

Kesten–McKay law for random subensembles of Paley equiangular tight frames

Mark Magsino

Dustin G. Mixon

Hans Parshall

Abstract

We apply the method of moments to prove a recent conjecture of Haikin, Zamir and Gavish [22] concerning the distribution of the singular values of random subensembles of Paley equiangular tight frames. Our analysis applies more generally to real equiangular tight frames of redundancy 2, and we suspect similar ideas will eventually produce more general results for arbitrary choices of redundancy.

2010 Mathematics Subject Classification. 42C15, 15B52, 94C30

Key words and phrases. equiangular tight frames, method of moments, Borel’s triangle

1 Introduction

Frame theory concerns redundant representations in a Hilbert space. A **frame** [16] is a sequence $\{\varphi_i\}_{i \in I}$ in a Hilbert space H for which there exist $\alpha, \beta \in (0, \infty)$ such that

$$\alpha\|x\|^2 \leq \sum_{i \in I} |\langle x, \varphi_i \rangle|^2 \leq \beta\|x\|^2,$$

for every $x \in H$. If every φ_i has unit norm, then we say the frame is **unit norm**, and if $\alpha = \beta$, we say the frame is **tight** [12]. In the special case where $H = \mathbb{R}^d$, a frame is simply a spanning set, but unit norm tight frames are still interesting and useful [5, 27]. For example, **equiangular tight frames** are unit norm tight frames with the additional property that $|\langle \varphi_i, \varphi_j \rangle|$ is constant over the choice of pair $\{i, j\}$. Equiangular tight frames are important because they necessarily span optimally packed lines, which in turn find applications in multiple description coding [33], digital fingerprinting [28], compressed sensing [2], and quantum state tomography [30]; see [18] for a survey.

Various applications demand control over the singular values of subensembles of frames. In quantum physics, Weaver’s conjecture [38] (equivalent to the Kadison–Singer problem [24, 10], and recently resolved in [25]) concerns the existence of subensembles of unit norm tight frames with appropriately small spectral norm. Compressed sensing [9, 14] has spurred the pursuit of explicit frames with the property that every subensemble is well conditioned [13, 6, 2], or at least that most subensembles are well conditioned [35, 8]. In this vein, Gurevich and Hadani [19] established that random subensembles of incoherent ensembles satisfy Wigner’s semicircle law when the subensemble size is a vanishing fraction of the full ensemble size. Motivated by applications in erasure-robust analog coding, Haikin, Zamir and Gavish [22, 21]

recently considered the case in which this fraction is fixed. Of particular interest are random subensembles of equiangular tight frames, and in this paper, we consider equiangular tight frames comprised of $2d$ vectors in $\mathbb{R}^d$, which correspond to symmetric conference matrices. (Note that such frames have already received some attention in the context of compressed sensing [2, 3].)

An $n \times n$ matrix S is said to be a **conference matrix** if

- (i) $S_{ii} = 0$ for every $i \in [n]$,
- (ii) $S_{ij} \in \{\pm 1\}$ for every $i, j \in [n]$ with $i \neq j$, and
- (iii) $S^\top S = (n - 1)I$.

A symmetric conference matrix of order n exists whenever $n - 1 \equiv 1 \pmod{4}$ is a prime power (by a Paley-based construction), and only if $n \equiv 2 \pmod{4}$ and $n - 1$ is a sum of two squares [23]. Explicitly, the Paley conference matrices are obtained by building a circulant matrix from the Legendre symbol and then padding with ones, for example:

$$\left\{ \left(\frac{x}{5} \right) \right\}_{x=0}^4 = (0, +, -, -, +) \quad \implies \quad S = \begin{bmatrix} 0 & + & + & + & + & + \\ + & 0 & + & - & - & + \\ + & + & 0 & + & - & - \\ + & - & + & 0 & + & - \\ + & - & - & + & 0 & + \\ + & + & - & - & + & 0 \end{bmatrix}, \quad (1)$$

where “ $\pm$ ” denotes ± 1 . One may verify that the above example satisfies $S^2 = 5I$. For every $n \times n$ symmetric conference matrix S , it holds that $I + \frac{1}{\sqrt{n-1}}S$ is the Gram matrix of an equiangular tight frame consisting of n vectors in $\mathbb{R}^{n/2}$ [33]. In particular, the equiangular tight frames that arise from the Paley conference matrices are known as Paley equiangular tight frames. In what follows, we consider random principal submatrices of symmetric conference matrices with the understanding that they may be identified with the Gram matrix of a random subensemble of the corresponding equiangular tight frame.

Given an $n \times n$ symmetric matrix Z with eigenvalues $\lambda_1 \leq \dots \leq \lambda_n$, we let μ_Z denote the uniform probability measure over the spectrum of Z (counted with multiplicity):

$$\mu_Z := \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}.$$

This is known as the **empirical spectral distribution** of Z . If Z is a random matrix, then its empirical spectral distribution μ_Z is a random measure. We say a sequence $\{\zeta_i\}_{i=1}^\infty$ of random measures **converges weakly almost surely** to a non-random measure μ if for every bounded continuous function $f: \mathbb{R} \rightarrow \mathbb{R}$, it holds that the random variable $\int_{\mathbb{R}} f(x) d\zeta_i(x)$ converges to $\int_{\mathbb{R}} f(x) d\mu(x)$ almost surely.

We are interested in random matrices of a particular form. Let $\mathcal{I}$ denote a random subset of $[n]$ such that the events $\{1 \in \mathcal{I}\}, \dots, \{n \in \mathcal{I}\}$ have probability p and are independent. Then for any fixed $n \times n$ matrix A , we write $X \sim \text{Sub}(A, p)$ to denote the (random) principal submatrix of A with rows and columns indexed by $\mathcal{I}$. For instance, take S from (1), put

$p = 1/3$, and draw $X \sim \text{Sub}(S, p)$. Then the random index subset $\mathcal{I} \subseteq \{1, \dots, 6\}$ equals $\{2, 4\}$ with probability $p^2(1-p)^4 = 16/729$, in which case $X = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Note that the size of the random matrix X is a random variable with binomial distribution.

Following [15], we define the **Kesten–McKay distribution** with parameter $v \geq 2$ by

$$d\mu_{\text{KM}(v)} = \begin{cases} \frac{v\sqrt{4(v-1)-x^2}}{2\pi(v^2-x^2)}, & \text{if } x^2 \leq 4(v-1) \\ 0, & \text{otherwise} \end{cases} dx.$$

Recall that a **lacunary sequence** is a set $\{n_i : i \in \mathbb{N}\}$ of natural numbers for which there exists $\lambda > 1$ such that $n_{i+1} \geq \lambda n_i$ for every i . We are now ready to state our main result; see Figure 1 for an illustration.

Theorem 1. *Fix $p \in (0, \frac{1}{2})$, take any lacunary sequence L for which there exists a sequence $\{S_n\}_{n \in L}$ of symmetric conference matrices of increasing size n , and consider the corresponding random matrices $X_n \sim \text{Sub}(S_n, p)$. Then the empirical spectral distribution of $\frac{1}{p\sqrt{n}}X_n$ converges weakly almost surely to the Kesten–McKay distribution with parameter $v = 1/p$.*

We note that Theorem 1 is the first proven instance of an assortment of conjectures posed by Haikin, Zamir and Gavish in [22]. In general, we expect random submatrices of the Gram matrix of an equiangular tight frame of n vectors in $\mathbb{C}^d$ to have empirical spectral distribution corresponding to the Wachter distribution [37] (also known as the MANOVA distribution) with parameters determined by both the sample probability p and the redundancy $\frac{n}{d}$. As partial progress toward proving these conjectures, Haikin, Zamir and Gavish [21] demonstrated that for any sequence of equiangular tight frames of any fixed redundancy, the first four moments converge in expectation to the corresponding Wachter moments. In her masters thesis [20], Haikin made additional combinatorial observations to facilitate an eventual proof by the method of moments. The Kesten–McKay distribution is a special case of the Wachter distribution, and as illustrated in [15], its moments are far less complicated. This is our reason for focusing on the redundancy-2 case, and we credit our success to the relative simplicity of these moments.

