J Fourier Anal Appl (2011) 17:733–753
DOI 10.1007/s00041-010-9156-z

Limiting Empirical Singular Value Distribution of Restrictions of Discrete Fourier Transform Matrices

Brendan Farrell

Received: 17 March 2010 / Revised: 11 October 2010 / Published online: 16 November 2010

Abstract We determine the limiting empirical singular value distribution for discrete Fourier transform (DFT) matrices when a random set of columns and rows is removed.

Keywords Singular values · Restrictions of Fourier matrices · Limiting Distribution · Restrictions of Unitary Matrices

Mathematics Subject Classification (2010) Primary 48B06 · 46B09 · 60B20 · Secondary 42A61

1 Introduction

An $n \times n$ Hermitian matrix A determines a distribution on the real line by

$$f_A(x) = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i}(x),$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A. Wigner was the first to determine the limiting behavior of such a distribution when the matrix A is random [18]. He initially considered symmetric matrices with 0's on the diagonal and independent plus or minus 1's in the upper-triangle and showed that when scaled by $\frac{1}{\sqrt{n}}$ the empirical

Communicated by Roman Vershynin.

B. Farrell (✉)
Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, Arcisstr. 21, 80333 München, Germany
e-mail: farrell@tum.de

distribution converges in probability to the *Semicircular Law*

$$f_W(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2} & \text{when } |x| \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Wigner later proved that the same limiting distribution holds for symmetric random variables with finite moments [19].

The second classical type of random matrix is the Wishart matrix [20]. Let $H \in \mathbb{C}^{m\times n}$ have independent Gaussian entries with variance $\frac{1}{n}$. Then HH^* is called a Wishart matrix. Marčenko and Pastur showed that the empirical distribution of HH^* converges in probability to

$$f_{HH^*}(x) = \max(0, 1-\rho)\delta(x) + \frac{\sqrt{(x-c_-)(c_+-x)}}{2\pi x} \cdot I_{[c,c_+]},$$

where $c_\pm = (1 \pm \sqrt{\frac{1}{\rho}})^2$ and ρ is the limiting ratio of n to m [10]. Independently, Silverstein and Grenander used a similar technique and proved almost sure convergence [9].

This paper applies the approach of Marčenko and Pastur to a question originating in geometric functional analysis: we address the singular values of random submatrices of discrete Fourier transform (DFT) matrices when a random set of columns and rows is removed and determine their limiting empirical singular value distribution. The $n \times n$ DFT matrix has entries

$$F_{jk} = \frac{1}{\sqrt{n}} e^{-2\pi i(j-1)(k-1)/n}.$$

Let T and Ω be subsets of $\{1, \dots, n\}$. We define $F_{\Omega T}$ to be the matrix obtained from F_n by removing rows with indices not in Ω and columns with indices not in T.

We show that when each index is included in Ω independently with probability $(1-q)$ and in T independently with probability $(1-p)$, then the limiting empirical distribution of $F_{\Omega T}F_{\Omega T}^*$ depends only on the parameters p and q and converges almost surely to

$$f_{p,q}(x) = \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)(1-\max(p,q))} \cdot I_{(r_-,r_+)}(x) + \frac{\max(0, 1-(p+q))}{1-\max(p,q)} \cdot \delta(x-1),$$

where

$$r_- = \left(\sqrt{p(1-q)} - \sqrt{q(1-p)}\right)^2$$

and

$$r_+ = \left(\sqrt{p(1-q)} + \sqrt{q(1-p)}\right)^2.$$

This is formally stated as Theorem 3.1. As a corollary, we show in Sect. 3.1 that the same distribution holds when random unitary matrices with Haar distribution take the role of the DFT matrices above.

The interest in the spectrum of these matrices from the perspective of geometric functional analysis is in the largest and smallest eigenvalues of $F_{\Omega T}F^*_{\Omega T}$, initially perhaps asymptotically, but ideally for finite dimension. These eigenvalues are related to discrete uncertainty principles, as well as random projections and embeddings. More discussion of their significance is given following Theorem 3.1. The first works on the extremal eigenvalues in the Wishart case were [3, 6, 21], and in the Wigner case it was [2]. It is important to note that the limiting empirical distributions were determined in these two cases before the behavior of the extremal eigenvalues was proved and were essential in that effort. We hope that the distribution presented here leads to similar developments.

1.1 Notation

To make notation easiest, a single subscript will denote the dimension of a square matrix, while a double index will refer to an entry of the matrix. Thus F_n will denote the $n \times n$-dimensional DFT matrix, and

$$F_{jk} = \frac{1}{\sqrt{n}} e^{-2\pi i (j-1)(k-1)/n}$$

will denote its entry at index (j, k). When we want to make the original dimension apparent, we write $F_{\Omega_n T_n}$. We find it helpful to also work with matrices with rows and columns set to zero rather than removed. For clarity we make the following definitions.

Definition 1.1 A square matrix is called a *diagonal projection matrix* if its off-diagonal entries are all zero and its diagonal entries are zero or one.

Definition 1.2 A random diagonal projection matrix will be called a *Bernoulli* diagonal projection matrix if the diagonal entries are independent and equal to 1 with probability $1 - p$ and equal to 0 with probability p.

The matrices P_n and Q_n will denote independent Bernoulli diagonal projection matrices. Asymptotically, P_n and Q_n randomly "erase" the percentage p and q respectively of a vector. For a matrix A, A^* denotes the conjugate transpose of A. We use tr to denote the normalized trace.

Note 1.3 Throughout this paper we take the square root of a complex number to be uniquely defined by having argument in $[0, \pi)$. The reader will see that this is justified.

Note 1.4 When either p or q is 0 or 1, the corresponding matrix is trivial. For the convergence of several sums in later proofs, we assume that $p, q \in (0, 1)$.

2 The Stieltjes and η Transforms

Our main tool is the Stieltjes transform, which is only defined for real random variables. Thus, we will determine the limiting eigenvalue distribution of

$$P_n F_n Q_n F_n^* P_n,$$

which of course is real and contained in [0, 1].

The Stieltjes transform of a real random variable X with distribution function $F_X(x)$ is a function $m_X : \mathbb{C}^+ \to \mathbb{C}$ defined by

$$m_X(z) = \mathbb{E}_X\left[\frac{1}{X-z}\right].$$

If F_X is continuous at x, then $f_X(x)$ can be recovered by the Stieltjes inversion formula [1]

$$f_X(x) = \frac{1}{\pi}\lim_{\omega\to 0} \Im\, m_X(x+i\omega). \tag{1}$$

We will determine the Stieltjes transforms of $P_n F_n Q_n F_n^* P_n$ by first using the η-transform, which was introduced by Tulino and Verdú in [15]. For a real valued random variable X, the η-transform is also a function $\eta_X : \mathbb{C}^+ \to \mathbb{C}$ defined by

$$\eta_X(z) = \mathbb{E}_X\left[\frac{1}{1+zX}\right].$$

Note that for z in an appropriate region of convergence

$$m_X(z) = -\frac{1}{z}\sum_{k=0}^{\infty}(z)^{-k}\mathbb{E}\big[X^k\big]$$

and

$$\eta_X(z) = \sum_{k=0}^{\infty}(-z)^k\mathbb{E}\big[X^k\big], \tag{2}$$

so that

$$m_X(z) = -\frac{1}{z}\eta_X\left(-\frac{1}{z}\right). \tag{3}$$

In this section we determine the η-transform for the matrices $P_n F_n Q_n F_n^* P_n$, which is given in Proposition 2.4. We require several lemmas en route to this proposition.

