsto \frac{1}{2\pi x\sigma^2}\, \sqrt{(4\sigma^2 - x)x}\, \mathrm{I}_{[0,4\sigma^2]}(x). \tag{3}$$

Sub-dominant eigenvalue

The precise asymptotic behavior of the extremal eigenvalues of non-normal random matrices is not well understood, even in the i.i.d. centered entries case, see for instance [30, 29]. The fact that non-centered entries produce an explosive extremal eigenvalue was already noticed in various situations, see [1], [35], [13, th. 1.4], [12], and [14]. It is natural to ask about the asymptotic behavior (convergence and fluctuations) of the sub-dominant eigenvalue $\lambda_2(\mathbf{M})$ when $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$. The reader may find some answers in [20] and [21], and may forge new conjectures from our simulations (see figures 2 and 3). For instance, one can expect that $\sqrt{n}|\lambda_2(\mathbf{M})| \to 1$ a.s. as $n \to \infty$ and that $b_n(|\lambda_2(\mathbf{M})| - a_n)$ converges in distribution as $n \to \infty$ to some Gumbel type extreme distribution, for some deterministic sequences (a_n) and (b_n). Goldberg and Neumann have shown [20] that if $\mathbf{X}$ is an $n \times n$ random matrix with i.i.d. rows such that for every $1 \leqslant i, j, j' \leqslant n$,

$$\mathbb{E}[\mathbf{X}_{i,j}] = \frac{1}{n}, \quad \text{and} \quad \mathrm{Var}(\mathbf{X}_{i,j}) = \mathcal{O}\left(\frac{1}{n^2}\right), \quad \text{and} \quad |\mathrm{Cov}(\mathbf{X}_{i,j}, \mathbf{X}_{i,j'})| = \mathcal{O}\left(\frac{1}{n^3}\right)$$

then $\mathbb{P}(\lambda_2(\mathbf{X}) \leqslant 1 + \varepsilon) \geqslant p$ for any $p \in (0, 1)$, any $\varepsilon > 0$, and large enough n.

Other distributions

The Dirichlet distribution of dimension n and mean $(\frac{1}{n}, \ldots, \frac{1}{n})$ is the uniform distribution on the simplex Λ_n defined by (1). One can replace the uniform distribution by a Dirichlet distribution of dimension n and arbitrary mean. The argument used in the proof of theorem 1.4 remains the same due to the very similar construction of Dirichlet distributions by projection from i.i.d. Gamma random variables. One can also replace the $\|\cdot\|_1$-norm by any other $\|\cdot\|_p$-norm, and investigate the limiting spectral distribution of the corresponding random matrices. This case can be handled with the construction of the uniform distribution by projection proposed in [32]. Replacing the non-negative portion of spheres by the non-negative portion of balls is also possible by using [8]. More generally, one can consider random matrices with independent rows. The case of the uniform distribution on the whole unit $\|\cdot\|_p$-ball of $\mathbb{R}^n$ is considered for instance in [2] by using [8] together with random matrices results for i.i.d. centered entries. It is crucial here to have an explicit construction of the distribution from an i.i.d. array. For the link with the sampling of convex bodies, see [3].

Bistochastic matrices

The Birkhoff polytope is the set of $n \times n$ bistochastic matrices, i.e. matrices which are Markov and have a Markov transpose. This polytope is a convex compact subset of $\mathcal{M}_n$ of zero Lebesgue measure in $\mathbb{R}^{n^2}$ and $(n-1)^2$ degrees of freedom if $n > 1$. As for $\mathcal{M}_n$, one can define the uniform distribution as the normalized trace of the Lebesgue measure. However, we ignore if this distribution has a probabilistic representation that allows simulation as for $\mathcal{U}(\mathcal{M}_n)$. The spectral behavior of random bistochastic matrices was considered in the Physics literature, see for instance [10]. On the purely discrete side, the Birkhoff polytope is also related to magic squares, transportation polytopes and contingency tables, see

[17, 16] and [18]. Notice also that if $\mathbf{M}$ is Markov, then $\mathbf{M}\mathbf{M}^\top$ and $\frac{1}{2}(\mathbf{M}+\mathbf{M}^\top)$ are not Markov in general. However, it is true when $\mathbf{M}$ is bistochastic.

Another interesting polytope of matrices is the set of symmetric $n \times n$ Markov matrices, which is a convex compact polytope of zero Lebesgue measure in $\mathbb{R}^{n^2}$ with $\frac{1}{2}n(n-1)$ degrees of freedom if $n > 1$. As for $\mathcal{M}_n$, one can define the uniform distribution as the normalized trace of the Lebesgue measure. However, we ignore if this distribution has a probabilistic representation that allows simulation as for $\mathcal{U}(\mathcal{M}_n)$. One can ask about the spectral properties of the corresponding random symmetric Markov matrices. Notice that these matrices are bistochastic, but the converse is false except when $n = 1$ or $n = 2$.

Let $\mathbf{M}$ be as in theorem 1.4. Numerical simulations suggest that almost surely, the ESD of the symmetric matrix $\frac{1}{2}(\mathbf{M}+\mathbf{M}^\top)$ tends, as $n \to \infty$, to the semi-circle Wigner distribution of density (2) with $\sigma^2 = 1/2$.

If $\mathbf{U}$ is an $n \times n$ unitary matrix, then $(|\mathbf{U}_{i,j}|^2)_{1\leqslant i,j\leqslant n}$ is a bistochastic matrix. These bistochastic matrices are called *uni-stochastic* or *unitary-stochastic*. There exists bistochastic matrices which are not uni-stochastic, see [9] and [36]. However, every permutation matrix is orthogonal and thus uni-stochastic. The Haar measure on the unitary group induces a probability distribution on the set of uni-stochastic matrices. How about the asymptotic spectral properties of the corresponding random matrices?

Perron-Frobenius eigenvector

If $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$, then almost surely, all the entries of $\mathbf{M}$ are non-zero, and in particular, $\mathbf{M}$ is almost surely recurrent irreducible and aperiodic. By a standard theorem of Perron and Frobenius, it follows that almost surely, the eigenspace of $\mathbf{M}^\top$ associated to the eigenvalue 1 is of dimension 1 and contains a unique vector with non-negative entries and unit $\|\cdot\|_1$-norm. One can ask about the asymptotic behavior of this vector as $n \to \infty$. For a fixed n, the distribution of this vector is the distribution of the rows of the infinite product of random matrices $\lim_{k\to\infty} \mathbf{M}^k$.