While the hypothesis $p < \frac{1}{2}$ in Theorem 1 may seem artificial at first glance, this constraint appears in the original Haikin–Zamir–Gavish conjectures [22], and for good reason. In terms of the equiangular tight frame $\Phi \in \mathbb{R}^{n/2 \times n}$ such that $\Phi^\top \Phi = I + \frac{1}{\sqrt{n-1}}S$, this restriction biases $\mathcal{I}$ to be typically smaller than the dimension $\frac{n}{2}$ so that the submatrix $\Phi_{\mathcal{I}}$ of columns indexed by $\mathcal{I}$ has a tall, skinny aspect ratio. When $|\mathcal{I}| > \frac{n}{2}$, then we expect a different spectral behavior since the submatrix $\Phi_{\mathcal{I}}$ now has a nontrivial nullspace. In this case of large $\mathcal{I}$, if we let $S_{\mathcal{I}}$ denote the submatrix of S with rows and columns indexed by $\mathcal{I}$, then it is straightforward to show that the eigenvalues of $S_{\mathcal{I}}$ are precisely the eigenvalues of $S_{\mathcal{I}^c}$, though with an additional $|\mathcal{I}| - \frac{n}{2}$ eigenvalues equal to each of $\pm\sqrt{n-1}$. Indeed, this follows from the facts that (i) the eigenvalues of $A^\top A$ are identical to the eigenvalues of $AA^\top$ modulo additional eigenvalues equal to 0, and (ii) $\Phi_{\mathcal{I}}\Phi_{\mathcal{I}}^\top + \Phi_{\mathcal{I}^c}\Phi_{\mathcal{I}^c}^\top = \Phi\Phi^\top = 2I_{n/2}$.

As an alternative method of drawing the random index subset $\mathcal{I}$, one might be inclined to fix k and draw $\mathcal{I}$ uniformly from the subsets of $[n]$ of order k . We suspect that the same convergence holds for this model by taking $k = \lfloor p/n \rfloor$, say; in fact, this might be provable from Theorem 1 by a perturbation argument. We instead focus on the model used

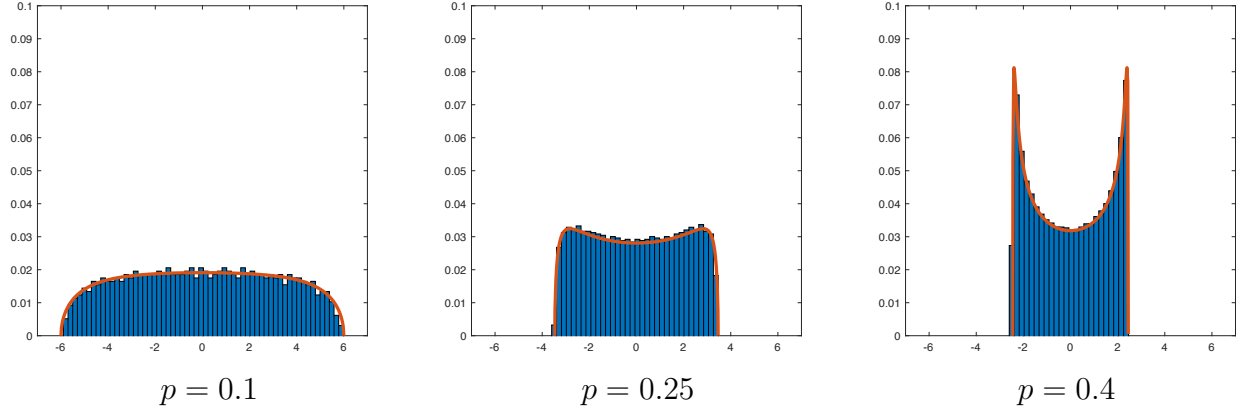


Figure 1: Consider the Paley conference matrix S of order $n = 10,010$. For each choice of $p \in \{0.1, 0.25, 0.4\}$, we draw $X \sim \text{Sub}(S, p)$ and plot a histogram of the spectrum of $\frac{1}{p\sqrt{n}}X$ along with a suitably scaled version of the Kesten–McKay density for $v = 1/p$. The similarity between these distributions was first observed by Haikin, Zamir and Gavish [22]. Our main result (Theorem 1) explains this phenomenon.

in the original Haikin–Zamir–Gavish conjectures, and our proof makes use of the fact that $\{i \in \mathcal{I}\}$ is independent of $\{j \in \mathcal{I}\}$ for $i \neq j$ in this model. One might also make progress on this problem by using a transform method instead of the method of moments. In fact, such techniques have proven effective in demonstrating that random submatrices of complex Hadamard matrices satisfy a Wachter law [36, 17, 1]. As an adjacent pursuit, there has been some work on constructing explicit matrices that satisfy Wigner’s semicircle law [31, 32], and it would be interesting to transfer these ideas to our setting.

In the next section, we prove Theorem 1 using the method of moments, saving the more technical portions for Section 3.

1.1 Notation

Given $x \in \mathbb{R}^n$, let $\text{diag}(x)$ denote the $n \times n$ diagonal matrix whose diagonal entries are the entries of x . Given $Z \in \mathbb{R}^{m \times n}$, let $\|Z\|_{2 \rightarrow 2}$ denote the induced 2-norm of Z (i.e., the largest singular value of Z), and let $\|Z\|_{S^p}$ denote the Schatten p -norm of Z (i.e., the p -norm of the singular values of Z). Throughout this paper, we will investigate how quantities relate as $n \rightarrow \infty$. For example, suppose we are interested in a quantity $f(n, \theta) \geq 0$ that depends on both $n \in \mathbb{N}$ and some additional parameters $\theta \in \mathbb{R}^m$. Then we write $f(n, \theta) = o(g(n, \theta))$ if for every $\theta \in \mathbb{R}^m$, it holds that $f(n, \theta)/g(n, \theta) \rightarrow 0$ as $n \rightarrow \infty$. We write $f(n, \theta) \lesssim g(n, \theta)$ if there exists $c > 0$ such that $f(n, \theta) \leq c \cdot g(n, \theta)$ for all $n \in \mathbb{N}$ and $\theta \in \mathbb{R}^m$, and we write $f(n, \theta) \lesssim_\theta g(n, \theta)$ if for every $\theta \in \mathbb{R}^m$, there exists $c(\theta) > 0$ such that $f(n, \theta) \leq c(\theta) \cdot g(n, \theta)$ for all $n \in \mathbb{N}$. Finally, we write $f(n, \theta) \asymp g(n, \theta)$ if both $f(n, \theta) \lesssim g(n, \theta)$ and $g(n, \theta) \lesssim f(n, \theta)$.

2 Proof of the main result

Our proof makes use of a standard sufficient condition for the almost sure weak convergence of random measures, namely, Proposition 2 below. We say that a probability measure μ is (a, b) -**subgaussian** if $\mu([-t, t]^c) \leq ae^{-bt^2}$ for every $t > 0$. A probability measure is called subgaussian if it is (a, b) -subgaussian for some $a, b > 0$. A sequence of probability measures is said to be **uniformly subgaussian** if there exists $a, b > 0$ for which every probability measure in the sequence is (a, b) -subgaussian.

Proposition 2. *Let $\{\zeta_i\}_{i=1}^\infty$ be a sequence of random probability measures that are almost surely uniformly subgaussian, and let μ be a non-random subgaussian probability measure. Suppose that for every $k \in \mathbb{N}$, it holds that*

$$(i) \quad \mathbb{E} \int_{\mathbb{R}} x^k d\zeta_i(x) \rightarrow \int_{\mathbb{R}} x^k d\mu(x), \text{ and}$$

$$(ii) \quad \sum_{i=1}^{\infty} \text{Var} \left(\int_{\mathbb{R}} x^k d\zeta_i(x) \right) < \infty.$$

Then ζ_i converges weakly almost surely to μ .

The proof of Proposition 2 is standard (cf. Exercise 2.4.6 in [34]): use Chebyshev's inequality and the Borel–Cantelli lemma to show $\int_{\mathbb{R}} x^k d\zeta_i(x) \rightarrow \int_{\mathbb{R}} x^k d\mu(x)$ almost surely for each $k \in \mathbb{N}$, and then apply the moment continuity theorem to points in the convergence event. As we will see, verifying hypothesis (i) in our case reduces to a combinatorics problem, whereas hypothesis (ii) can be treated separately with the help of Talagrand concentration:

Proposition 3 (Talagrand concentration, Theorem 2.1.13 in [34]). *There exists a universal constant $c > 0$ for which the following holds: Suppose $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is both convex and σ -Lipschitz in $\|\cdot\|_2$, and let X be a random vector in $\mathbb{R}^n$ with independent coordinates satisfying $|X_i| \leq b$ almost surely. Then for every $t \geq 0$, it holds that*

$$\mathbb{P}\{|f(X) - \mathbb{E}f(X)| \geq bt\} \lesssim e^{-t^2/c\sigma^2}.$$

Throughout, S_n denotes an $n \times n$ symmetric conference matrix, we draw $X_n \sim \text{Sub}(S_n, p)$ and put $Z_n := \frac{1}{p\sqrt{n}}X_n$. We typically suppress the subscript n . While the size of Z is random, its average size is pn , and so we use $\frac{1}{pn} \text{tr}(Z^k)$ as a proxy for $\int_{\mathbb{R}} x^k d\mu_Z(x)$. As one might expect, this is a good approximation:

Lemma 4. *Put $V := \frac{1}{pn} \text{tr}(Z^k)$ and $W := \int_{\mathbb{R}} x^k d\mu_Z(x)$. Then*

$$|\mathbb{E}V - \mathbb{E}W| \lesssim_p \frac{1}{\sqrt{n}}, \quad |\text{Var}(V) - \text{Var}(W)| \lesssim_p \frac{1}{\sqrt{n}}.$$