Lemma 2.1 *Let $\bar{P}$ be a mean-zero, random diagonal matrix with independent entries in $[-1, 1]$ and of dimension n. Then there exists a constant C_m such that for dimension n and all $1 \le k, l \le n$,*

$$\mathbb{E}\big|\big[F^*\bar{P}F\big]_{k,l}\big|^m \le C_m n^{-m/2}.$$

The constant C_m increases with m.

Proof We look at $(F^*\bar{P}F)_{k,l} = \frac{1}{n}\sum_{j=1}^{n}\bar{P}_{j,j}e^{2\pi i(j-1)(l-k)/n}$. Set $Y_j = \Re\bar{P}_{j,j}e^{2\pi i(j-1)(l-k)/n}$ and $Z_j = \Im\bar{P}_{j,j}e^{2\pi i(j-1)(l-k)/n}$. The Y_j, $j=1,\dots,n$, are independent random variables, and $-1 \le Y_j \le 1$ for all j. The same also holds for the Z_j. Since the $\bar{P}_j$ are mean-zero, $\mathbb{E}\sum_{k=1}^{n}Y_j = 0$ and $\mathbb{E}\sum_{k=1}^{n}Z_j = 0$. We have

$$\begin{aligned}\mathbb{P}\left(\left|\frac{1}{n}\sum_{j=1}^{n}(Y_j+iZ_j)\right| > t\right) &\le \mathbb{P}\left(\left|\frac{1}{n}\sum_{j=1}^{n}Y_j\right| > \frac{t}{\sqrt{2}} \cup \left|\frac{1}{n}\sum_{j=1}^{n}Z_j\right| > \frac{t}{\sqrt{2}}\right)\\ &\le \mathbb{P}\left(\left|\frac{1}{n}\sum_{j=1}^{n}Y_j\right| > \frac{t}{\sqrt{2}}\right) + \mathbb{P}\left(\left|\frac{1}{n}\sum_{j=1}^{n}Z_j\right| > \frac{t}{\sqrt{2}}\right)\\ &\le 4e^{-t^2\frac{n}{4}},\end{aligned}$$

where the last inequality is Hoeffding's inequality.

Lastly,

$$\begin{aligned}\mathbb{E}\left|\frac{1}{n}\sum_{j=1}^{n}(Y_j+iZ_j)\right|^m &= \int_0^\infty ms^{m-1}\mathbb{P}\left(\left|\frac{1}{n}\sum_{j=1}^{n}(Y_j+iZ_j)\right| \ge s\right)ds\\ &\le 4m\int_0^\infty s^{m-1}e^{-\frac{s^2n}{4}}\,ds\\ &= 4m\int_0^\infty 2^{m-1}\left(\frac{t}{n}\right)^{(m-1)/2}e^{-t}(nt)^{-1/2}\,dt\\ &= m2^{m+1}n^{-m/2}\int_0^\infty t^{\frac{m}{2}-1}e^{-t}\,dt\\ &= m2^{m+1}n^{-m/2}\Gamma\left(\frac{m}{2}-1\right).\end{aligned}$$

□

We define the matrix $W_n = P_nF_n$; however, in what follows we will not write the subscript n. We use w_i to denote the ith column of W_n, and define the following quantity

$$\begin{aligned}A_i &= I + zPFQF^*P - zQ_{i,i}w_iw_i^* &(4)\\ &= I + z\sum_{j\neq i}Q_{j,j}w_jw_j^*. &(5)\end{aligned}$$

Lemma 2.2 *For $|z|<1$ the random variable $w_i^*A_i^{-1}w_i$ defined in (5) equals a deterministic constant $D(z)$ independent of the dimension n plus a random part, which we denote $X_{n,i}$. There exists a constant C_z depending on z but independent of i such that for all n and $1\le i\le n$*

$$\mathbb{P}(|X_{n,i}|>\epsilon) \le C_z\epsilon^{-4}n^{-2}.$$

Proof We arbitrarily select an index i and denote it i^*. If $|z|<1$, then for any realization of P_n and Q_n every entry of the following sum converges:

$$\begin{aligned} W^* A_{i^*}^{-1} W &= F^* P\big(I + zPFQF^*P - zQ_{(i^*,i^*)} w_{i^*} w_{i^*}^*\big)^{-1} PF \\ &= F^* P \sum_{k=0}^{\infty} (-z)^k \big(PFQF^*P - Q_{(i^*,i^*)} w_{i^*} w_{i^*}^*\big)^k PF. \end{aligned}$$

Since each entry of these matrices is bounded in absolute value by 1, for $|z|<1$ and for any $\delta>0$ we may choose K such that

$$\left|\left[F^* P \sum_{k=K+1}^{\infty} (-z)^k \big(PFQF^*P\big)^k PF\right]_{(i^*,i^*)}\right| < \delta$$

for any realization of Q, independent of n. For now we just take K to be a large integer. Observe that for the index i^*, the definition of A_{i^*} in (4) requires that the random matrix Q has a deterministic zero at the entry (i^*,i^*). Therefore we set $\tilde{Q}_{i,i} = Q_{i,i} - (1-q)$ for $i \neq i^*$ and $\tilde{Q}_{i^*,i^*} = -(1-q)$. Note that $PF(1-q)IF^*P = (1-q)P$. For a fixed K we now consider

$$\begin{aligned} &F^* P \sum_{k=0}^{K} (-z)^k \big(PFQF^*P - Q_{(i^*,i^*)} w_{i^*} w_{i^*}^*\big)^k PF \\ &= \sum_{k=0}^{K} (-z)^k F^* P\big[PF\tilde{Q}F^*P + (1-q)P\big]^k PF \\ &= \sum_{k=0}^{K} (-z)^k \sum_{j=0}^{k} \binom{k}{j} (1-q)^j F^* P\big[PF\tilde{Q}F^*P\big]^{k-j} P^j PF \\ &= \sum_{k=0}^{K} (-z)^k \sum_{j=0}^{k} \binom{k}{j} (1-q)^j F^* P\big[F\tilde{Q}F^*P\big]^{k-j} F. \end{aligned} \tag{6}$$

Now we center the P matrices. Set $\tilde{P} = P - (1-p)I$. Then

$$\begin{aligned} F^* P\big[F\tilde{Q}F^*P\big]^k F &= F^*\big(\tilde{P} + (1-p)I\big)\big[F\tilde{Q}F^*\big(\tilde{P} + (1-p)I\big)\big]^k F \\ &= \sum_{\alpha} (1-p)^{k+1-|\alpha|} F^* \tilde{P}^{\alpha_0} F\tilde{Q}F^* \tilde{P}^{\alpha_1} F\tilde{Q} \cdots \tilde{P}^{\alpha_k} F \end{aligned} \tag{7}$$

for $\alpha_0,\dots,\alpha_k$ equaling 0 or 1 and $|\alpha|$ equalling the sum of the α_i's. In the case that all α_i's are 0, we recall that $\tilde{Q}_{i^*,i^*} = -(1-q)$, and thus the (i^*,i^*) entry of (7) is deterministic in that case. Thus, the (i^*,i^*) entry of (7) equals a constant independent of n plus a linear combination of the (i^*,i^*) entries of matrices of the form

$$F^* \tilde{P}^{\alpha_0} F\tilde{Q}F^* \tilde{P}^{\alpha_1} F\tilde{Q} \cdots \tilde{P}^{\alpha_k} F$$

with $\alpha_i \neq 0$ for at least one i. When an $\alpha_i = 0$, then we obtain a product of $\tilde{Q}$ matrices, which is then not centered. As earlier, we replace the non-centered random diagonal matrix with a multiple of the identity plus a centered random diagonal matrix. We continue centering each random diagonal matrix in this way such that eventually we only have constant terms, independent of n, and terms of the form

$$F^* \bar{P}^{(1)} F \bar{Q}^{(1)} \cdots F^* \bar{P}^{(k)} F \bar{Q}^{(k)} F^* \bar{P}^{(k+1)} F \tag{8}$$

for some centered matrices $\bar{P}^{(1)}, \dots, \bar{P}^{(k+1)}$ and (except for the (i^*, i^*)-entry) $\bar{Q}^{(1)} \dots \bar{Q}^{(k)}$ and some $1 < k \leq K$. Note that the dimension n plays no role in these expansions. Thus the term in (6) has a deterministic part independent of the dimension and a random part that is a sum of terms of the form (8). The number of such terms depends only on K; call this quantity K_1.