2 Structure of the Dirichlet Markov Ensemble

Let Λ_n be as in (1). For any $a \in (0,\infty)^n$, the Dirichlet distribution $\mathcal{D}_n(a_1,\ldots,a_n)$, supported by Λ_n, is defined as the distribution of

$$\frac{1}{\|G\|_1}G = \left(\frac{G_1}{G_1+\cdots+G_n},\ldots,\frac{G_n}{G_1+\cdots+G_n}\right)$$

where G is a random vector of $\mathbb{R}^n$ with independent entries with $G_i \sim \mathrm{Gamma}(1, a_i)$ for every $1 \leqslant i \leqslant n$. Here $\mathrm{Gamma}(\lambda, a)$ has density

$$t \mapsto \frac{\lambda^a}{\Gamma(a)} t^{a-1} e^{-\lambda t} \, \mathrm{I}_{(0,\infty)}(t),$$

where $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} \, dt$ is the Euler Gamma function. Let $P \sim \mathcal{D}_n(a_1, \ldots, a_n)$. For every partition $I_1, \ldots, I_k$ of $\{1, \ldots, n\}$ into k non empty subsets, we have

$$\left(\sum_{i \in I_1} P_i, \ldots, \sum_{i \in I_k} P_i \right) \sim \mathcal{D}_k \left(\sum_{i \in I_1} a_i, \ldots, \sum_{i \in I_k} a_i \right).$$

The mean and covariance matrix of $\mathcal{D}_n(a_1, \ldots, a_n)$ are given by

$$\frac{1}{\|a\|_1} a \quad \text{and} \quad \frac{1}{\|a\|_1^2 (1 + \|a\|_1)} (\|a\|_1 \mathrm{diag}(a) - a a^\top)$$

where $a = (a_1, \ldots, a_n)^\top$ and $\mathrm{diag}(a)$ is the diagonal matrix with diagonal given by a. For any non-empty subset I of $\{1, \ldots, n\}$, we have

$$\sum_{i \in I} P_i \sim \mathrm{Beta} \left(\sum_{i \in I} a_i, \sum_{i \notin I} a_i \right),$$

where $\mathrm{Beta}(\alpha, \beta)$ denotes the Euler Beta distribution on $[0, 1]$ of Lebesgue density

$$t \mapsto \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha-1} (1 - t)^{\beta-1} \, \mathrm{I}_{[0,1]}(t).$$

If $P_I = (P_i)_{i \in I}$, $P_{I^c} = (P_i)_{i \notin I}$, $a_I = (a_i)_{i \in I}$, and $|I| = \mathrm{card}(I)$, then

$$\frac{1}{\sum_{i \in I} P_i} P_I \text{ and } P_{I^c} \text{ are independent and } \frac{1}{\sum_{i \in I} P_i} P_I \sim \mathcal{D}_{|I|}(a_I),$$

For any $\alpha > 0$, the Dirichlet distribution $\mathcal{D}_n(\alpha, \ldots, \alpha)$ is exchangeable, with negatively correlated components. More generally, if $P \sim \mu$ where μ is an exchangeable probability distribution on the simplex Λ_n with $n > 1$, then

$$0 = \mathrm{Var}(1) = \mathrm{Var}(P_1 + \cdots + P_n) = n \mathrm{Var}(P_1) + n(n-1) \mathrm{Cov}(P_1, P_2).$$

Consequently, $\mathrm{Cov}(P_1, P_2) = -(n-1)^{-1} \mathrm{Var}(P_1)$ and in particular $\mathrm{Cov}(P_1, P_2) \leqslant 0$.

Proof of theorem 1.1. As a subset of $\mathbb{R}^n$, the simplex Λ_n defined by (1) is of zero Lebesgue measure. However, by considering Λ_n as a convex subset of the hyperplane of equation $x_1 + \cdots + x_n = 1$ or by using the general notion of Hausdorff measure, one can see that in fact, the Dirichlet distribution $\mathcal{D}_n(1,\dots,1)$ is the normalized trace of the Lebesgue measure of $\mathbb{R}^n$ on the simplex Λ_n. In other words, $\mathcal{D}_n(1,\dots,1)$ can be seen as the uniform distribution on Λ_n, see [32].

We identify $\mathcal{M}_n$ with $(\Lambda_n)^n = \Lambda_n \times \cdots \times \Lambda_n$ where Λ_n is repeated n times. The trace of the Lebesgue measure of $\mathbb{R}^{n^2} = (\mathbb{R}^n)^n$ on $(\Lambda_n)^n$ is the n-tensor product of the trace of the Lebesgue measure of $\mathbb{R}^n$ on Λ_n, i.e. the n-tensor product measure $\mathcal{D}_n(1,\dots,1)^{\otimes n}$. Consequently, for every positive integer n,

$$(\mathcal{M}_n, \mathcal{U}(\mathcal{M}_n)) = ((\Lambda_n)^n, \mathcal{D}_n(1,\dots,1)^{\otimes n}).$$

This gives the invariance of $\mathcal{U}(\mathcal{M}_n)$ by permutation of rows. If $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$, then the rows of $\mathbf{M}$ are i.i.d. and follow the Dirichlet distribution $\mathcal{D}_n(1,\dots,1)$. In particular, for every $1 \leqslant i,j \leqslant n$, $\mathbf{M}_{i,j} \sim \mathrm{Beta}(1, n-1)$ and

$$\mathbb{E}[\mathbf{M}_{i,j}] = \frac{1}{n}$$

and

$$\mathrm{Cov}(\mathbf{M}_{i,j}, \mathbf{M}_{i',j'}) = \begin{cases} 0 & \text{if } i \neq i' \\ \frac{n-1}{n^2(n+1)} & \text{if } i = i' \text{ and } j = j' \\ -\frac{1}{n^2(n+1)} & \text{if } i = i' \text{ and } j \neq j' \end{cases}$$

for every $1 \leqslant i, i', j, j' \leqslant n$. The entries $\mathbf{M}_{i,j}$ and $\mathbf{M}_{i',j'}$ are independent if and only if $i \neq i'$. Finally, the invariance of $\mathcal{U}(\mathcal{M}_n)$ by permutation of columns comes from the exchangeability of the Dirichlet distribution $\mathcal{D}_n(1,\dots,1)$. $\square$