Proof. Since X is a submatrix of S , it holds that

$$|V| \lesssim_p \frac{1}{n} \sum_i |\lambda_i(Z)|^k \leq \|Z\|_{2 \rightarrow 2}^k \asymp_p \frac{1}{n^{k/2}} \|X\|_{2 \rightarrow 2}^k \leq \frac{1}{n^{k/2}} \|S\|_{2 \rightarrow 2}^k \leq 1,$$

almost surely. Similarly, $|W| \leq \|Z\|_{2 \rightarrow 2}^k \lesssim_p 1$ almost surely. Next, let N denote the (random) size of Z . Then $V = \frac{N}{pn} \cdot W$, and so our bound on $|W|$ gives

$$\mathbb{E}|V - W| = \mathbb{E}\left(\left|\frac{N}{pn} - 1\right| \cdot |W|\right) \lesssim_p \mathbb{E}\left|\frac{N}{pn} - 1\right| \leq \frac{1}{pn} (\mathbb{E}(N - pn)^2)^{1/2} \lesssim_p \frac{1}{\sqrt{n}},$$

where the last step applies the fact that N has binomial distribution. This immediately implies the desired bound on $|\mathbb{E}V - \mathbb{E}W|$. Finally, since $|V|, |W| \lesssim_p 1$ almost surely, we have

$$\begin{aligned} |\text{Var}(V) - \text{Var}(W)| &\leq |\mathbb{E}V^2 - \mathbb{E}W^2| + |(\mathbb{E}V)^2 - (\mathbb{E}W)^2| \\ &\leq \mathbb{E}(|V + W||V - W|) + |\mathbb{E}V + \mathbb{E}W||\mathbb{E}V - \mathbb{E}W| \\ &\lesssim_p \mathbb{E}|V - W| \lesssim_p \frac{1}{\sqrt{n}}, \end{aligned}$$

which completes the result. $\square$

As such, to demonstrate hypothesis (i) from Proposition 2 in our case, it suffices to prove

$$\mathbb{E} \frac{1}{pn} \text{tr}(Z^k) \rightarrow \int_{\mathbb{R}} x^k d\mu_{\text{KM}(1/p)}(x). \quad (2)$$

The Kesten–McKay moments are implicitly computed in [26], and are naturally expressed in terms of entries of **Catalan’s triangle**:

$$C(n, k) := \frac{(n+k)!(n-k+1)}{k!(n+1)!}.$$

Proposition 5 (Lemma 2.1 in [26]). *For every $v \geq 2$ and $k \in \mathbb{N}$, it holds that*

$$\int_{\mathbb{R}} x^k d\mu_{\text{KM}(v)}(x) = \begin{cases} \sum_{j=1}^{k/2} C(k/2 - 1, k/2 - j) v^j (v-1)^{k/2-j}, & \text{if } k \text{ is even} \\ 0, & \text{if } k \text{ is odd.} \end{cases}$$

Recalling that $Z = \frac{1}{p\sqrt{n}}X$, then Proposition 5 gives that (2) is equivalent to

$$\frac{1}{n^{k/2+1}} \mathbb{E} \text{tr}(X^k) \rightarrow \begin{cases} \sum_{t=k/2+1}^k (-1)^{t-k/2-1} \cdot B(k/2 - 1, t - k/2 - 1) \cdot p^t, & \text{if } k \text{ is even} \\ 0, & \text{if } k \text{ is odd,} \end{cases} \quad (3)$$

where $B(n, k)$ denotes an entry of **Borel’s triangle**:

$$B(n, k) := \sum_{j=k}^n \binom{j}{k} C(n, j).$$

To compute these limits, we first find a convenient expression for $\frac{1}{n^{k/2+1}} \mathbb{E} \operatorname{tr}(X^k)$. To this end, recall that X is the submatrix of S with index set $\mathcal{I}$, and let P denote the random $n \times n$ diagonal matrix such that $P_{ii} = 1_{\{i \in \mathcal{I}\}}$. Then

$$\begin{aligned} \operatorname{tr}(X^k) &= \operatorname{tr}((PSP)^k) = \operatorname{tr}((PS)^k) = \sum_{a_1, \dots, a_k \in [n]} (PS)_{a_1 a_2} (PS)_{a_2 a_3} \cdots (PS)_{a_k a_1} \\ &= \sum_{a_1, \dots, a_k \in [n]} S_{a_1 a_2} S_{a_2 a_3} \cdots S_{a_k a_1} \cdot \prod_{i=1}^k 1_{\{a_i \in \mathcal{I}\}}. \end{aligned}$$

Considering $\mathbb{E} \prod_{i=1}^k 1_{\{a_i \in \mathcal{I}\}} = p^{|\{a_1, \dots, a_k\}|}$, it follows that

$$\frac{1}{n^{k/2+1}} \mathbb{E} \operatorname{tr}(X^k) = \sum_{t=1}^k \left(\frac{1}{n^{k/2+1}} \sum_{\substack{a_1, \dots, a_k \in [n] \\ |\{a_1, \dots, a_k\}|=t}} S_{a_1 a_2} S_{a_2 a_3} \cdots S_{a_k a_1} \right) \cdot p^t. \quad (4)$$

It remains to show that these coefficients converge to the corresponding coefficients in (3).

First, we introduce some additional notation. Taking inspiration from Bargmann invariants [4], it is convenient to write

$$\Delta(a_1, a_2, a_3, \dots, a_k) := S_{a_1 a_2} S_{a_2 a_3} \cdots S_{a_k a_1}.$$

Next, we say π is a **partition** of $[k]$ into t blocks if $\pi = \{B_1, \dots, B_t\}$ such that $B_1 \sqcup \cdots \sqcup B_t = [k]$, and we let $\Pi(k, t)$ denote the set of all such partitions. For each partition π of $[k]$, we consider the set of functions $a: [k] \rightarrow [n]$ whose level sets are the blocks of π , namely

$$L_n(\pi) := \{a: [k] \rightarrow [n] : \{a^{-1}(a(i)) : i \in [k]\} = \pi\}.$$

With this, we define

$$V_n(\pi) := \frac{1}{n^{k/2+1}} \sum_{a \in L_n(\pi)} \Delta(a(1), \dots, a(k)).$$

Considering (4), it therefore holds that

$$\frac{1}{n^{k/2+1}} \mathbb{E} \operatorname{tr}(X^k) = \sum_{t=1}^k \left(\sum_{\pi \in \Pi(k, t)} V_n(\pi) \right) \cdot p^t. \quad (5)$$

As such, to demonstrate (3), it suffices to determine the limit of $V_n(\pi)$ for every partition π of $[k]$. We start with a quick calculation:

Lemma 6. *For every $\pi \in \Pi(k, t)$ with $t < k/2 + 1$, it holds that $V_n(\pi) \rightarrow 0$.*

Proof. Estimate $|V_n(\pi)|$ using the triangle inequality to obtain a sum of $|L_n(\pi)| \leq n^t = o(n^{k/2+1})$ terms, each of size at most 1. $\square$

For each $t < k/2 + 1$, this establishes that the coefficient of p^t in (5) approaches zero, i.e., the corresponding coefficient in (3). Now we wish to tackle the limiting value of $V_n(\pi)$ in general. In light of the related literature [29], it comes as no surprise that $V_n(\pi)$ depends on whether π is a so-called *crossing partition*. We say a partition π of $[k]$ is **crossing** if there exist $A, B \in \pi$ with $A \neq B$ for which there exist $a_1, a_2 \in A$ and $b_1, b_2 \in B$ such that $a_1 < b_1 < a_2 < b_2$. Otherwise, π is said to be **non-crossing**. Next, for each $x \in [k]$, we let $\pi(x)$ denote the unique member of π such that $x \in \pi(x)$. Consider the graph G_π with vertex set π and edges given by $\pi(x) \leftrightarrow \pi(x+1)$ for every $x \in [k]$; here, we interpret $x+1$ modulo k so that $k+1 = 1$. Let $\text{EC}(k, t)$ denote the set of non-crossing $\pi \in \Pi(k, t)$ for which the edges of G_π partition into simple even cycles. See Figure 2 for an example of such a partition π and its graph G_π . Finally, let $C_n := \frac{1}{n+1} \binom{2n}{n}$ denote the n th **Catalan number**. With these notions, we can describe the limit of each $V_n(\pi)$:

Lemma 7 (Key combinatorial lemma).