We set $T^{(j)} = F^* \bar{P}^{(j)} F$ and consider a term of the form (8). In the following we use only one index for the diagonal element of the $\bar{Q}$ matrices.

$$\begin{aligned}
&\mathbb{E}\big|\big(T^{(1)} \bar{Q}^{(1)} \cdots T^{(k)} \bar{Q}^{(k)} T^{(k+1)}\big)_{(i^*,i^*)}\big|^4. \\
&\quad = \sum_{l,m,r,s} \big\{\mathbb{E} \bar{Q}^{(1)}_{l_1} \cdots \bar{Q}^{(k)}_{l_k} \bar{Q}^{(1)}_{m_1} \cdots \bar{Q}^{(k)}_{m_k} \bar{Q}^{(1)}_{r_1} \cdots \bar{Q}^{(k)}_{r_k} \bar{Q}^{(1)}_{s_1} \cdots \bar{Q}^{(k)}_{s_k} \\
&\qquad \times \mathbb{E} T^{(1)}_{i^*,l_1} \cdots T^{(k+1)}_{l_k,i^*} T^{(1)}_{i^*,m_1} \cdots T^{(k+1)}_{m_k,i^*} T^{(1)}_{i^*,r_1} \cdots T^{(k+1)}_{r_k,i^*} T^{(1)}_{i^*,s_1} \cdots T^{(k+1)}_{s_k,i^*}\big\}
\end{aligned} \tag{9}$$

By Lemma 2.1, $\mathbb{E}_{P_n,Q_n} |T^{(l)}_{i,j}|^m \leq C_m n^{-m/2}$ for $1 \leq i, j \leq n$, so that

$$\begin{aligned}
&\mathbb{E}\big|T^{(1)}_{i^*,l_1} \cdots T^{(k+1)}_{l_k,i^*} T^{(1)}_{i^*,m_1} \cdots T^{(k+1)}_{m_k,i^*} T^{(1)}_{i^*,r_1} \cdots T^{(k+1)}_{r_k,i^*} T^{(1)}_{i^*,s_1} \cdots T^{(k+1)}_{s_k,i^*}\big| \\
&\quad \leq \big(\mathbb{E}\big|T^{(1)}_{i^*,l_1}\big|^{4(k+1)} \cdots \mathbb{E}\big|T^{(k)}_{s_k,i^*}\big|^{4(k+1)}\big)^{1/(4(k+1))} \\
&\quad \leq C_{4(k+1)} n^{-(4(k+1))/2} \\
&\quad \leq C_{4(K_1+1)} n^{-(2k+2)}.
\end{aligned}$$

We now bound

$$\sum_{l,m,r,s} \mathbb{E} \bar{Q}^{(1)}_{l_1} \cdots \bar{Q}^{(k)}_{l_k} \bar{Q}^{(1)}_{m_1} \cdots \bar{Q}^{(k)}_{m_k} \bar{Q}^{(1)}_{r_1} \cdots \bar{Q}^{(k)}_{r_k} \bar{Q}^{(1)}_{s_1} \cdots \bar{Q}^{(k)}_{s_k}. \tag{10}$$

The expectations in line (10) are all less than or equal 1, so the sum may be bounded by the number of nonzero terms there are. Since the $\bar{Q}^{(i)}$'s are independent and centered, the expectation of a product of $\bar{Q}_i$'s is zero if there is not at least the square of each term or the non-zero term $\bar{Q}^{(i)}_{(i^*,i^*)}$. Regardless of whether i^* is an index in an expectation, the number of possible other indices is at most $2k$, and for $j = 1, \dots, 2k$, there are nonzero expectations for j terms different from i^*. Once j integers out of $\{1, \dots, n\}$ are chosen, the number of ways to assign them to $4k$ positions is independent of n. Call this number $C_{j,4k}$. The number of ways to choose j different numbers out of n is $\binom{n}{j} \le n^j$. Set $C'_k = 2\max_{1\le j\le 2k} C_{j,2k}$. Then the term in line (10) is less than or equal to

$$\begin{aligned}\sum_{j=1}^{2k} C_{j,4k} n^j &\le 2k \max_{1\le j\le 2k} C_{j,4k} n^{2k}\\ &\le K_1 C'_{K_1} n^{2k}.\end{aligned}$$

Then

$$\begin{aligned}&\mathbb{P}(|\text{sum of all terms of the form (8)}| > \epsilon)\\ &\quad\le K_1 \cdot \mathbb{P}\left(|\text{one such term}| > \frac{\epsilon}{K_1}\right)\\ &\quad= K_1 \cdot \mathbb{P}\left(|\text{one such term}|^4 > \left(\frac{\epsilon}{K_1}\right)^4\right)\\ &\quad\le K_1 \cdot \left(K_1^4 \epsilon^{-4} C_{K_1} n^{-(2k+2)}\right) \cdot \left(K_1 C'_{K_1} n^{2k}\right)\\ &\quad= K_1^6 \epsilon^{-4} C_{K_1} C'_{K_1} n^{-2}.\end{aligned} \tag{11}$$

Since the terms (11) are summable with respect to n, the Borel-Cantelli lemma implies that the sum of the random terms

$$\sum_{l\in\{1,\dots,n\}^k} T^{(1)}_{i^*,l_1} \bar{Q}^{(1)}_{l_1} T^{(2)}_{l_1,l_2} \cdots T^{(k)}_{l_{k-1},l_k} \bar{Q}^{(k)}_{l_k} T^{(k+1)}_{l_{k+1},i^*}$$

converges almost surely to 0 for all k as $n \to \infty$.

This argument is independent of which index we denote i^*, so the term in line (6) converges almost surely to the same constant independent of the index. Our centering process isolates the nonrandom part of the sum in line (6), and we denote it $D_K(z)$. Thus the term in line (6) converges almost surely to $D_K(z)$. Since $|D_K(z) - D_{K+1}(z)| < |z|$, $\{D_k(z)\}_{k=1}^{\infty}$ is a Cauchy sequence, and, if we denote its limit $D(z)$, then we have $|D_k(z) - D(z)| \le \frac{|z|^k}{1-|z|}$.