Recursive simulation

The simulation of $\mathcal{U}(\mathcal{M}_n)$ follows from the simulation of n i.i.d. realizations of $\mathcal{D}_n(1,\dots,1)$ by using n^2 i.i.d. exponential random variables. The elements of Dyson's classical Gaussian ensembles GUE and GOE can be simulated recursively by adding a new independent line/column. It is natural to ask about a recursive method for the Dirichlet Markov Ensemble. If

$$X \sim \mathcal{D}_{n-1}(a_2,\dots,a_n) \quad \text{and} \quad Y \sim \mathrm{Beta}(a_1, a_2 + \cdots + a_n)$$

are independent, then

$$(Y, (1-Y)X) \sim \mathcal{D}_n(a_1,\dots,a_n).$$

This recursive simulation of Dirichlet distributions is known as the *stick-breaking* algorithm [34]. It allows to simulate $\mathcal{U}(\mathcal{M}_n)$ recursively on n. Namely, if $\mathbf{M}$ is such that $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$, then

$$\begin{pmatrix} Y & (1-Y)\cdot\mathbf{M} \\ Z_1 & Z_2 \ \cdots \ Z_n \end{pmatrix} \sim \mathcal{U}(\mathcal{M}_{n+1})$$

where Z is a random row vector of $\mathbb{R}^{n+1}$ with $Z \sim \mathcal{D}_{n+1}(1,\ldots,1)$ and Y is a random column vector of $\mathbb{R}^n$ with i.i.d. entries of law $\mathrm{Beta}(1,n)$, with $\mathbf{M}, Y, Z$ independent. Here $((1-Y)\cdot\mathbf{M})_{i,j} := (1-Y)_i\mathbf{M}_{i,j}$ for every $1 \leqslant i,j \leqslant n$.

Semi-group structure and translation invariance

The set $\mathcal{M}_n$ is a semi-group for the usual matrix product. In particular, for every $\mathbf{T} \in \mathcal{M}_n$, the set $\mathcal{M}_n$ is stable by the left translation $\mathbf{M} \mapsto \mathbf{TM}$ and the right translation $\mathbf{M} \mapsto \mathbf{MT}$. When $\mathbf{T}$ is a permutation matrix, then these translations are bijective maps, and the left translation (respectively right) translation corresponds to rows (respectively columns) permutations.

For some fixed $\mathbf{T} \in \mathcal{M}_n$, let us consider the left translation $\mathbf{M} \mapsto \mathbf{TM}$, where $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$. By linearity, we have

$$\mathbb{E}[\mathbf{TM}] = \mathbf{T}\mathbb{E}[\mathbf{M}] = \mathbf{T}\frac{1}{n}\mathbf{1}_n = \frac{1}{n}\mathbf{1}_n$$

where $\mathbf{1}_n$ is the $n \times n$ matrix full of ones. Thus, the left translation by $\mathbf{T}$ leaves the mean invariant.

Proof of theorem 1.2. First of all, the case $n = 1$ is trivial and one can assume that $n > 1$ in the rest of the proof. A probability distribution μ on $\mathcal{M}_n$ is invariant by the left translation $\mathbf{M} \mapsto \mathbf{PM}$ for every permutation matrix $\mathbf{P}$ of size $n \times n$ if and only if μ is row exchangeable. Similarly, μ is invariant by the right translation $\mathbf{M} \mapsto \mathbf{MP}$ for every permutation matrix $\mathbf{P}$ of size $n \times n$ if and only μ is column exchangeable. Theorem 1.1 gives then the invariance of $\mathcal{U}(\mathcal{M}_n)$ by left and right translations with respect to permutation matrices. Notice however that as a probability distribution over $\mathbb{R}^{n^2}$, $\mathcal{U}(\mathcal{M}_n)$ is not exchangeable. The permutation of rows and columns correspond to a proper subset of the group of permutations of the n^2 entries.

Conversely, let us assume that the law $\mathcal{U}(\mathcal{M}_n)$ is invariant by the left translation $\mathbf{M} \mapsto \mathbf{TM}$ for some $\mathbf{T} \in \mathcal{M}_n$. If $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$, and since the components of the first column $\mathbf{M}_{\cdot,1}$ of $\mathbf{M}$ are i.i.d. we have

$$\begin{aligned}
\mathrm{Var}((\mathbf{TM})_{1,1}) &= \mathrm{Var}\left(\sum_{k=1}^n \mathbf{T}_{1,k}\mathbf{M}_{k,1}\right) \\
&= \sum_{k=1}^n (\mathbf{T}_{1,k})^2 \mathrm{Var}(\mathbf{M}_{k,1}) \\
&= \mathrm{Var}(\mathbf{M}_{1,1}) \sum_{k=1}^n (\mathbf{T}_{1,k})^2.
\end{aligned}$$

The invariance hypothesis implies in particular that $\mathrm{Var}(\mathbf{M}_{1,1}) = \mathrm{Var}((\mathbf{TM})_{1,1})$. Since $\mathrm{Var}(\mathbf{M}_{1,1}) = (n-1)/(n^2(n+1)) > 0$, we get $1 = \sum_{k=1}^n (\mathbf{T}_{1,k})^2$. Now, $\mathbf{T}$ is Markov and thus $\sum_{k=1}^n \mathbf{T}_{1,k} = 1$, which gives

$$\sum_{k=1}^n (\mathbf{T}_{1,k} - (\mathbf{T}_{1,k})^2) = 0.$$

Since $\mathbf{T}$ is Markov, its entries are in $[0,1]$ and hence $\mathbf{T}_{1,k} \in \{0,1\}$ for every $1 \leqslant k \leqslant n$. The condition $\sum_{k=1}^n \mathbf{T}_{1,k} = 1$ gives then that the first line of $\mathbf{T}$ is an element of the canonical basis of $\mathbb{R}^n$. The same argument used for $(\mathbf{TM})_{k,1}$ for every $1 \leqslant k \leqslant n$ shows that every line of $\mathbf{T}$ is an element of the canonical basis, and thus $\mathbf{T}$ is a binary matrix with exactly a unique 1 on each line. Since $\mathbf{TM} \sim \mathcal{U}(\mathcal{M}_n)$, it has independent rows, and thus the position of the 1's on the rows of $\mathbf{T}$ are pairwise different, which means that $\mathbf{T}$ is a permutation matrix as expected.