- (i) Suppose $\pi \in \Pi(k, t) \setminus \text{EC}(k, t)$. Then $V_n(\pi) \rightarrow 0$.
- (ii) Suppose $\pi \in \text{EC}(k, t)$ and the edges of G_π partition into m simple cycles of sizes $2s_1, \dots, 2s_m$. Then $m = k - t + 1$ and

$$V_n(\pi) \rightarrow (-1)^{k/2-m} \cdot C_{s_1-1} \cdots C_{s_m-1}.$$

The proof of Lemma 7 is rather technical (involving multiple rounds of induction), and so we save it for Section 3. (Curiously, a reviewer of this paper noted that these signed products of Catalan numbers are the values of the Möbius function of the non-crossing partition lattice [29], and this observation may eventually lead to a more bijective combinatorial proof of Lemma 7.) In the meantime, we demonstrate how Lemma 7 can be applied to prove that the coefficients in (5) converge to the coefficients in (3). Recall that a **Dyck path** of semi-length n is a path in the plane from $(0, 0)$ to $(2n, 0)$ consisting of n steps along the vector $(1, 1)$, called **up-steps**, and n steps along the vector $(1, -1)$, called **down-steps**, that never goes below the x -axis. We say a Dyck path is **strict** if none of the path's interior vertices reside on the x -axis. Each (strict) Dyck path determines a sequence of $2n$ letters from $\{U, D\}$ that represent up- and down-steps in the path; this sequence is known as a **(strict) Dyck word**. With these notions, we may prove the following result by leveraging the fact that Borel's triangle counts so-called *marked Dyck paths* [7]; see Figure 2 for an illustration.

Lemma 8. *It holds that*

$$\sum_{\pi \in \Pi(k, t)} V_n(\pi) \rightarrow \begin{cases} (-1)^{t-k/2-1} \cdot B(k/2-1, t-k/2-1), & \text{if } k \text{ is even and } t \geq k/2+1 \\ 0, & \text{otherwise.} \end{cases}$$

Proof. When $t < k/2 + 1$, the result follows from Lemma 6, and when k is odd, the k edges in each G_π fail to partition into even simple cycles, and so the result follows from Lemma 7(i). Now suppose k is even and $t \geq k/2 + 1$. For $\pi \in \text{EC}(k, t)$, recall that the edges of G_π are indexed by $[k]$ and partitioned into simple even cycles. Define $\text{MD}(\pi)$ to be the words $w: [k] \rightarrow \{U, U', D\}$ such that for every simple cycle in G_π with edges indexed by $T \subseteq [k]$,

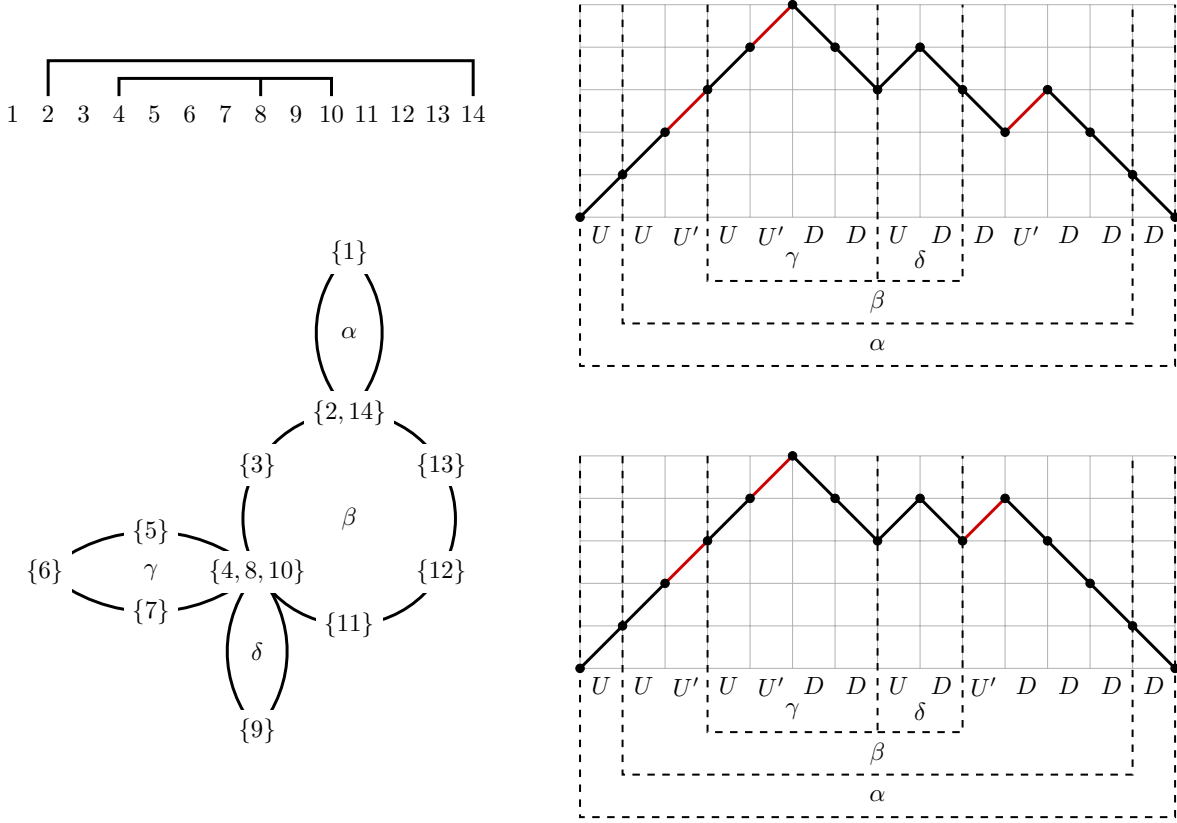


Figure 2: **(top left)** Select $k = 14$ and $t = 11$, and consider the partition $\pi \in \Pi(k, t)$ with all singleton blocks except for $\{2, 14\}$ and $\{4, 8, 10\}$. Observe that π is a non-crossing partition. **(bottom left)** We depict the corresponding graph G_π , whose vertices are the blocks of π . By definition, blocks are adjacent in G_π when they contain cyclicly adjacent members of $[k]$. In this case, the edges of G_π partition into four simple cycles, which we label α , β , γ and δ . **(right)** Each simple cycle of G_π is assigned a strict Dyck word of the cycle's length, and we mark all but the first up-steps. The only choice for α and δ is UD , and the only choice for γ is $UU'DD$; here, U' denotes a marked up-step. Meanwhile, β has $C_2 = 2$ choices: $UU'DU'DD$ and $UU'U'DDD$. For each selection, we traverse G_π from $\pi(1)$ to $\pi(2)$, to $\pi(3)$, etc., to $\pi(14)$ and back to $\pi(1)$, labelling the edges of G_π with the next letter from the current cycle's Dyck word. The result is a Dyck word with $t - k/2 - 1 = 3$ marked up-steps, none of which at ground level. We illustrate the corresponding marked Dyck paths above. Notice that G_π can be recovered from either marked Dyck path since a cycle is born with each un-marked up-step and dies once the Dyck path returns to its height from the birth of that cycle. By Theorem 2 in [7], marked Dyck paths are counted by entries in Borel's triangle, which explains their appearance in (3).

the restriction $w|_T$ is a strict Dyck word with all but its first up-steps marked (here, U' denotes a marked up-step). Note that strict Dyck words of semi-length s are in one-to-one correspondence with Dyck words of semi-length $s - 1$, and so there are C_{s-1} of them. As such, Lemma 7 implies that for every $\pi \in \Pi(k, t)$, it holds that

$$(-1)^{t-k/2-1} \cdot V_n(\pi) \rightarrow \begin{cases} |\text{MD}(\pi)|, & \text{if } \pi \in \text{EC}(k, t) \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

Let $\text{MD}(k, t)$ denote the set of marked Dyck words $w: [k] \rightarrow \{U, U', D\}$ with $t - k/2 - 1$ marked up-steps, none of which are at ground level. We observe that

$$\text{MD}(k, t) = \bigsqcup_{\pi \in \text{EC}(k, t)} \text{MD}(\pi). \quad (7)$$

Then equations (6) and (7) together give

$$(-1)^{t-k/2-1} \sum_{\pi \in \Pi(k, t)} V_n(\pi) \rightarrow \sum_{\pi \in \text{EC}(k, t)} |\text{MD}(\pi)| = |\text{MD}(k, t)| = B(k/2 - 1, t - k/2 - 1),$$

where the last step applies Theorem 2 in [7]. $\square$

At this point, we are in a position to verify hypothesis (i) from Proposition 2 in our case. For hypothesis (ii), we follow the approach suggested by Remark 2.4.5 in [34] of leveraging Talagrand concentration to bound the variance. First, we pass to a setting that is more amenable to analysis with Talagrand concentration. Here and throughout, for each $n \in L$, we fix an $n \times n$ matrix F such that $F^\top F = I + \frac{1}{\sqrt{n}} S_n$.