To show the almost sure convergence we let $\epsilon > 0$ be given. Choose K large enough so that $\frac{|z|^K}{1-|z|} \le \epsilon/4$ and $|D_K(z) - D(z)| \le \epsilon/4$. In the following calculation the matrices are of dimension n, as indexed by the sum. Using (11) we have

$$\begin{aligned}&\sum_{n=1}^{\infty} \mathbb{P}\left(\left|\left[W^* A_i^{-1} W\right]_{ii} - D(z)\right| > \epsilon\right)\\ &\quad= \sum_{n=1}^{\infty} \mathbb{P}\left(\left|\left[F^* P \sum_{k=0}^{\infty} (-z)^k \left(PFQF^*P\right)^k PF\right]_{ii} - D(z)\right| > \epsilon\right)\end{aligned}$$

$$\leq \sum_{n=1}^{\infty} \mathbb{P}\Bigg(\Bigg| \Bigg[F^* P \sum_{k=0}^{K} (-z)^k \big(P F Q F^* P \big)^k P F \Bigg]_{ii} - D_K(z) \Bigg|$$
$$+ |D_K(z) - D(z)| + \Bigg| \Bigg[F^* P \sum_{k=K+1}^{\infty} (-z)^k \big(P F Q F^* P \big)^k P F \Bigg]_{ii} \Bigg| > \epsilon \Bigg)$$
$$\leq \sum_{n=1}^{\infty} \mathbb{P}\Bigg(\Bigg| \Bigg[F^* P \sum_{k=0}^{K} (-z)^k \big(P F Q F^* P \big)^k P F \Bigg]_{ii} - D_K(z) \Bigg| > \epsilon/2 \Bigg)$$
$$\leq \sum_{n=1}^{\infty} 2^4 K_1^6 \epsilon^{-4} C_{K_1} C'_{K_1} n^{-2},$$

which is finite. The Borel-Cantelli lemma now gives the almost sure convergence of $[W^* A_i^{-1} W]_{ii}$ to $D(z)$ for all i. □

The following lemma is essentially due to Tulino, Verdú, Caire and Shamai, and was developed in their work on (deterministic) Toeplitz matrices conjugated by a random Bernoulli projection matrix, which they call 'erasure matrices' [16]. The manipulations and the insight concerning the term $[W^* A_i^{-1} W]_{ii}$ are theirs. However, $[W^* A_i^{-1} W]_{ii}$ has a different form in the work presented here. As a consequence, the proof of Lemma 2.3 requires the preceding lemmas, which are our own. The proof given here is also self-contained. Equation (12) certainly holds more broadly than just the case covered in [16] and the work presented here. Similar general settings where such equations hold are proved in [17].

Lemma 2.3 *The η-transform of $P_n F_n Q_n F_n^* P_n$ converges almost surely to a function, which we denote $\eta_{p,q}$. This function is a solution to the implicit equation*

$$\eta_{p,q}(z) = \eta_Q\left(z - z \frac{p}{\eta_{p,q}(z)} \right) \tag{12}$$

for $z \in \mathbb{C}^+$, where η_Q is the asymptotic η-transform of Q_n.

Proof We return to the set-up given in (4) and (5). Since $\eta_{p,q}$ and η_Q are analytic functions, it suffices to prove (12) for $|z| < 1/2$ and $\Im z \neq 0$, which is contained in the region where we may apply Lemma 2.2. Recall that $W_n = P_n F_n$ and that w_i is the ith column of W_n. As defined in (4) and (5)

$$\begin{aligned} A_i &= I + z P F Q F^* P - z Q_{i,i} w_i w_i^* \\ &= I + z \sum_{j \neq i} Q_{j,j} w_j w_j^*. \end{aligned}$$

A_i is invertible for $z \notin [-\infty, 0]$, and one can verify directly that

$$(I + z P F Q F^* P)^{-1} = A_i^{-1} - \frac{z Q_{i,i}}{1 + z Q_{i,i} w_{i^*} A_i^{-1} w_i} A_i^{-1} w_i w_{i^*} A_i^{-1}. \tag{13}$$

Now we multiply both sides of (13) by $zQ_{i,i}w_iw_{i^*}$ and obtain

$$\begin{aligned}
&zQ_{i,i}w_iw_{i^*}(I+zPFQF^*P)^{-1}\\
&\quad = zQ_{i,i}w_iw_{i^*}A_i^{-1} - \frac{z^2Q_{i,i}^2w_iw_{i^*}}{1+zQ_{i,i}w_{i^*}A_i^{-1}w_i}A_i^{-1}w_iw_{i^*}A_i^{-1}\\
&\quad = zQ_{i,i}w_iw_{i^*}A_i^{-1}\left(1-\frac{zQ_{i,i}w_{i^*}A_i^{-1}w_i}{1+zQ_{i,i}w_{i^*}A_i^{-1}w_i}\right)\\
&\quad = \frac{zQ_{i,i}}{1+zQ_{i,i}w_{i^*}A_i^{-1}w_i}w_iw_{i^*}A_i^{-1}.
\end{aligned}$$

Summing over i gives

$$\begin{aligned}
zPFQF^*P(I+zPFQF^*P)^{-1} &= \sum_{i=1}^n zQ_{i,i}w_iw_{i^*}(I+zPFQF^*P)^{-1} &(14)\\
&= \sum_{i=1}^n \frac{zQ_{i,i}}{1+zQ_{i,i}w_{i^*}A_i^{-1}w_i}w_iw_{i^*}A_i^{-1}. &(15)
\end{aligned}$$

Using (15),

$$\begin{aligned}
\mathrm{tr}\big((I+zPFQF^*P)^{-1}\big) &= 1-\mathrm{tr}\big(zQF^*P(I+zPFQF^*P)^{-1}PF^*\big)\\
&= 1-\mathrm{tr}\left(\sum_{i=1}^n \frac{zQ_{i,i}w_iw_{i^*}A_i^{-1}}{1+zQ_{i,i}w_{i^*}A_i^{-1}w_i}\right)\\
&= 1-\frac{1}{n}\sum_{i=1}^n \frac{zQ_{i,i}w_{i^*}A_i^{-1}w_i}{1+zQ_{i,i}w_{i^*}A_i^{-1}w_i}\\
&= \frac{1}{n}\sum_{i=1}^n \frac{1}{1+zQ_{i,i}w_{i^*}A_i^{-1}w_i}.
\end{aligned}$$

We note

$$\begin{aligned}
&\mathrm{tr}\big(F^*P(I+zPFQF^*P)^{-1}PF\big)\\
&\quad = \mathrm{tr}\big((I+zPFQF^*P)^{-1}\big) - \frac{1}{n}\sum_{i=1}^n 1\{P_{i,i}=0\}\big[(I+zPFQF^*P)^{-1}\big]_{i,i}\\
&\quad = \mathrm{tr}\big((I+zPFQF^*P)^{-1}\big) - \frac{1}{n}\sum_{i=1}^n 1\{P_{i,i}=0\}, \qquad (16)
\end{aligned}$$

where the last line follows from using that $[(I+zPFQF^*P)^{-1}]_{i,i}=1$ when $P_{i,i}=0$.

Lemma 2.2 states that $w_i^* A_i^{-1} w_i$ converges almost surely to a number $D(z)$ independent of i as $n \to \infty$. Thus

$$\eta_{p,q}(z) = \lim_{n\to\infty} \mathop{\mathbb{E}}_{P_n,Q_n} \operatorname{tr}\big((I + zPFQF^*P)^{-1}\big)$$

$$= \lim_{n\to\infty} \mathop{\mathbb{E}}_{P_n,Q_n} \frac{1}{n}\sum_{i=1}^{n} \frac{1}{1 + zQ_{i,i} w_{i^*} A_i^{-1} w_i}.$$

For each i we write $w_{i^*} A_i^{-1} w_i = D(z) + X_{n,i}$, where $D(z)$ is the deterministic part independent of dimension, as in Lemma 2.2. Since $|z| < 1/2$, $\|A_i^{-1}\| \le 2$ and $|w_{i^*} A_i^{-1} w_i| < 2$ for every i, all n and every realization of P_n and Q_n. Thus $|D(z)| < 2$ as well for $|z| < 1/2$. This allows us to bound the denominator in what follows from below by $1/4$. As in the proof of Lemma 2.2, we truncate the expansion of each $X_{n,i}$ at K terms, such that the remainder is smaller in absolute value than $\epsilon/4$, regardless of the realization. We denote the truncation $X_{n,i}^{(K)}$.