Let us consider now the case where the law $\mathcal{U}(\mathcal{M}_n)$ is invariant by the right translation $\mathbf{M} \mapsto \mathbf{MT}$ for some $\mathbf{T} \in \mathcal{M}_n$. If $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$, we can first take a look at the mean. Namely, $\mathbb{E}[\mathbf{MT}] = \mathbb{E}[\mathbf{M}]\mathbf{T} = \frac{1}{n}\mathbf{S}$ where $\mathbf{S}$ is defined by

$$\mathbf{S}_{i,j} = \sum_{k=1}^n \mathbf{T}_{k,j}$$

for every $1 \leqslant i,j \leqslant n$. Now, the invariance hypothesis gives on the other hand

$$\mathbb{E}[\mathbf{MT}] = \mathbb{E}[\mathbf{M}] = \frac{1}{n}\mathbf{1}_n$$

and thus $\mathbf{S} = \mathbf{1}_n$, which means that $\mathbf{T}$ is bistochastic, i.e. both $\mathbf{T}$ and $\mathbf{T}^\top$ are Markov. The invariance hypothesis implies also that

$$\mathrm{Var}((\mathbf{MT})_{1,1}) = \mathrm{Var}(\mathbf{M}_{1,1}) = (n-1)/(n^2(n+1)).$$

But since the first line $\mathbf{M}_{1,\cdot}$ of $\mathbf{M}$ is $\mathcal{D}_n(1,\ldots,1)$ distributed,

$$\begin{aligned}\mathrm{Var}((\mathbf{MT})_{1,1}) &= \sum_{1\leqslant i,j\leqslant n} \mathbf{T}_{i,1}\mathbf{T}_{j,1}\mathrm{Cov}(\mathbf{M}_{1,i};\mathbf{M}_{1,j}) \\ &= \frac{n-1}{n^2(1+n)}\sum_{i=1}^n (\mathbf{T}_{i,1})^2 - \frac{2}{n^2(n+1)}\sum_{1\leqslant i<j\leqslant n} \mathbf{T}_{i,1}\mathbf{T}_{j,1}.\end{aligned}$$

Since $\mathbf{T}$ is bistochastic, we have $1 = \sum_{i=1}^n \mathbf{T}_{i,1}$ and thus

$$(n-1)\sum_{i=1}^n (\mathbf{T}_{i,1} - (\mathbf{T}_{i,1})^2) = -2\sum_{1\leqslant i<j\leqslant n} \mathbf{T}_{i,1}\mathbf{T}_{j,1}.$$

The terms of the left and right hand side have opposite signs, which gives that $\mathbf{T}_{i,1} \in \{0,1\}$ for every $1 \leqslant i \leqslant n$. The same method used for $(\mathbf{MT})_{1,k}$ for every $1 \leqslant k \leqslant n$ shows that $\mathbf{T}$ is a binary matrix. Since $\mathbf{T}$ is bistochastic, it follows that $\mathbf{T}$ is actually a permutation matrix, as expected. □

The set of $n \times n$ permutation matrices is a discrete subgroup of the orthogonal group of $\mathbb{R}^n$, isomorphic to the symmetric group Σ_n. The group of permutation matrices plays for the Dirichlet Markov Ensemble the role played by the orthogonal group for Dyson's GOE or COE, and the role played by the unitary group for Dyson's GUE or CUE. In some sense, we replaced an L^2 Gaussian structure by an L^1 Dirichlet structure while maintaining the permutation invariance.

A very natural question is to ask about the existence of a convolution idempotent probability distribution on the compact semi-group $\mathcal{M}_n$. Recall that a probability distribution μ on a semi-group $\mathfrak{S}$ is idempotent if and only if $\mu * \mu = \mu$. Here the convolution $\mu * \nu$ of two probability distributions μ and ν on $\mathfrak{S}$ is defined, for every bounded continuous $f : \mathfrak{S} \to \mathbb{R}$, by

$$\int_{\mathfrak{S}} f(s) d(\mu * \nu)(s) = \int_{\mathfrak{S}} \left(\int_{\mathfrak{S}} f(s_l s_r)\, d\mu(s_l) \right) d\nu(s_r).$$

Actually, the structure of compact semi-groups and their idempotent measures was deeply investigated in the 1960's, see [31, pages 158-160] for a historical account. In particular, one can find in [31, Lemma 3 page 141] the following result.

Lemma 2.1. *Let μ be a regular probability distribution over a compact Hausdorff semi-group $\mathfrak{S}$ such that the support of μ generates $\mathfrak{S}$. Then the mass of the convolution sequence μ^{*n} concentrates on the kernel $K(\mathfrak{S})$ of $\mathfrak{S}$. More precisely, for every open set O containing K and every $\varepsilon > 0$, there exists a positive integer n_ε such that $\mu^{*n}(O) > 1 - \varepsilon$ for every $n \geqslant n_\varepsilon$.*

Here μ^{*n} denotes the convolution product $\mu * \cdots * \mu$ of n copies of μ. If μ^{*n} tends to μ as $n \to \infty$ then μ is convolution idempotent, that is $\mu * \mu = \mu$. The kernel $K(\mathfrak{S})$ of $\mathfrak{S}$ is the sub-semi-group of $\mathfrak{S}$ obtained by taking the intersection of the family of two sided ideals of $\mathfrak{S}$, see [31, Theorem 1 page 140]. A direct consequence of lemma 2.1 is the absence of a translation invariant probability measure μ on $\mathfrak{S}$ with full support such that the kernel of $\mathfrak{S}$ is a μ-proper sub-semi-group of $\mathfrak{S}$. By μ-proper sub-semi-group here we mean that its μ-measure is < 1. This result can be easily understood intuitively since the translation associated to a non-invertible element of $\mathfrak{S}$ gives a strict contraction of the support.