Lemma 9. *It holds that $\text{Var}(\text{tr}((\frac{1}{p\sqrt{n}}X)^k)) \lesssim_{p,k} \max_{j \in [k]} \text{Var}(\|FP\|_{S^{2j}}^{2j}) + n^{1/2}$.*

Proof. Define $Y := \frac{1}{p\sqrt{n}}PSP$, and observe that

$$\text{tr}(Y^k) = \text{tr}(\frac{1}{p\sqrt{n}}X^k), \quad \text{tr}((Y + \frac{1}{p}P)^k) = \text{tr}((\frac{1}{p}PF^\top FP)^k) = \frac{1}{p^k} \|FP\|_{S^{2k}}^{2k}.$$

Since Y commutes with P and $YP = Y$, the binomial theorem gives

$$\text{tr}((Y + \frac{1}{p}P)^k) = \text{tr} \sum_{j=0}^k \binom{k}{j} Y^j (\frac{1}{p}P)^{k-j} = \sum_{j=1}^k \binom{k}{j} \frac{1}{p^{k-j}} \text{tr}(Y^j) + \frac{1}{p^k} \text{tr}(P),$$

and so rearranging gives

$$\begin{aligned} \text{tr}(\frac{1}{p\sqrt{n}}X^k) &= \text{tr}(Y^k) = \text{tr}((Y + \frac{1}{p}P)^k) - \sum_{j=1}^{k-1} \binom{k}{j} \frac{1}{p^{k-j}} \text{tr}(Y^j) - \frac{1}{p^k} \text{tr}(P) \\ &= \frac{1}{p^k} \|FP\|_{S^k}^{2k} - \sum_{j=1}^{k-1} \binom{k}{j} \frac{1}{p^{k-j}} \text{tr}(\frac{1}{p\sqrt{n}}X^j) - \frac{1}{p^k} \text{tr}(P). \end{aligned} \quad (8)$$

The following estimate holds for any choice of random variables $\{X_i\}_{i \in [m]}$:

$$\begin{aligned} \text{Var} \left(\sum_{i=1}^m X_i \right) &= \sum_{i=1}^m \sum_{j=1}^m \text{Cov}(X_i, X_j) \leq \sum_{i=1}^m \sum_{j=1}^m |\text{Cov}(X_i, X_j)| \\ &\leq \sum_{i=1}^m \sum_{j=1}^m \sqrt{\text{Var}(X_i) \text{Var}(X_j)} \leq m^2 \cdot \max_{i \in [m]} \text{Var}(X_i). \end{aligned}$$

The lemma follows from applying this estimate to (8) by induction on k . $\square$

Next, we establish the convexity and Lipschitz continuity required by Talagrand:

Lemma 10. *For each $k \in \mathbb{N}$, consider the mapping $f: \{x \in \mathbb{R}^n : \|x\|_\infty < 2\} \rightarrow \mathbb{R}$ defined by $f(x) = \|F \text{diag}(x)\|_{S^{2k}}^{2k}$. Then f is convex and $(8^k k n^{1-1/2k})$ -Lipschitz.*

Proof. We adopt the shorthand notation $D_x := \text{diag}(x)$. First, f is convex since $\|\cdot\|_{S^{2k}}$ satisfies the triangle inequality and $t \mapsto t^{2k}$ is convex:

$$f(\lambda x + (1-\lambda)y) \leq \left(\lambda \|FD_x\|_{S^{2k}} + (1-\lambda) \|FD_y\|_{S^{2k}} \right)^{2k} \leq \lambda f(x) + (1-\lambda)f(y).$$

To compute a Lipschitz bound, we apply the factorization

$$u^{2k} - v^{2k} = (u-v)(u+v) \sum_{j=0}^{k-1} u^{2(k-1-j)} v^{2j},$$

with $u := \|FD_x\|_{S^{2k}}$ and $v := \|FD_y\|_{S^{2k}}$ to get

$$|f(x) - f(y)| = |u^{2k} - v^{2k}| = \left((u+v) \sum_{j=0}^{k-1} u^{2(k-1-j)} v^{2j} \right) \cdot |u-v| \leq 8^k k n^{1-1/2k} \cdot |u-v|,$$

where the last step follows from the fact that $\|FD_x\|_{2 \rightarrow 2} \leq 2\|F\|_{2 \rightarrow 2} \leq 2\sqrt{2}$, meaning $u \leq 2\sqrt{2}n^{1/2k}$ (and similarly for v). Next, we apply the reverse triangle inequality to get

$$\begin{aligned} |u-v| &= \left| \|FD_x\|_{S^{2k}} - \|FD_y\|_{S^{2k}} \right| \\ &\leq \|FD_x - FD_y\|_{S^{2k}} \leq \|F(D_x - D_y)\|_F \leq \|F\|_{2 \rightarrow 2} \cdot \|x - y\|_2 \leq \sqrt{2} \cdot \|x - y\|_2, \end{aligned}$$

which implies the result. $\square$

Finally, we apply Talagrand concentration to obtain a variance bound:

Lemma 11. *It holds that $\text{Var} \left(\frac{1}{pn} \text{tr} \left(\left(\frac{1}{p\sqrt{n}} X \right)^k \right) \right) \lesssim_{p,k} n^{-1/k}$.*

Proof. Given the mapping f from Lemma 10, define $\tilde{f}: \mathbb{R}^n \rightarrow \mathbb{R}$ in terms of subgradients by

$$\tilde{f}(x) := \sup_{x_0 \in R} \sup_{z \in \partial f(x_0)} \left(f(x_0) + \langle z, x - x_0 \rangle \right).$$

This is known as the smallest convex extension of f to $\mathbb{R}^n$, and it is straightforward to verify that $\tilde{f}$ is convex and $(8^k k n^{1-1/2k})$ -Lipschitz with $\tilde{f}|_R = f$. Let $B \in \mathbb{R}^n$ have independent entries, each equal to 1 with probability p and 0 otherwise. Since $B \in R$ almost surely, it holds that $\tilde{f}(B)$ has the same distribution as $\|FP\|_{S^{2k}}^{2k}$, and we let E denote its expectation. By Talagrand concentration (Proposition 3), there exists $c > 0$ such that

$$\begin{aligned} \text{Var}(\|FP\|_{S^{2k}}^{2k}) &= \mathbb{E}\left[(\|FP\|_{S^{2k}}^{2k} - E)^2\right] = \int_0^\infty \mathbb{P}\left\{(\|FP\|_{S^{2k}}^{2k} - E)^2 \geq u\right\} du \\ &= \int_0^\infty \mathbb{P}\left\{|\tilde{f}(B) - E| \geq \sqrt{u}\right\} du \\ &\lesssim \int_0^\infty \exp\left(\frac{-u}{c \cdot 8^{2k} k^2 n^{2-1/k}}\right) du = c \cdot 8^{2k} k^2 n^{2-1/k}. \end{aligned}$$

Combining with Lemma 9 then gives

$$\begin{aligned} \text{Var}\left(\frac{1}{pn} \text{tr}\left(\left(\frac{1}{p\sqrt{n}}X\right)^k\right)\right) &= \frac{1}{p^2 n^2} \text{Var}\left(\text{tr}\left(\left(\frac{1}{p\sqrt{n}}X\right)^k\right)\right) \\ &\lesssim_{p,k} \frac{1}{n^2} \left(\max_{j \in [k]} \text{Var}(\|FP\|_{S^{2j}}^{2j}) + n^{1/2}\right) \lesssim_k n^{-1/k}, \end{aligned}$$

as desired. $\square$

We may now verify hypotheses (i) and (ii) from Proposition 2 in our case.

Proof of Theorem 1. Put $Z_n := \frac{1}{p\sqrt{n}}X_n$ and $\mu := \mu_{\text{KM}(1/p)}$. First, we modify the random measure μ_{Z_n} so that we may apply Proposition 2 to prove the result. Indeed, μ_{Z_n} fails to be a probability measure with probability $(1-p)^n$, since $\mu_{Z_n} = 0$ when $\mathcal{I} = \mathcal{I}_n$ is the empty set. To rectify this, we define

$$\zeta_n := \begin{cases} \mu_{Z_n}, & \text{if } \mathcal{I}_n \neq \emptyset \\ \delta_0, & \text{otherwise.} \end{cases}$$

Then it suffices to prove $\zeta_n \rightarrow \mu$ weakly almost surely, since the Borel–Cantelli lemma implies $1_{\{\mathcal{I}_n = \emptyset\}} \rightarrow 0$ almost surely, and so for every bounded continuous $f: \mathbb{R} \rightarrow \mathbb{R}$, it holds that

$$\int_{\mathbb{R}} f(x) d\mu_{Z_n}(x) = \int_{\mathbb{R}} f(x) d\zeta_n(x) - 1_{\{\mathcal{I}_n = \emptyset\}} \cdot f(0) \xrightarrow{\text{a.s.}} \int_{\mathbb{R}} f(x) d\mu(x).$$

Conveniently, for every $n \in L$ and $k \in \mathbb{N}$, it holds that

$$\int_{\mathbb{R}} x^k d\zeta_n(x) = \int_{\mathbb{R}} x^k d\mu_{Z_n}(x),$$

almost surely, and so the left-hand side inherits moments from the right-hand side.