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+zQ_{i,i}(D(z)+X_{n,i})} - \eta_Q\big(zD(z)\big)\right| > \epsilon\right)$$

$$\le \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+zQ_{i,i}(D(z)+X_{n,i})} - \frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+zQ_{i,i}D(z)}\right| + \left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+zQ_{i,i}D(z)} - \eta_Q\big(D(z)z\big)\right| > \epsilon\right)$$

$$\le \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+zQ_{i,i}(D(z)+X_{n,i})} - \frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+zQ_{i,i}D(z)}\right| > \frac{\epsilon}{2}\right) + \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+zQ_{i,i}D(z)} - \eta_Q\big(D(z)z\big)\right| > \frac{\epsilon}{2}\right) \tag{17}$$

$$\le 2\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\frac{zQ_{i,i}X_{n,i}^{(K)}}{(1+zQ_{i,i}D(z))(1+zQ_{i,i}(D(z)+X_{n,i}))}\right|^4 > \left(\frac{\epsilon}{4}\right)^4\right) \tag{18}$$

$$\le 2^{17}\epsilon^{-4}|z|^4 \frac{1}{n^4} \sum_{i_1,\ldots,i_4=1}^{n} \mathbb{E}_{P_n,Q_n}\big|X_{n,i_1}^{(K)} \cdots X_{n,i_4}^{(K)}\big| \tag{19}$$

$$\le 2^{17}\epsilon^{-4}|z|^4 C_z' n^{-2}, \tag{20}$$

where we use the exponential Bernstein bound for the term in line (17) and incorporate it into the 2 in line (18). In line (19) we have products of sums of terms of the form (9) with powers summing to 4, and so the work of Lemma 2.2 applies. The constant may differ, and so we denote it C_z' in line (20). Using the Borel-Cantelli lemma, we now have the almost sure convergence

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+zQ_{i,i}w_{i^*}A_i^{-1}w_i}\stackrel{a.s}{\to}\eta_Q\big(zD(z)\big).$$

Using this and Lemma 2.2 again,

$$\begin{aligned}D(z)\eta_{p,q}(z) &= \lim_{n\to\infty}\mathbb{E}_{P_n,Q_n}\frac{1}{n}\sum_{i=1}^{n}\frac{w_{i^*}A_i^{-1}w_i}{1+zQ_{i,i}w_{i^*}A_i^{-1}w_i}\\ &= \lim_{n\to\infty}\mathbb{E}_{P_n,Q_n}\frac{1}{n}\sum_{i=1}^{n}w_{i^*}(I+zWQW^*)^{-1}w_i\\ &= \lim_{n\to\infty}\mathbb{E}_{P_n,Q_n}\mathrm{tr}\big(W^*(I+zWQW^*)^{-1}W\big)\\ &= \eta_{p,q}(z)-p, \qquad (21)\end{aligned}$$

where (21) follows from taking the limit with respect to n of the expectation of (16). □

Proposition 2.4 *Let P_n and Q_n be independent Bernoulli diagonal projection matrices with expected traces $1-p\in(0,1)$ and $1-q\in(0,1)$ respectively. Then the η-transform of $P_nF_nQ_nF_n^*P_n$ converges almost surely to the asymptotic η-transform*

$$\eta_{p,q}(z)=\frac{1+(p+q)z+\sqrt{1+(2(p+q)-4pq)z+((p+q)^2-4pq)z^2}}{2(1+z)}.$$

Proof The matrices $F_nQ_nF_n^*$ have only eigenvalues 1 and 0, and their limiting η-transform is

$$\eta_{FQF^*}(z)=\eta_Q(z)=\frac{1+qz}{1+z}.$$

Applying (12) from Lemma 2.3 yields

$$\eta_{p,q}(z)=\frac{1+q\big(z-z\frac{p}{\eta_{p,q}(z)}\big)}{1+\big(z-z\frac{p}{\eta_{p,q}(z)}\big)},$$

which leads to the equation

$$(1+z)\eta_{p,q}^2(z)-\big(1+(p+q)z\big)\eta_{p,q}(z)+pqz=0.$$

This equation has the solutions

$$\frac{1+(p+q)z\pm\sqrt{(1+(p+q)z)^2-4(1+z)pqz}}{2(1+z)}.$$

Noting that $\eta_{p,q}(0)$ must equal 1, we choose the solution with addition. We thus have

$$\eta_{p,q}(z) = \frac{1+(p+q)z+\sqrt{1+(2(p+q)-4pq)z+(p-q)^2z^2}}{2(1+z)}. \tag{22}$$

□

3 Limiting Empirical Distributions

Theorem 3.1 *For $i = 1, \dots, n$ let i be contained in Ω_n independently with probability $(1-q)$ and, also independently, let i be included in T_n with probability $(1-p)$. Then the empirical distribution of the $\min(|T_n|, |\Omega_n|)$ largest eigenvalues of $F_{\Omega_n T_n} F^*_{\Omega_n T_n}$ converges almost surely to*

$$\begin{aligned} f_{p,q}(x) &= \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)(1-\max(p,q))} \cdot I_{(r_-,r_+)}(x) + \frac{\max(0, 1-(p+q))}{1-\max(p,q)} \cdot \delta(x-1) \end{aligned} \tag{23}$$

where

$$r_- = \left(\sqrt{p(1-q)} - \sqrt{q(1-p)}\right)^2$$

and

$$r_+ = \left(\sqrt{p(1-q)} + \sqrt{q(1-p)}\right)^2.$$

Note that $r_- = 0$ only when $p = q$; that is, when $F_{\Omega_n T_n} F^*_{\Omega_n T_n}$ is asymptotically square. Therefore, when $p \neq q$ the support of the limiting distribution begins at $r_- > 0$. Also, $r_+ = 1$ only when $p+q = 1$, so when $p+q \neq 1$ there is a gap in the support of the limiting distribution from r_+ to 1. When $r_- \neq 0$ and $r_+ \neq 1$ the continuous part of the measure begins at r_- and makes an arc ending at r_+. When $r_- = 0$ or $r_+ = 1$, the continuous part of (23) tends to ∞ as $x \to 0$ or $x \to 1$. When $p+q < 1$ there is a point mass of measure $1-(p+q)$ at 1, and when $p+q > 1$, the support stops at r_+ and there is no point mass at 1. In Fig. 1 the continuous part of the asymptotic distribution is plotted against empirical values for several parameter pairs p and q.