Proof of theorem 1.3. The kernel of the semi-group $\mathcal{M}_n$ is constituted by the $n \times n$ Markov matrices with equal rows, which are the $n \times n$ idempotent Markov matrices (i.e. $\mathbf{M}^2 = \mathbf{M}$). The reader may find more details in [31, page 146]. Since the kernel of $\mathcal{M}_n$ is a $\mathcal{U}(\mathcal{M}_n)$-proper sub-semi-group of $\mathcal{M}_n$, lemma 2.1 implies the absence of any convolution idempotent probability distribution on $\mathcal{M}_n$, absolutely continuous with respect to $\mathcal{U}(\mathcal{M}_n)$ and with full support. The proof is finished by noticing that if a probability distribution on $\mathcal{M}_n$ is invariant by every left (or right) translation, then it is convolution idempotent. Notice by the way that the Wedderburn matrix $\frac{1}{n}\mathbf{1}_n$ belongs to the kernel of $\mathcal{M}_n$, and also that this kernel is equal to $\{\lim_{k\to\infty}\mathbf{M}^k; \mathbf{M}\in\mathcal{A}_n\}$ where $\mathcal{A}_n$ is the collection of irreducible aperiodic elements of $\mathcal{M}_n$. The reader may find in [31, Chapter 5] the structure of non fully supported idempotent probability distributions on compact semi-groups and in particular on $\mathcal{M}_n$. $\square$

3 Proofs of theorem 1.4

The following theorem can be found for instance in [6, th. 3.6 p. 43].

Theorem 3.1 (Singular values of large dimensional non-centered random arrays). *Let $(X_{i,j})_{1\leqslant i,j<\infty}$ be an infinite array of i.i.d. real random variables with mean m and variance $\sigma^2\in(0,\infty)$. If $\mathbf{X}=(X_{i,j})_{1\leqslant i,j\leqslant n}$, then almost surely, the ESD of $n^{-1}\mathbf{X}\mathbf{X}^\top$ tends, as $n\to\infty$, to the Marchenko-Pastur distribution* (3).

Theorem 3.1 implies in particular that almost surely

$$\frac{1}{n}\,\mathrm{Card}\{1\leqslant k\leqslant n \text{ such that } \lambda_k(n^{-1}\mathbf{X}\mathbf{X}^\top)>4\sigma^2\}\underset{n\to\infty}{\longrightarrow}0.$$

However, it can be shown (see [6]) that if $m\neq 0$, then almost surely,

$$\lambda_1(n^{-1}\mathbf{X}\mathbf{X}^\top)\underset{n\to\infty}{\longrightarrow}+\infty.$$

The following lemma is a consequence of [7, le. 2] (see also [6, le. 5.13 p. 102]).

Lemma 3.2 (Uniform law of large numbers). *If $(X_{i,j})_{1\leqslant i,j<\infty}$ is an infinite array of i.i.d. random variables of mean m, then by denoting $S_{i,n}=\sum_{j=1}^n X_{i,j}$,*

$$\max_{1\leqslant i\leqslant n}\left|\frac{S_{i,n}}{n}-m\right|\xrightarrow[n\to\infty]{a.s.}0$$

and in the case where $m\neq 0$, we have also

$$\max_{1\leqslant i\leqslant n}\left|\frac{n}{S_{i,n}}-\frac{1}{m}\right|\xrightarrow[n\to\infty]{a.s.}0.$$

The following lemma is a consequence of the Courant-Fischer variational formulas for singular values, see [25]. Also, we leave the proof to the reader.

Lemma 3.3 (Singular values of diagonal multiplicative perturbations). *For every $n \times n$ matrix $\mathbf{A}$, every $n \times n$ diagonal matrix $\mathbf{D}$, and every $1 \leqslant k \leqslant n$,*

$$s_n(\mathbf{D})^2 s_k(\mathbf{A}) \leqslant s_k(\mathbf{D}\mathbf{A}) \leqslant s_1(\mathbf{D})^2 s_k(\mathbf{A}).$$

We are now able to prove theorem 1.4.

Proof of theorem 1.4. We use the method of Aubrun [2], by replacing the unit $\|\cdot\|_1$-ball by the portion of the unit $\|\cdot\|_1$-sphere with non-negative coordinates. Let $\mathbf{M}$ be as in theorem 1.4. We have $\mathbf{M} = \mathbf{D}\mathbf{E}$ where $\mathbf{E} = (X_{i,j})_{1\leqslant i,j\leqslant n}$ and $\mathbf{D}$ is the $n \times n$ diagonal matrix given for every $1 \leqslant i \leqslant n$ by

$$\mathbf{D}_{i,i} = \frac{1}{\sum_{j=1}^n X_{i,j}}.$$

The fact that $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$ follows immediately from theorem 1.1 combined with the construction of the Dirichlet distribution $\mathcal{D}_n(1,\ldots,1)$ from i.i.d. exponential random variables. It remains to prove the convergence of the ESD of $n\mathbf{M}\mathbf{M}^\top$ as $n \to \infty$ to the Marchenko-Pastur distribution. We have

$$n\mathbf{M}\mathbf{M}^\top = (n\mathbf{D}n^{-1/2}\mathbf{E})(n\mathbf{D}n^{-1/2}\mathbf{E})^\top.$$

By lemma 3.3, we get for every $1 \leqslant k \leqslant n$,

$$s_n(n\mathbf{D})^2\lambda_k(n^{-1}\mathbf{E}\mathbf{E}^\top) \leqslant \lambda_k(n\mathbf{M}\mathbf{M}^\top) \leqslant s_1(n\mathbf{D})^2\lambda_k(n^{-1}\mathbf{E}\mathbf{E}^\top).$$

Now, lemma 3.2 gives that almost surely, $s_1(n\mathbf{D}) \to 1$ and $s_n(n\mathbf{D}) \to 1$ as $n \to \infty$. By theorem 3.1, almost surely, the ESD of $n^{-1}\mathbf{E}\mathbf{E}^\top$ converges as $n \to \infty$ to the Marchenko-Pastur distribution (3) with $\sigma^2 = 1$. It follows that almost surely, the ESD of $n\mathbf{M}\mathbf{M}^\top$ tends as $n \to \infty$ to the same distribution. ☐