To apply Proposition 2, we first observe that

$$\|Z_n\|_{2 \rightarrow 2} = \frac{1}{p\sqrt{n}} \|X_n\|_{2 \rightarrow 2} \leq \frac{1}{p\sqrt{n}} \|S_n\|_{2 \rightarrow 2} \leq \frac{1}{p},$$

almost surely, and so $\{\zeta_n\}_{n \in L}$ are almost surely uniformly bounded, and therefore almost surely uniformly subgaussian. Similarly, μ is bounded and therefore subgaussian. Fix $k \in \mathbb{N}$. As a consequence of Lemma 8, it holds that

$$\mathbb{E} \frac{1}{pn} \operatorname{tr}(Z_n^k) \rightarrow \int_{\mathbb{R}} x^k d\mu(x),$$

and so by Lemma 4, we have

$$\mathbb{E} \int_{\mathbb{R}} x^k d\zeta_n(x) = \mathbb{E} \int_{\mathbb{R}} x^k d\mu_{Z_n}(x) \rightarrow \int_{\mathbb{R}} x^k d\mu(x).$$

As such, $\{\zeta_n\}_{n \in L}$ satisfies hypothesis (i) from Proposition 2. Next, Lemma 11 establishes that $\operatorname{Var} \left(\frac{1}{pn} \operatorname{tr}(Z_n^k) \right) \lesssim_{p,k} n^{-1/k}$, and so Lemma 4 implies

$$\operatorname{Var} \left(\int_{\mathbb{R}} x^k d\zeta_n(x) \right) = \operatorname{Var} \left(\int_{\mathbb{R}} x^k d\mu_{Z_n}(x) \right) \lesssim_{p,k} n^{-1/k} + n^{-1/2} \lesssim n^{-1/(k+1)}.$$

Writing $L = \{n_i : i \in \mathbb{N}\}$, select $\lambda > 1$ such that $n_{i+1} \geq \lambda n_i$ for every $i \in \mathbb{N}$. Then

$$\begin{aligned} \sum_{n \in L} \operatorname{Var} \left(\int_{\mathbb{R}} x^k d\zeta_n(x) \right) &\lesssim_{p,k} \sum_{n \in L} n^{-1/(k+1)} \\ &\leq \sum_{i=0}^{\infty} (\lambda^i n_1)^{-1/(k+1)} = n_1^{-1/(k+1)} \sum_{i=0}^{\infty} (\lambda^{-1/(k+1)})^i < \infty. \end{aligned}$$

As such, $\{\zeta_n\}_{n \in L}$ also satisfies hypothesis (ii) from Proposition 2, and so $\zeta_n \rightarrow \mu$ weakly almost surely, as desired. $\square$

3 Proof of Lemma 7

It remains to compute, for each $\pi \in \Pi(k, t)$, the limit of

$$V_n(\pi) = \frac{1}{n^{k/2+1}} \sum_{a \in L_n(\pi)} \Delta(a(1), \dots, a(k)),$$

where $L_n(\pi)$ is the set of $a: [k] \rightarrow [n]$ whose level sets are the blocks of π and

$$\Delta(a(1), \dots, a(k)) = S_{a(1)a(2)} S_{a(2)a(3)} \cdots S_{a(k)a(1)}.$$

We begin with some basic properties of Δ .

Lemma 12. *For every $a_1, \dots, a_k \in [n]$, each of the following holds:*

- (i) *If $a_1 \neq a_2$, then $\Delta(a_1, a_2) = 1$.*
- (ii) *If $a_j = a_{j+1}$ for any $j \in [k-1]$ or $a_k = a_1$, then $\Delta(a_1, \dots, a_k) = 0$.*
- (iii) *If σ is any cyclic permutation of $[k]$, then $\Delta(a_{\sigma(1)}, \dots, a_{\sigma(k)}) = \Delta(a_1, \dots, a_k)$.*

(iv) If $a_1 \neq a_{k-1}$, then $\sum_{b \in [n]} \Delta(a_1, \dots, a_{k-1}, b) = 0$.

(v) If $a_1 = a_{k-1}$ and $a_1 \neq a_k$, then $\Delta(a_1, \dots, a_k) = \Delta(a_1, \dots, a_{k-2})$.

Proof. First, (i) follows from the fact that S is symmetric with off-diagonal entries in $\{\pm 1\}$. Next, (ii) follows from the fact that the diagonal entries of S are 0. Recalling the definition of Δ , then (iii) follows from commutativity. Next suppose $a_{k-1} \neq a_1$. Then

$$\sum_{b \in [n]} \Delta(a_1, \dots, a_{k-1}, b) = S_{a_1 a_2} \cdots S_{a_{k-2} a_{k-1}} \sum_{b \in [n]} S_{a_{k-1} b} S_{b a_1},$$

and (iv) follows since $\sum_{b \in [n]} S_{a_{k-1} b} S_{b a_1}$ is the (a_{k-1}, a_1) entry of $S^2 = (n-1)I$. Finally, in the case where $a_1 = a_{k-1}$, we have

$$\Delta(a_1, \dots, a_k) = S_{a_1 a_2} \cdots S_{a_{k-2} a_1} S_{a_1 a_k} S_{a_k a_1},$$

and (v) follows since $S_{a_1 a_k} S_{a_k a_1} = 1$ provided $a_1 \neq a_k$. $\square$

Let π be a partition of $[k]$. Recall that for $j \in [k]$, we let $\pi(j)$ denote the block of π containing j . We extend this notation to any integer j by considering $\pi(j)$ to be the block of π containing a representative of j modulo k . For convenience, we record the following immediate consequence of Lemma 12(iii).

Lemma 13. *Let π be a partition of $[k]$ and fix $j \in \mathbb{Z}$. Define π' to be the partition of $[k]$ with $\pi'(i) = \pi(i - j)$ for all $i \in [k]$. Then $V_n(\pi') = V_n(\pi)$.*

To establish Lemma 7(i), we will show separately that $V_n(\pi) \rightarrow 0$ for every crossing partition $\pi \in \Pi(k, t)$ and that $V_n(\pi) \rightarrow 0$ for every non-crossing partition $\pi \in \Pi(k, t)$ such that G_π contains an odd cycle.

Lemma 14. *Let $\pi \in \Pi(k, t)$ be a crossing partition. Then $V_n(\pi) \rightarrow 0$.*

Proof. For $\pi \in \Pi(k, t)$ to be a crossing partition, it must hold that $t \geq 2$ and $k \geq 4$. Observe that the case $t = 2$ follows immediately from Lemma 6 since $k \geq 4$. Now consider $t > 2$ and suppose the lemma has been established for every crossing partition on $t-1$ blocks. By Lemma 6, we may further suppose that k satisfies $t \geq k/2 + 1$. Then for $\pi \in \Pi(k, t)$, the pigeonhole principle guarantees that π contains a singleton block $\{j\} \in \pi$. By Lemma 13, we may assume $\{k\} \in \pi$. We proceed in cases:

Case I: $\pi(1) = \pi(k-1)$. We may apply Lemma 12(v) to obtain

$$\begin{aligned} V_n(\pi) &= \frac{1}{n^{k/2+1}} \sum_{a \in L_n(\pi)} \Delta(a(1), \dots, a(k-2), a(1), a(k)) \\ &= \frac{1}{n^{k/2}} \sum_{a \in L_n(\pi \setminus \{k\})} \Delta(a(1), \dots, a(k-2)) + o(1). \end{aligned}$$

The restriction of $\pi \setminus \{k\}$ to $[k-2]$ results in a crossing partition π' of $[k-2]$ into $t-1$ blocks. Moreover, the above expression for $V_n(\pi)$ implies

$$V_n(\pi) = V_n(\pi') + o(1),$$

and so our induction hypothesis provides $V_n(\pi) \rightarrow 0$.