We discuss the relationship between Theorem 3.1 and uncertainty principles and other areas before turning to the proof. In the following we assume that the dimension is n. The norm of a DFT matrix with a set of rows and columns removed equals 1 if and only if there exists a vector with time support corresponding to the remaining columns and frequency support corresponding to the remaining rows. That is, denoting $Fx = \hat{x}$, in the notation of this paper, $\|F_{\Omega T}\| = 1$ if and only if there exists $x \in \mathbb{C}^n$

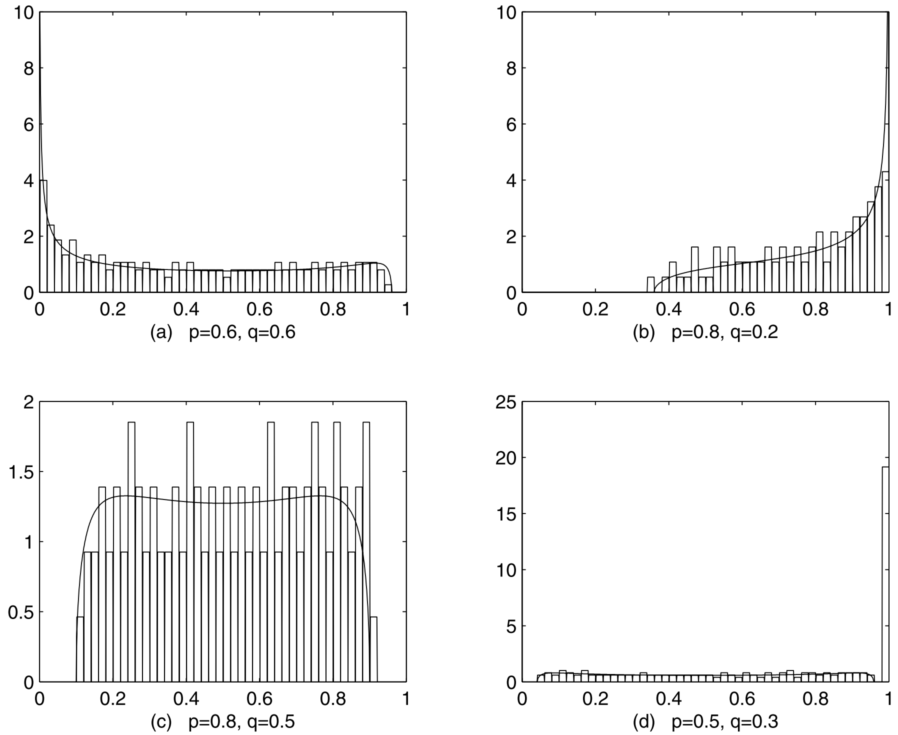

Fig. 1 Empirical eigenvalue distribution for *one* realization plotted against continuous part of asymptotic distribution. In each case, the original DFT matrix had dimensions 500×500

such that $\operatorname{supp}(x) \subset T$ and $\operatorname{supp}(\hat{x}) \subset \Omega$. One is then interested in determining conditions on the cardinality of $|\Omega|$ and $|T|$ such that $\|F_{\Omega T}\| < 1$, where generally at least one set is random and a statement is made in some probabilistic form. This non-asymptotic question has been studied intensively over the last ten years. Recent results and a general discussion can be found in [14]. While Theorem 3.1 does not assert the non-existence of any vectors, it does state when certain vectors do exist and sheds light on one of the main theorems in this area, namely that of Tao.

Theorem 3.2 (Tao [13]) *If n is prime and $|\Omega| + |T| < n$, then $\|F_{\Omega T}\| < 1$.*

Since Tao's theorem requires n to be prime, it precludes the case that $\frac{|\Omega|}{n} + \frac{|T|}{n} = 1$. Theorem 3.1 says that if $\frac{|\Omega_n|}{n} + \frac{|T_n|}{n} \to 1$, then $\|F_{\Omega_n T_n}\|$ converges almost surely to 1. Almost sure convergence and the existence of infinitely many primes imply that for all $\epsilon_1, \epsilon_2 > 0$, there exists a prime n and sets $\Omega_n, T_n \subset \{1, \dots, n\}$ such that $\frac{|\Omega_n|}{n} + \frac{|T_n|}{n} < 1 + \epsilon_1$ and $\|F_{\Omega_n T_n}\| > 1 - \epsilon_2$. In fact, the proportion of subsets for which $\|F_{\Omega_n T_n}\| < 1 - \epsilon_2$ converges to 0 as $n \to \infty$.

A further area of interest is the smallest eigenvalue of $F_{\Omega_n T_n} F^*_{\Omega_n T_n}$; in particular, one would like to bound the smallest eigenvalue away from 0. See [12] for recent results in the setting of independent matrix entries. While Theorem 3.1 does not make any statement on when the smallest eigenvalue is strictly positive, it does say that if $\frac{|\Omega_n|}{|Tn|} \to 1$, then the smallest eigenvalue of $F_{\Omega_n T_n} F^*_{\Omega_n T_n}$ converges to 0. This corresponds to the behavior of square matrices with independent entries, though in that case non-asymptotic bounds away from 0 exist [12].

Theorem 3.1 and some numerical experiments suggest the obvious conjecture that the largest and smallest eigenvalues converge to the edge of the limiting support. This would imply that Tao's result gives the general uncertainty principle behavior for DFT and random unitary matrices, and that the submatrices that do not have this behavior have measure zero asymptotically. As was the case for the Wigner and Wishart distributions, we hope that the limiting empirical distribution is helpful in determining the behavior of the extremal eigenvalues.

We note, lastly, that a potential further step in this direction is restricted isometry properties. Here one set, say T_n, is taken at random and one seeks to bound the extremal eigenvalues of $F_{\Omega_n T_n} F^*_{\Omega_n T_n}$ when Ω_n ranges over all subsets of certain cardinality. This property of random matrices is central to compressed sensing and has received enormous attention in recent years; see [4] for a recent overview.

Proof of Theorem 3.1 By Proposition 2.4 we have

$$\left| \frac{1}{n} \sum_{i=1}^{n} \frac{1}{1 + z\lambda_i} - \eta_{p,q}(z) \right| \xrightarrow{a.s.} 0 \tag{24}$$

for all $z \in \mathbb{C}^+$. By applying this to $-\frac{1}{z}$, multiplying both terms in (24) by $-\frac{1}{z}$ and using (3), we obtain

$$\left| \frac{1}{n} \sum_{i=1}^{n} \frac{1}{z - \lambda_i} - m_{p,q}(z) \right| \xrightarrow{a.s.} 0.$$

Thus, the random measures induced by the eigenvalues at each dimension n converge almost surely to the probability measure corresponding to $m_{p,q}$ [1] (Theorem 2.4.4(c) in [1] also holds for almost sure convergence). By (3) the limiting Stieltjes transform is

$$\begin{aligned} m_{p,q}(z) &= \frac{-1}{z} \cdot \frac{1 - (p+q)\frac{1}{z} + \sqrt{1 + (2(p+q) - 4pq)(\frac{-1}{z}) + (p-q)^2(\frac{1}{z})^2}}{2(1 - \frac{1}{z})} \\ &= \frac{1 - (p+q)\frac{1}{z} + \sqrt{1 - (2(p+q) - 4pq)\frac{1}{z} + (p-q)^2\frac{1}{z^2}}}{2(1 - z)}, \end{aligned}$$

and we are interested in the inverse Stieltjes transform

$$\lim_{\omega \to 0^+} \frac{1}{\pi} \Im\, m_{p,q}(x + i\omega).$$

$m_{p,q}(x+i\omega)$ is a continuous function of both x and ω for $x \in (0,1)$. Thus, using equation (1) for $x \in (0,1)$,

$$f_{p,q}(x) = \frac{1}{\pi}\Im \frac{1-(p+q)\frac{1}{x}+\sqrt{1-(2(p+q)-4pq)\frac{1}{x}+(p-q)^2\frac{1}{x^2}}}{2(1-x)}$$
$$= \Im \frac{\sqrt{1-(2(p+q)-4pq)\frac{1}{x}+(p-q)^2\frac{1}{x^2}}}{2\pi(1-x)}. \tag{25}$$

Imitating Marčenko and Pastur [10], we denote the roots of the equation $x^2-(2(p+q)-4pq)x+(p-q)^2=0$ by r_- and r_+. These values are $r_\pm = (\sqrt{p(1-q)} \pm \sqrt{q(1-p)})^2$, as defined in the statement of the theorem. For $x \in (0,1)$ equation (25) is now

$$f_{p,q}(x) = \Im \frac{\sqrt{1-(2(p+q)-4pq)\frac{1}{x}+(p-q)^2\frac{1}{x^2}}}{2\pi(1-x)}$$
$$= \Im \frac{\sqrt{(1-\frac{r_-}{x})(1-\frac{r_+}{x})}}{2\pi(1-x)}$$
$$= \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)} \cdot I_{[r_-,r_+]}(x). \tag{26}$$