There is no equivalent of lemma 3.3 for the eigenvalues instead of the singular values, and thus the method used to prove theorem 1.4 fails for conjecture 1.5. Notice that by lemma 3.2 used with the exponential distribution of mean $m = 1$,

$$\|n\mathbf{D} - \mathbf{I}\|_2 = \max_{1\leqslant i\leqslant n}\left|\frac{n}{\sum_{j=1}^n X_{i,j}} - 1\right| \underset{n\to\infty}{\longrightarrow} 0 \quad \text{a.s.}$$

where $\|\mathbf{A}\|_2 = s_1(\mathbf{A}) = \max_{\|x\|_2=1}\|\mathbf{A}x\|_2$ is the Euclidean operator norm of $\mathbf{A}$. If $\mathbf{A}$ is diagonal, then we simply have $\|\mathbf{A}\|_2 = s_1(\mathbf{A}) = \max_{1\leqslant k\leqslant n}|\mathbf{A}_{k,k}|$, and when $\mathbf{A}$ is diagonal and invertible, $\|\mathbf{A}^{-1}\|_2^{-1} = s_n(\mathbf{A}) = \min_{1\leqslant k\leqslant n}|\mathbf{A}_{k,k}|$. Now, by the circular law theorem for non-central random matrices [14], we get that almost surely, the ESD of $n^{-1/2}\mathbf{E}$ converges, as $n \to \infty$, to the uniform distribution on $D(0,1)$. It is then natural to decompose $\sqrt{n}\mathbf{M}$ as

$$\sqrt{n}\mathbf{M} = n\mathbf{D}n^{-1/2}\mathbf{E} = (n\mathbf{D} - \mathbf{I})n^{-1/2}\mathbf{E} + n^{-1/2}\mathbf{E}.$$

Unfortunately, since $m = 1 \neq 0$, we have almost surely (see [14])

$$\left\|n^{-1/2}\mathbf{E}\right\|_2 = s_1(n^{-1/2}\mathbf{E}) \underset{n\to\infty}{\longrightarrow} +\infty.$$

This suggests that $\sqrt{n}\mathbf{M}$ cannot be seen as a perturbation of $n^{-1/2}\mathbf{E}$ with a matrix of small norm. Actually, even if it was the case, the relation between the two spectra is unknown since $\mathbf{E}$ is not normal. One can think about using logarithmic potentials to circumvent the problem. The strength of the logarithmic potential approach is that it allows to study the asymptotic behavior of the ESD (i.e. eigenvalues) of non-normal matrices via the singular values of a family of matrices indexed by $z \in \mathbb{C}$. The details are given in [28] and [14] for instance. The logarithmic potential of the ESD of $\sqrt{n}\,\mathbf{M}$ at point z is

$$\begin{aligned} U_n(z) &= -\frac{1}{n}\log\left|\det(\sqrt{n}\mathbf{M} - z\mathbf{I})\right| \\ &= -\frac{1}{n}\log\left|\det(n\mathbf{D})\right| - \frac{1}{n}\log\left|\det(n^{-1/2}\mathbf{E} - z(n\mathbf{D})^{-1})\right|. \end{aligned}$$

Now, by lemma 3.2,

$$\frac{1}{n}\log\left|\det(n\mathbf{D})\right| \underset{n\to\infty}{\longrightarrow} 0 \quad \text{a.s.}$$

By the circular law theorem for non-central random matrices (see [14]) together with the lower envelope theorem (see [28] and [14]), almost surely, for quasi-every $z \in \mathbb{C}$, the quantity

$$\liminf_{n\to\infty} -\frac{1}{n}\log\left|\det(n^{-1/2}\mathbf{E} - z\mathbf{I})\right|$$

is equal to the logarithmic potential at point z of the uniform distribution on the unit disc $D(0,1)$. It is thus enough to show that almost surely, for every $z \in \mathbb{C}$,

$$\frac{1}{n}\log\left|\det(n^{-1/2}\mathbf{E} - z(n\mathbf{D})^{-1})\right| - \frac{1}{n}\log\left|\det(n^{-1/2}\mathbf{E} - z\mathbf{I})\right| \underset{n\to\infty}{\longrightarrow} 0.$$

Unfortunately, we ignore how to prove that, mainly because we ignore how to bound the extremal singular values of an additive diagonal perturbation of a non-normal matrix.

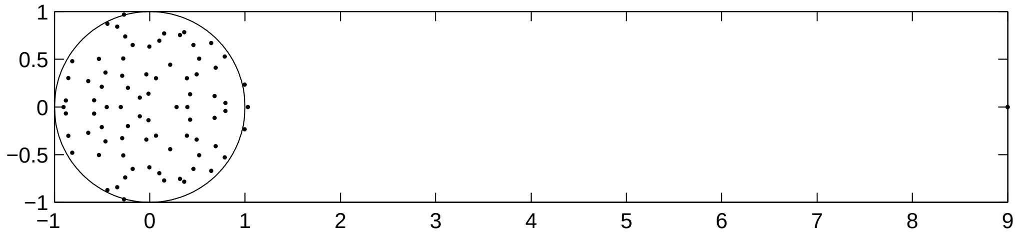

Figure 1: Plot of the spectrum of a single realization of $\sqrt{n}\mathbf{M}$ where $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$ with $n = 81$. We see one isolated eigenvalue $\lambda_1(\sqrt{n}\mathbf{M}) = \sqrt{n} = 9$ while the rest of the spectrum remains near the unit disc and seems uniformly distributed, in accordance with conjecture 1.5.

Acknowledgements. Part of this work was done during two visits to Laboratoire Jean Dieudonné in Nice, France. The author would like to thank Pierre Del Moral and Persi Diaconis for their kind hospitality there. Many thanks to Zhidong Bai, Franck Barthe, Włodzimierz Bryc, Mireille Capitaine, Delphine Féral, Michel Ledoux, and Gérard Letac for exchanging some ideas on the subject. This work benefited from many stimulating discussions with Neil O'Connell when he visited the Institut de Mathématiques de Toulouse.

References

[1] A. L. Andrew, *Eigenvalues and singular values of certain random matrices*, J. Comput. Appl. Math. **30** (1990), no. 2, 165–171.

[2] G. Aubrun, *Random points in the unit ball of l_p^n*, Positivity **10** (2006), no. 4, 755–759.

[3] ———, *Sampling convex bodies: a random matrix approach*, Proc. Amer. Math. Soc. **135** (2007), no. 5, 1293–1303 (electronic).