Case II: $\pi(1) \neq \pi(k-1)$. Writing out $\pi = \{B_1, \dots, B_{t-1}, \{k\}\}$, we choose representatives $j_1, \dots, j_{t-1} \in [k-1]$ with $\pi(j_i) = B_i$. Then by Lemma 12(iv), we have

$$\begin{aligned} V_n(\pi) &= \frac{1}{n^{k/2+1}} \sum_{a \in L_n(\pi \setminus \{k\})} \sum_{\substack{a_k \in [n] \\ a_k \notin a([k-1])}} \Delta(a(1), \dots, a(k-1), a_k) \\ &= -\frac{1}{n^{k/2+1}} \sum_{a \in L_n(\pi \setminus \{k\})} \sum_{i=1}^{t-1} \Delta(a(1), \dots, a(k-1), a(j_i)). \end{aligned}$$

For $i, j \in [t-1]$, we define new blocks

$$B_j^i := \begin{cases} B_j \cup \{k\}, & \text{if } j = i \\ B_j, & \text{if } j \neq i, \end{cases}$$

and the corresponding crossing partition $\pi^i = \{B_1^i, \dots, B_{t-1}^i\} \in \Pi(k, t-1)$. Then

$$V_n(\pi) = -\sum_{i=1}^{t-1} \frac{1}{n^{k/2+1}} \sum_{a \in L_n(\pi \setminus \{k\})} \Delta(a(1), \dots, a(k-1), a(j_i)) = -\sum_{i=1}^{t-1} V_n(\pi^i).$$

Since each $V_n(\pi^i) \rightarrow 0$ by our induction hypothesis, we see $V_n(\pi) \rightarrow 0$ as well. $\square$

For non-crossing partitions, we will study the structure of the graph G_π for $\pi \in \Pi(k, t)$, which we recall has vertex set π and edges $\pi(j) \leftrightarrow \pi(j+1)$ for all $j \in [k]$. Observe that if G_π has a loop, then $\pi(j) = \pi(j+1)$ for some $j \in [k]$, and so $V_n(\pi) = 0$ by Lemma 12(ii). For this reason, we direct our attention to **loop-free** partitions π , that is, partitions π for which G_π is loop-free.

Given a loop-free graph G on vertices V with edges E , we say $v \in V$ is a **cut vertex** if the induced subgraph of G on $V \setminus \{v\}$ is disconnected. A graph with no cut vertices is called **biconnected**, and the **biconnected components** of a graph are its maximal biconnected subgraphs. When the biconnected components of G are all simple cycles, we call G a **cactus**. A reviewer of this paper noted that cactus graphs also emerge in the context of *traffic freeness* [11], but at the time of this writing, we have not fully investigated how deep this connection is.

Lemma 15. *If $\pi \in \Pi(k, t)$ is a loop-free non-crossing partition, then $t \geq k/2 + 1$ and G_π is a cactus whose edges partition into $k - t + 1$ simple cycles.*

Proof. First, suppose G_π is a cactus whose edges partition into $k - t + 1$ simple cycles. Since G_π has no loops, the number of cycles is at least at least half the number of edges, that is, $k - t + 1 \geq k/2$. Rearranging then gives $t \geq k/2 + 1$. It remains to verify that G_π is, indeed, a cactus whose edges partition into $k - t + 1$ simple cycles.

Fixing k , we proceed by induction on $k - t$. If $\pi \in \Pi(k, k)$, then G_π is itself a simple cycle and hence a cactus. For $k - t > 0$, we now consider a loop-free non-crossing partition $\pi \in \Pi(k, t)$. By the pigeonhole principle, we may select $B \in \pi$ such that B contains at least

two elements of $[k]$. Let j denote the least element of B . Writing $B = \{j\} \sqcup B'$, we consider $\pi' \in \Pi(k, t+1)$ defined by $\pi' = (\pi \setminus B) \cup \{B', \{j\}\}$. Since π' is also loop-free and non-crossing, our induction hypothesis guarantees that $G_{\pi'}$ is a cactus with $k - t$ simple cycles. Our task is to use this information to show that G_{π} is a cactus with $k - t + 1$ simple cycles.

Suppose first that $\{j\}$ and B' reside in the same simple cycle of $G_{\pi'}$. Then the simple cycles of $G_{\pi'}$ not containing $\{j\}$ and B' remain simple cycles and biconnected components of G_{π} . Moreover, by identifying $\{j\}$ and B' , we see that the simple cycle of $G_{\pi'}$ containing $\{j\}$ and B' corresponds to two simple cycles of G_{π} sharing the cut vertex $B = \{j\} \sqcup B'$. As such, G_{π} contains $k - t + 1$ biconnected components, each of which is a simple cycle.

We now claim that $\{j\}$ and B' must reside in the same simple cycle of $G_{\pi'}$, in which case we are done by the previous paragraph. Suppose instead that there exists a cut vertex $X \in \pi' \setminus \{\{j\}, B'\}$ that separates $\{j\}$ and B' within $G_{\pi'}$, and select $j' \in [k]$ with $\pi(j') = B'$. Since j is the least element of B , we necessarily have $j < j'$. Furthermore, since X separates $\{j\}$ and B' , we can traverse along a trail in $G_{\pi'}$ from $\{j\}$ to X , to B' , and back to X to obtain indices $i \in (j, j')$ and $i' \in (j', k + j)$ with $\pi'(i) = \pi'(i') = X$. These indices $j < i < j' < i'$ satisfy $\pi(j) = \pi(j')$ and $\pi(i) = \pi(i')$, contradicting our assumption that π is non-crossing. Hence, $\{j\}$ and B' must reside in the same simple cycle of $G_{\pi'}$ as claimed. $\square$

Lemma 16. *For every loop-free non-crossing $\pi \in \Pi(k, t)$, each of the following holds:*

(i) *If G_{π} contains any odd cycles, then $V_n(\pi) \rightarrow 0$.*

(ii) *If the edges of G_{π} partition into m simple cycles of sizes $2s_1, \dots, 2s_m$, then*

$$V_n(\pi) \rightarrow (-1)^{k/2-m} \cdot C_{s_1-1} \cdots C_{s_m-1}.$$

Proof. Any loop-free non-crossing partition must have at least two blocks. When $t = 2$, we may assume $k = 2$ by Lemma 6 so that the only partition under consideration is $\{\{1\}, \{2\}\}$, in which case $m = 1$ and $s_1 = 1$. Lemma 12(i) allows us to verify the result in this case:

$$V_n(\{\{1\}, \{2\}\}) = \frac{1}{n^2} \sum_{\substack{a: [2] \rightarrow [n] \\ a(1) \neq a(2)}} \Delta(a(1), a(2)) = \frac{n(n-1)}{n^2} \rightarrow 1 = (-1)^{k/2-m} \cdot C_{s_1-1}.$$

Now consider $t > 2$, and suppose the lemma has been established for every loop-free non-crossing partition on $t-1$ blocks. By Lemma 15, we may assume that k satisfies $t \geq k/2 + 1$. Then for $\pi \in \Pi(k, t)$, the pigeonhole principle guarantees that π contains a singleton block $\{j\} \in \pi$. By Lemma 13, we may assume $\{k\} \in \pi$. We proceed in cases:

Case I: $\pi(1) = \pi(k-1)$. We may apply Lemma 12(v) to obtain

$$\begin{aligned} V_n(\pi) &= \frac{1}{n^{k/2+1}} \sum_{a \in L_n(\pi)} \Delta(a(1), \dots, a(k-2), a(1), a(k)) \\ &= \frac{1}{n^{k/2}} \sum_{a \in L_n(\pi \setminus \{k\})} \Delta(a(1), \dots, a(k-2)) + o(1). \end{aligned}$$

The restriction of $\pi \setminus \{k\}$ to $[k-2]$ results in a loop-free non-crossing partition π' of $[k-2]$ into $t-1$ blocks. Moreover, the above expression for $V_n(\pi)$ implies

$$V_n(\pi) = V_n(\pi') + o(1). \quad (9)$$

For (i), observe that if G_π contains any odd cycles, then $G_{\pi'}$ must also contain odd cycles. In this case, we may apply our induction hypothesis to $V_n(\pi')$ to conclude $V_n(\pi) \rightarrow 0$. For (ii), the edges of G_π partition into m simple cycles of sizes $2s_1, \dots, 2s_m$ with $s_m = 1$, and so the edges of $G_{\pi'}$ partition into $m-1$ simple cycles of sizes $2s_1, \dots, 2s_{m-1}$. Then (9) and our induction hypothesis together imply

$$V_n(\pi) \rightarrow (-1)^{k/2-m} \cdot C_{s_1-1} \cdots C_{s_{m-1}-1}.$$

Since $C_0 = 1$, this establishes (ii).