We now determine the density at $x=0$, for which we need to find the asymptotic proportion of zero eigenvalues of $P_nF_nQ_nF_n^*P_n$. This proportion is given by $\lim_{r\to\infty}\eta_{p,q}(r)$, since the latter quantity gives the measure of the set $\{0\}$ with respect to the limiting distribution. We have

$$\lim_{r\to\infty}\eta_{p,q}(r) = \lim_{r\to\infty}\frac{1+(p+q)r+\sqrt{1+(2(p+q)-4pq)r+(p-q)^2r^2}}{2(1+r)}$$
$$= \frac{(p+q)+|p-q|}{2}$$
$$= \max(p,q). \tag{27}$$

Lastly, we investigate the point $x=1$. We denote by $\mu(\{1\})$ the measure of the set $\{1\}$ with respect to the limiting distribution. Let X denote the random variable of the eigenvalues given by selecting one the n eigenvalues of a realization according to a uniform distribution. We first must address the convergence of $\sum_{k=0}^{\infty}(-z)^k\mathbb{E}X^k$. If $p+q \neq 1$ and $p \neq q$, then $0 < r_-$ and $r_+ < 1$, and

$$\mathbb{E}X^k \leq \left\| \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)} I_{[r_-,r_+]}\right\|_\infty \cdot \int_0^{r_+} r_+^k\,dx + \mu(1)$$
$$= \left\| \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)} I_{[r_-,r_+]}\right\|_\infty \cdot r_+^{k+1} + \mu(1).$$

If $p+q \neq 1$ and $p=q$, which implies $r_-=0$, then we pick a small ϵ and have

$$\mathbb{E}X^k \le \int_0^\epsilon \epsilon^k \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)}dx + \left\| \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)} I_{[\epsilon,r_+]} \right\|_\infty \cdot \int_0^{r_+} r_+^k dx + \mu(1).$$

When $p+q=1$ and when z belongs to a region of convergence to be determined shortly, equation (2) implies

$$\begin{aligned}\eta_{p,q}(z) &= \sum_{k=0}^\infty (-z)^k \mathbb{E}X^k \\ &= \sum_{k=0}^\infty (-z)^k \int_{r_-}^{r_+} x^k \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)}dx + \sum_{k=0}^\infty \mu(1)(-z)^k \\ &= \sum_{k=0}^\infty (-z)^k \int_{r_-}^{r_+} x^k \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)}dx + \frac{\mu(1)}{1+z}. \end{aligned} \tag{28}$$

Since we have assumed that neither p or q equals 1 or 0, we have $r_- < 1$. We show that the sum on the left side of (28) converges for $|z| \le 1/r_-$. We define

$$\lfloor k \rfloor_e = \begin{cases} k & k \text{ even} \\ k-1 & k \text{ odd.} \end{cases}$$

$$\begin{aligned}&\left| \sum_{k=0}^\infty (-z)^k \int_{r_-}^1 x^k \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)}dx \right| \\ &\quad \le \sum_{k=0}^2 |z|^k \mathbb{E}X^k + \sum_{k=3}^\infty |z|^k \int_{r_-}^1 x^k \frac{\sqrt{\frac{1}{x}-1}}{2\pi(1-x)}dx \\ &\quad = \sum_{k=0}^2 |z|^k \mathbb{E}X^k + \sum_{k=3}^\infty |z|^k \int_{r_-}^1 x^k \frac{\sqrt{\frac{1-x}{x}}}{2\pi(1-x)}dx \\ &\quad = \sum_{k=0}^2 |z|^k \mathbb{E}X^k + \frac{1}{2\pi}\sum_{k=3}^\infty |z|^k \int_{r_-}^1 x^{k-1/2}(1-x)^{-1/2}dx \\ &\quad \le \sum_{k=0}^2 |z|^k \mathbb{E}X^k + \frac{|z|}{2\pi}\sum_{k=2}^\infty |z|^k \int_{r_-}^1 x^k(1-x)^{-1/2}dx \end{aligned}$$

$$\leq \sum_{k=0}^{2} |z|^k \mathbb{E}X^k + \frac{|z|}{2\pi} \sum_{k=2}^{\infty} |z|^k \int_{r_-}^{1} x^{\lfloor k \rfloor_e} (1-x)^{-1/2} dx$$

$$= \sum_{k=0}^{2} |z|^k \mathbb{E}X^k + \frac{|z|}{2\pi} \sum_{k=2}^{\infty} |z|^k 2(1-r_-)^{1/2} \sum_{j=0}^{\lfloor k \rfloor_e} \frac{(-1)^j \binom{\lfloor k \rfloor_e}{j} (1-r_-)^{\lfloor k \rfloor_e - j}}{2(\lfloor k \rfloor_e - j) + 3}. \qquad (29)$$

The integral in line (29) is given by equation 2.221 in [8]. One may verify that

$$\frac{\binom{n}{k}(1-r_+)^{n-k}}{2(n-k)+3} - \frac{\binom{n}{k-1}(1-r_+)^{n-k+1}}{2(n-k+1)+3}$$
$$< \binom{n}{k}(1-r_+)^{n-k} - \binom{n}{k-1}(1-r_+)^{n-k+1},$$

so that

$$\sum_{j=0}^{\lfloor k \rfloor_e} \frac{(-1)^j \binom{\lfloor k \rfloor_e}{j} (1-r_-)^{\lfloor k \rfloor_e - j}}{2(\lfloor k \rfloor_e - j) + 3} \leq \sum_{j=0}^{\lfloor k \rfloor_e} (-1)^j \binom{\lfloor k \rfloor_e}{j} (1-r_-)^{\lfloor k \rfloor_e - j} = r_-^{\lfloor k \rfloor_e}.$$

Thus

$$(29) \leq \sum_{k=0}^{2} |z|^k \mathbb{E}X^k + \frac{|z|}{2\pi}(1-r_-)^{1/2} \sum_{k=2}^{\infty} |z|^k r_-^{\lfloor k \rfloor_e},$$

which converges for all $|z| < 1/r_-$. This gives

$$\mu(\{1\}) = (1+z)\eta_{p,q}(z) - (1+z)\sum_{k=0}^{\infty} z^k \int_{r_-}^{r_+} x^k \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)} dx, \qquad (30)$$

where equation (30) holds for all values $|z| < 1$, and the sum on the right side of equation (30) remains finite as $z \to -1$. Convergence of the necessary sums and integrals is now established for all $p, q \in (0, 1)$. Using the equation for $\eta_{p,q}(z)$ from Proposition 2.4,

$$(1+z)\eta_{p,q}(z) = \frac{1+(p+q)z+\sqrt{1+(2(p+q)-4pq)z+((p+q)^2-4pq)z^2}}{2}.$$

Allowing $z \to -1$, we have

$$\mu(\{1\}) = \frac{1-(p+q)+\sqrt{(1-(p+q))^2}}{2}$$
$$= \frac{1-(p+q)+|1-(p+q)|}{2}. \qquad (31)$$

When $p+q \geq 1$, (31) is equal to 0, and when $p+q < 1$, (31) is equal to $1-(p+q)$. From (27) and (31), it follows that when $p+q > 1$

$$\int_{r_-}^{r_+} \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)} dx = 1-\max(p,q),$$

and when $p+q \leq 1$,

$$\int_{r_-}^{r_+} \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)} dx = \min(p,q).$$

Now it only remains to remove the point mass at 0 and normalize the distribution by $1/(1-\max(p,q))$, that is by $1/(1-\mu(\{0\}))$. □

For the following corollary we define the n singular values of the matrix $P_n F_n Q_n$ to be the (positive) square roots of the n eigenvalues of the matrix $P_n F_n Q_n F_n^* P_n$. We thus have the following limiting distribution for the singular values of $P_n F_n Q_n$.