[4] Z. D. Bai, *Circular law*, Ann. Probab. **25** (1997), no. 1, 494–529.

[5] ———, *Methodologies in spectral analysis of large-dimensional random matrices, a review*, Statist. Sinica **9** (1999), no. 3, 611–677, With comments by G. J. Rodgers and J. W. Silverstein; and a rejoinder by the author.

[6] Z. D. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Mathematics Monograph Series 2, Science Press, Beijing, 2006.

[7] Z. D. Bai and Y. Q. Yin, *Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix*, Ann. Probab. **21** (1993), no. 3, 1275–1294.

[8] F. Barthe, O. Guédon, Sh. Mendelson, and A. Naor, *A probabilistic approach to the geometry of the ℓ_n^p ball*, The Annals of Probability **33** (2005), no. 2, 480–513, `arXiv.org:math.PR/0503650`.

[9] I. Bengtsson, Å. Ericsson, M. Kuś, W. Tadej, and K. Życzkowski, *Birkhoff's polytope and unistochastic matrices, $N = 3$ and $N = 4$*, Comm. Math. Phys. **259** (2005), no. 2, 307–324.

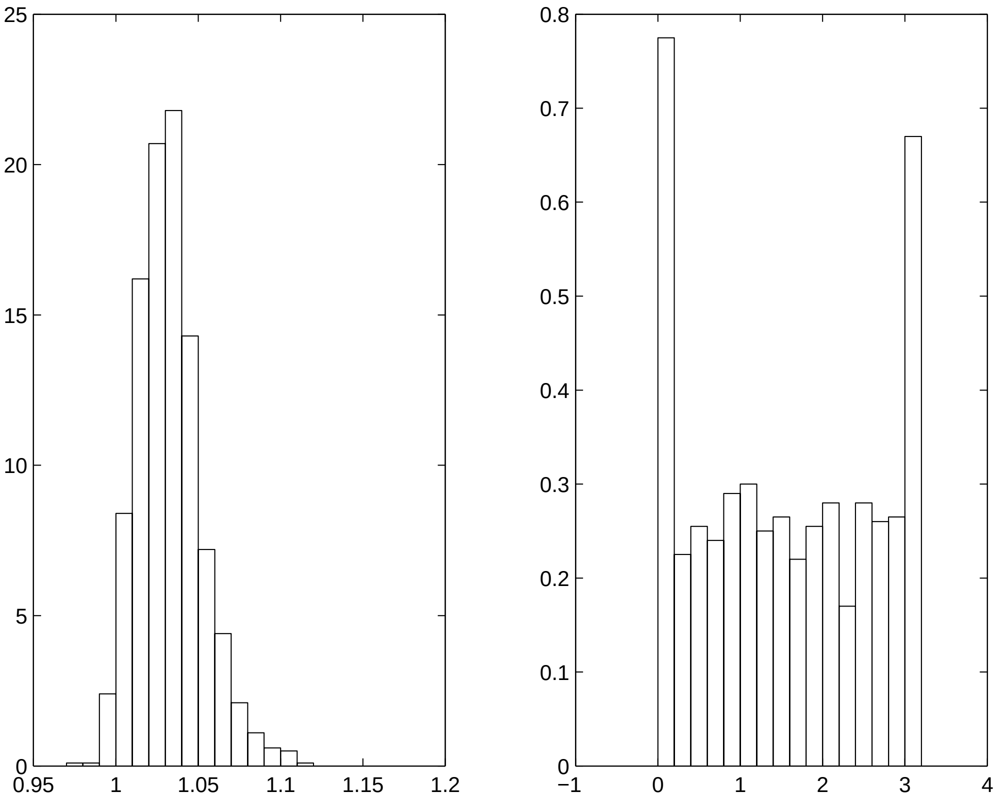

Figure 2: Here 1000 i.i.d. realizations of $\sqrt{n}\mathbf{M}$ where simulated where $\mathbf{M} \sim \mathcal{U}(\mathcal{M}_n)$ with $n = 300$. The first plot is the histogram of $|\lambda_2(\sqrt{n}\mathbf{M})|$, i.e. the module of the sub-dominant eigenvalue $\lambda_2(\sqrt{n}\mathbf{M})$. The second plot is the histogram of $|\mathfrak{Arg}(\lambda_2(\sqrt{n}\mathbf{M}))|$, i.e. the absolute value of the argument of the sub-dominant eigenvalue $\lambda_2(\sqrt{n}\mathbf{M})$. Recall that the spectrum is symmetric with respect to the real axis since the matrices are real. The first histogram suggests that the asymptotic fluctuations of the module is probably, up to scaling, a Gumbel type distribution as for the Ginibre Ensemble [30, 29]. The second histogram suggests that the argument is uniformly distributed outside the real axis.

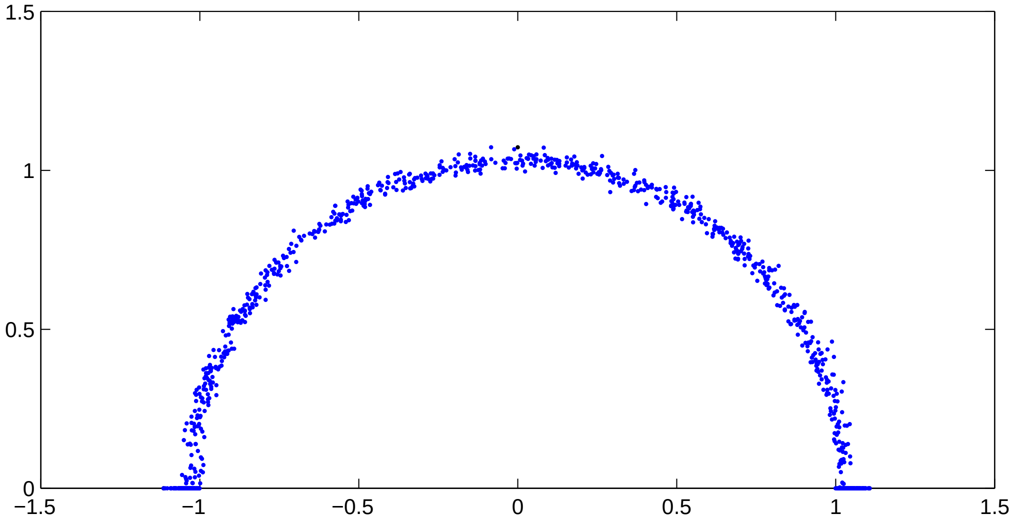

Figure 3: Here we reused the sample used for figure 2. The graphic is a plot of the 1000 i.i.d. realizations of the sub-dominant eigenvalue $\lambda_2(\sqrt{n}\mathbf{M})$. It suggests that apart the behavior near ± 1, the sub-dominant eigenvalue lies uniformly near the unit circle. Since we deal with real matrices, the spectrum is symmetric with respect to the real axis, and we plotted $(\mathfrak{Re}(\lambda_2), |\mathfrak{Im}(\lambda_2)|)$ in the complex plane.