Corollary 3.3 *For $i = 1, \ldots, n$ let i be contained in Ω_n independently with probability $(1-q)$ and, also independently, let i be included in T_n with probability $(1-p)$. Then the empirical distribution of the* $\min(|T_n|, |\Omega_n|)$ *largest singular values of $F_{\Omega_n T_n}$ converges almost surely to*

$$f_{p,q}^s(x) = \frac{\sqrt{x^2(1-\frac{r_-^s}{x})(\frac{r_+^s}{x}-1)}}{2\pi(1-x^2)(1-\max(p,q))} \cdot I_{(r_-^s, r_+^s)}(x) + \frac{\max(0, 1-(p+q))}{1-\max(p,q)} \cdot \delta(x-1)$$

where

$$r_\pm^s = |\sqrt{p(1-q)} \pm \sqrt{q(1-p)}|.$$

Proof The measures for the singular values 0 and 1 are the same as the eigenvalues 0 and 1 of the previous theorem. Also, the continuous part of the measure will clearly have support $[r_-^s, r_+^s] = [\sqrt{r_-}, \sqrt{r_+}]$. Using (26), we have

$$f_{p,q}^s(x) = 2x f_{p,q}\big(x^2\big) = \frac{\sqrt{x^2(1-\frac{r_-}{x^2})(\frac{r_+}{x^2}-1)}}{\pi(1-x^2)} \cdot I_{[\sqrt{r_-}, \sqrt{r_+}]}(x).$$

□

3.1 Random Unitary Matrices with Haar Distribution

Using the work of Garnaev and Gluskin, we can show that the results of the previous section hold as well for random submatrices of random unitary matrices with Haar distribution. Let $\mathcal{U}_n$ denote the $n \times n$ random unitary matrix with Haar distribution and define $\mathcal{U}_{\Omega_n T_n}$ analogously to $F_{\Omega_n T_n}$. We will show here that the limiting empirical singular value distribution of $\mathcal{U}_{\Omega_n T_n}$ is the same as $F_{\Omega_n T_n}$.

Here we note that the eigenvalue distribution of random unitary matrices with a fixed proportion of the bottom rows and right-most columns removed has already been studied. In the case when these proportions are equal, i.e. when the resulting matrix is square, the limiting empirical eigenvalue density was derived in [11], which builds on the work in [22].

Corollary 3.4 *For $i = 1, \ldots, n$ let i be contained in Ω_n independently with probability $(1-q)$ and, also independently, let i be included in T_n with probability $(1-p)$. Then the empirical distribution of the $\min(|T_n|, |\Omega_n|)$ largest eigenvalues of $\mathcal{U}_{\Omega_n T_n}\mathcal{U}^*_{\Omega_n T_n}$ converges almost surely to*

$$f_{p,q}(x) = \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)(1-\max(p,q))} \cdot I_{(r_-,r_+)}(x) + \frac{\max(0, 1-(p+q))}{1-\max(p,q)} \cdot \delta(x-1)$$

where

$$r_\pm = \left(\sqrt{p(1-q)} \pm \sqrt{q(1-p)}\right)^2.$$

Note that the discussion following Theorem 3.1 also applies here.

Proof (sketch) With high probability, the matrix P_n selects roughly $(1-p)n$ columns from the unitary matrix $\mathcal{U}_n$. The work of Garnaev and Gluskin [5, 7] gives that, with high probability, the $\|\cdot\|_{l_n^\infty}$-norm of each of these columns is less than a constant M. Additionally, the expectation of the inner-product of any two columns is zero. Thus, taking into account the constant M, these random vectors satisfy a bound analogous to Lemma 2.1 for the DFT vectors. Once this is established, all the other proofs then hold as well for random unitary matrices. □

Acknowledgement The author thanks Roland Speicher, Walid Hachem, Antonia Tulino and Sergio Verdú for helpful discussions and comments.

References

1. Anderson, G., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices. Cambridge University Press, Cambridge (2009)
2. Bai, Z.D., Yin, Y.Q.: Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. Ann. Probab. **16**(4), 1729–1741 (1988)
3. Bai, Z.D., Yin, Y.Q.: Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. Ann. Probab. **21**(3), 1275–1294 (1993)
4. Candès, E.J.: Compressive sampling. In: International Congress of Mathematicians, vol. III, pp. 1433–1452. Eur. Math. Soc., Zürich (2006)
5. Garnaev, A.Y., Gluskin, E.D.: On widths of the Euclidean ball. Sov. Math., Dokl. **30**, 200–204 (1984) (English. Russian original)

6. Geman, S.: A limit theorem for the norm of random matrices. Ann. Probab. **8**(2), 252–261 (1980)
7. Gluskin, E.D.: Norms of random matrices and widths of finite-dimensional sets. Math. USSR, Sb. **48**, 173–182 (1984) (English)
8. Gradshteyn, I.S., Ryzhik, I.M.: Table of Integrals, Series, and Products. Academic Press, New York (2000)
9. Grenander, U., Silverstein, J.: Spectral analysis of networks with random topologies. SIAM J. Appl. Math. **32**(2), 499–519 (1977)
10. Marčenko, V.A., Pastur, L.A.: Distribution of eigenvalues in certain sets of random matrices. Mat. Sb. (N.S.) **72**(114), 507–536 (1967)
11. Petz, D., Réffy, J.: Large deviation for the empirical eigenvalue density of truncated Haar unitary matrices. Probab. Theory Relat. Fields **133**(2), 175–189 (2005) (English)
12. Rudelson, M., Vershynin, R.: The Littlewood-Offord problem and invertibility of random matrices. Adv. Math. **218**(2), 600–633 (2008)
13. Tao, T.: An uncertainty principle for cyclic groups of prime order. Math. Res. Lett. **12**(1), 121–127 (2005)
14. Tropp, J.A.: On the linear independence of spikes and sines. J. Fourier Anal. Appl. **14**(5), 838–858 (2008)
15. Tulino, A., Verdú, S.: Random matrices and wireless communications. Found. Trends Inf. Theory **1**(1) (2004)
16. Tulino, A., Verdú, S., Caire, G., Shamai, S.: Capacity of the gaussian erasure channel. Preprint, Technion CCIT, no. 655 (2007). http://www3.ee.technion.ac.il/ccit/info/Publication/Scientific_e.asps
17. Tulino, A.M., Caire, G., Shamai, S., Verdu, S.: Capacity of channels with frequency-selective and time-selective fading. IEEE Trans. Inf. Theory **56**(3), 1187–1215 (2010)
18. Wigner, E.: Characteristic vectors of bordered matrices with infinite dimensions. Ann. Math. **62**(2), 548–564 (1955)
19. Wigner, E.: On the distribution of the roots of certain symmetric matrices. Ann. Math. **67**(2), 325–327 (1958)
20. Wishart, J.: The generalized product moment distribution in samples from a normal multivariate population. Biometrika **20A**, 32–52 (1928)
21. Yin, Y.Q., Bai, Z.D., Krishnaiah, P.R.: On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix. Probab. Theory Relat. Fields **78**(4), 509–521 (1988)
22. Życzkowski, K., Sommers, H.-J.: Truncations of random unitary matrices. J. Phys. A **33**(10), 2045–2057 (2000)