[10] G. Berkolaiko, *Spectral gap of doubly stochastic matrices generated from equidistributed unitary matrices*, J. Phys. A **34** (2001), no. 22, L319–L326.

[11] Ph. Biane, *On the free convolution with a semi-circular distribution*, Indiana Univ. Math. J. **46** (1997), no. 3, 705–718.

[12] A. Bose and A. Sen, *Spectral norm of random large dimensional noncentral Toeplitz and Hankel matrices*, Electron. Comm. Probab. **12** (2007), 29–35 (electronic).

[13] W. Bryc, A. Dembo, and T. Jiang, *Spectral measure of large random Hankel, Markov and Toeplitz matrices*, Ann. Probab. **34** (2006), no. 1, 1–38.

[14] D. Chafaï, *Circular law for non-central random matrices*, preprint `arXiv:0709.0036 [math.PR]`, 2007.

[15] F. Comets, S. Popov, G. Schütz, and M. Vachkovskaia, *Billiards in a general domain with random reflection*, preprint `arXiv:math/0612799v2 [math.PR]`, 2007.

[16] P. Diaconis and B. Efron, *Testing for independence in a two-way table: new interpretations of the chi-square statistic*, Ann. Statist. **13** (1985), no. 3, 845–913, With discussions and with a reply by the authors.

[17] ______, *Probabilistic-geometric theorems arising from the analysis of contingency tables*, Contributions to the theory and application of statistics, Academic Press, Boston, MA, 1987, pp. 103–125.

[18] P. Diaconis and A. Gangolli, *Rectangular arrays with fixed margins*, Discrete probability and algorithms (Minneapolis, MN, 1993), IMA Vol. Math. Appl., vol. 72, Springer, New York, 1995, pp. 15–41.

[19] K. Falconer, *Fractal geometry*, second ed., John Wiley & Sons Inc., Hoboken, NJ, 2003, Mathematical foundations and applications.

[20] G. Goldberg and M. Neumann, *Distribution of subdominant eigenvalues of matrices with random rows*, SIAM J. Matrix Anal. Appl. **24** (2003), no. 3, 747–761 (electronic).

[21] G. Goldberg, P. Okunev, M. Neumann, and H. Schneider, *Distribution of subdominant eigenvalues of random matrices*, Methodol. Comput. Appl. Probab. **2** (2000), no. 2, 137–151.

[22] F. Götze and A. Tikhomirov, *The Circular Law for Random Matrices*, preprint `arXiv:math/07093995v1 [math.PR]`, 2007.

[23] ———, *On the Circular Law*, preprint `arXiv:math/0702386v1 [math.PR]`, 2007.

[24] F. Hiai and D. Petz, *The semicircle law, free random variables and entropy*, Mathematical Surveys and Monographs, vol. 77, American Mathematical Society, Providence, RI, 2000.

[25] R. A. Horn and Ch. R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1990, Corrected reprint of the 1985 original.

[26] M. L. Mehta, *Random matrices*, third ed., Pure and Applied Mathematics (Amsterdam), vol. 142, Elsevier/Academic Press, Amsterdam, 2004.

[27] Sh. Mendelson and A. Pajor, *On singular values of matrices with independent rows*, Bernoulli **12** (2006), no. 5, 761–773.

[28] G. Pan and W. Zhou, *Circular law, extreme singular values and potential theory*, `arXiv:0705.3773v1 [math.PR]`, 2007.

[29] B. Rider, *A limit theorem at the edge of a non-Hermitian random matrix ensemble*, J. Phys. A **36** (2003), no. 12, 3401–3409, Random matrix theory.

[30] ———, *Order statistics and Ginibre's ensembles*, J. Statist. Phys. **114** (2004), no. 3-4, 1139–1148.

[31] M. Rosenblatt, *Markov processes. Structure and asymptotic behavior*, Springer-Verlag, New York, 1971, Die Grundlehren der mathematischen Wissenschaften, Band 184.

[32] G. Schechtman and J. Zinn, *On the volume of the intersection of two ℓ_n^p balls*, Proc. Amer. Math. Soc. **110** (1990), 217–224.

[33] A. Sen and A. Bose, *Another look at the moment method for large dimensional random matrices*, preprint `http://www.isical.ac.in/~statmath/html/publication/volume25.pdf`, May 2007.

[34] J. Sethuraman, *A constructive definition of Dirichlet priors*, Statist. Sinica **4** (1994), no. 2, 639–650.

[35] J. W. Silverstein, *The spectral radii and norms of large-dimensional non-central random matrices*, Comm. Statist. Stochastic Models **10** (1994), no. 3, 525–532.

[36] G. Tanner, *Unitary-stochastic matrix ensembles and spectral statistics*, J. Phys. A **34** (2001), no. 41, 8485–8500.

[37] T. Tao and V. Vu, *Random Matrices: The circular law*, preprint `arXiv:0708.2895v1 [math.PR]`, 2007.

[38] L. N. Trefethen and M. Embree, *Spectra and pseudospectra*, Princeton University Press, Princeton, NJ, 2005, The behavior of nonnormal matrices and operators.

Djalil Chafaï **e-mail address:** `chafai[@]math.ups-tlse.fr`

Institut de Mathématiques de Toulouse (UMR CNRS 5219)
Université Paul Sabatier Toulouse III
118 route de Narbonne, F-31062 Toulouse Cedex 9, France.

Unité de Physiologie et Toxicologie Expérimentales (UMR INRA 181)
École Nationale Vétérinaire de Toulouse
23 Chemin des Capelles, B.P. 87614, F-31076 Toulouse Cedex 3, France.