Case II: $\pi(1) \neq \pi(k-1)$. In this case, $\pi(k)$ necessarily resides in a cycle of length $\ell \geq 3$. Select representatives $k, j_2, \dots, j_t \in [k]$ with $\pi(k) = B_1$ and $\pi(j_i) = B_i$ so that the vertices in the cycle are given by $B_1, \dots, B_\ell$. Then we may apply Lemma 12(iv) to obtain

$$\begin{aligned} V_n(\pi) &= \frac{1}{n^{k/2+1}} \sum_{a \in L_n(\pi \setminus \{k\})} \sum_{\substack{a_k \in [n] \\ a_k \notin a([k-1])}} \Delta(a(1), \dots, a(k-1), a_k) \\ &= -\frac{1}{n^{k/2+1}} \sum_{a \in L_n(\pi \setminus \{k\})} \sum_{i=2}^t \Delta(a(1), \dots, a(k-1), a(j_i)). \end{aligned}$$

For $i, j \in [2, t]$, define new blocks

$$B_j^i = \begin{cases} B_j \cup \{k\}, & \text{if } j = i \\ B_j, & \text{if } j \neq i, \end{cases}$$

and the corresponding partitions $\pi^i = \{B_2^i, \dots, B_t^i\}$, we have

$$V_n(\pi) = -\sum_{i=2}^t V_n(\pi^i).$$

By Lemma 14, $V_n(\pi^i) \rightarrow 0$ whenever π^i is a crossing partition. Since π^i is obtained from π by merging blocks B_i and $\{k\}$, we can argue as in the proof of Lemma 15 to conclude that π^i is crossing if and only if B_i and $\{k\}$ do not reside in the same simple cycle of G_π . Hence,

$$V_n(\pi) = -\sum_{i=2}^{\ell} V_n(\pi^i) + o(1),$$

where each π^i is non-crossing for $2 \leq i \leq \ell$. Both π^2 and π^ℓ contain loops, so $V_n(\pi^2) = V_n(\pi^\ell) = 0$ by Lemma 12(ii). When $\ell = 3$, this gives $V_n(\pi) \rightarrow 0$, as desired by (i). Supposing for the remainder that $\ell \geq 4$, we must still compute the limit of

$$V_n(\pi) = -\sum_{i=3}^{\ell-1} V_n(\pi^i) + o(1). \quad (10)$$

Observe that our cycle $\{B_1, \dots, B_\ell\}$ in G_π of length ℓ corresponds to the two simple cycles in G_{π^i} of $\{B_2^i, \dots, B_i^i\}$ and $\{B_i^i, B_{i+1}^i, \dots, B_\ell^i\}$ with lengths $i-1$ and $\ell-i+1$ and share the cut vertex B_i^i . Moreover, all other simple cycles are identical between the two graphs.

If ℓ is odd, then for each $i \in [3, \ell-1]$, either $i-1$ or $\ell-i+1$ is odd, and so G_{π^i} must have an odd cycle. Since each π^i has $t-1$ blocks and an odd cycle, we can apply our induction hypothesis to conclude that each $V_n(\pi^i) \rightarrow 0$ so that $V_n(\pi) \rightarrow 0$, as desired by (i). Suppose instead that ℓ is even, but G_π has an odd cycle. This odd cycle is also contained in each G_{π^i} for $i \in [3, \ell-1]$, and again we can apply our induction hypothesis to conclude that $V_n(\pi) \rightarrow 0$, thereby establishing (i).

Finally, for (ii), suppose that ℓ is even and that the edges of G_π partition into m cycles of lengths $2s_1, \dots, 2s_m$ with $2s_m = \ell$. Notice that if $i-1$ is odd, then G_{π^i} contains an odd cycle, and $V_n(\pi^i) \rightarrow 0$. Since the contribution of these terms is negligible, we must compute the limit of

$$V_n(\pi) = - \sum_{i=1}^{\ell/2-1} V_n(\pi^{2i+1}) + o(1).$$

The cycles of lengths $2s_1, \dots, 2s_{m-1}$ are common to both G_π and $G_{\pi^{2i+1}}$, while the cycle of length $\ell = 2s_m$ in G_π corresponds to two cycles of length $2i$ and $\ell-2i$ in $G_{\pi^{2i+1}}$. Applying our induction hypothesis, we have

$$V_n(\pi) \rightarrow (-1)^{k/2-m} \cdot C_{s_1-1} \cdots C_{s_{m-1}-1} \sum_{i=1}^{\ell/2-1} C_{i-1} C_{\ell/2-i-1}.$$

Reindexing and applying the convolution identity for Catalan numbers, we have

$$\sum_{i=1}^{\ell/2-1} C_{i-1} C_{\ell/2-i-1} = \sum_{i=0}^{\ell/2-2} C_i C_{\ell/2-i-2} = C_{\ell/2-1}.$$

Hence, $V_n(\pi) \rightarrow (-1)^{k/2-m} \cdot C_{s_1-1} \cdots C_{s_{m-1}-1}$, thereby establishing (ii). $\square$

Proof of Lemma 7. To prove (i), consider $\pi \in \Pi(k, t) \setminus \text{EC}(k, t)$. Then either π is crossing, G_π contains a loop, or G_π contains an odd cycle. If π is crossing, then $V_n(\pi) \rightarrow 0$ by Lemma 14. If G_π contains a loop, then $V_n(\pi) = 0$ by Lemma 12(ii). If π is loop-free and non-crossing but G_π contains an odd cycle, then $V_n(\pi) \rightarrow 0$ by Lemma 16(i). This establishes (i). Finally, (ii) follows from applying both Lemmas 15 and 16(ii). $\square$

Acknowledgments

The authors thank Ram Zamir and the anonymous reviewers for extensive feedback that greatly improved the presentation of both our results and the relevant literature. MM and DGM were partially supported by AFOSR FA9550-18-1-0107. DGM was also supported by NSF DMS 1829955 and the Simons Institute of the Theory of Computing. HP was partially supported by an AMS-Simons Travel Grant.

References

- [1] G. W. Anderson, B. Farrell, Asymptotically liberating sequences of random unitary matrices, *Adv. Math.* 255 (2014) 381–413.
- [2] A. S. Bandeira, M. Fickus, D. G. Mixon, P. Wong, The road to deterministic matrices with the restricted isometry property, *J. Fourier Anal. Appl.* 19 (2013) 1123–1149.
- [3] A. S. Bandeira, D. G. Mixon, J. Moreira, A conditional construction of restricted isometries, *Int. Math. Res. Notices* 2017 (2017) 372–381.
- [4] V. Bargmann, Note on Wigner’s theorem on symmetry operations, *J. Math. Phys.* 5 (1964) 862–868.
- [5] J. J. Benedetto, M. Fickus, Finite normalized tight frames, *Adv. Comput. Math.* 18 (2003) 357–385.
- [6] J. Bourgain, S. Dilworth, K. Ford, S. Konyagin, D. Kutzarova, Explicit constructions of RIP matrices and related problems, *Duke Math. J.* 159 (2011) 145–185.
- [7] Y. Cai, C. Yan, Counting with Borel’s triangle, *Discrete Math.* 342 (2019) 529–539.
- [8] R. Calderbank, S. Howard, S. Jafarpour, Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property, *IEEE J. Sel. Topics Signal Process.* 4 (2010) 358–374.
- [9] E. J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inf. Theory* 52 (2006) 489–509.
- [10] P. G. Casazza, M. Fickus, J. C. Tremain, E. Weber, The Kadison–Singer problem in mathematics and engineering—a detailed account, in: *Operator Theory, Operator Algebras, and Applications*, D. Han, P. E. T. Jorgensen, D. R. Larson (eds.), *Contemporary Mathematics*, vol. 414, Providence, RI: American Mathematical Society, 2006, pp. 299–355.
- [11] G. Cébron, A. Dahlqvist, C. Male, Universal constructions for spaces of traffics, *arXiv:1601.00168*
- [12] I. Daubechies, A. Grossmann, Y. Meyer, Painless nonorthogonal expansions, *J. Math. Phys.* 27 (1986) 1271–1283.
- [13] R. A. DeVore, Deterministic constructions of compressed sensing matrices, *J. Complexity* 23 (2007) 918–925.
- [14] D. L. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory* 52 (2006) 1289–1306.
- [15] A. Dubbs, A. Edelman, Infinite random matrix theory, tridiagonal bordered Toeplitz matrices, and the moment problem, *Linear Algebra Appl.* 467 (2015) 188–201.

- [16] R. J. Duffin, A. C. Schaeffer, A class of nonharmonic Fourier series, *Trans. Am. Math. Soc.* 72 (1952) 341–366.
- [17] B. Farrell, Limiting empirical singular value distribution of restrictions of discrete Fourier transform matrices *J. Fourier Anal. Appl.* 17 (2011) 733–753.
- [18] M. Fickus, D. G. Mixon, Tables of the existence of equiangular tight frames, [arXiv:1504.00253](https://arxiv.org/abs/1504.00253)
- [19] S. Gurevich, R. Hadani, The statistical restricted isometry property and the Wigner semicircle distribution of incoherent dictionaries, <https://hilbert.math.wisc.edu/~shamgar/SRIP-AAP-6-2-09.pdf>
- [20] M. Haikin, Analog Coding Frame-work Tel Aviv University, Masters thesis, 2018.
- [21] M. Haikin, R. Zamir, M. Gavish, Frame Moments and Welch Bound with Erasures, *ISIT 2018*, 2057–2061.
- [22] M. Haikin, R. Zamir, M. Gavish, Random subsets of structured deterministic frames have MANOVA spectra, *Proc. Natl. Acad. Sci. U.S.A.* 114 (2017) E5024–E5033.
- [23] Y. J. Ionin, H. Kharachani, Balanced generalized weighing matrices and conference matrices, In: *Handbook of Combinatorial Designs*, 2nd ed., 2007, 419–435.
- [24] R. Kadison, I. Singer, Extensions of pure states, *Am. J. Math.* 81 (1959) 383–400.
- [25] A. W. Marcus, D. A. Spielman, N. Srivastava, Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem, *Ann. Math.* (2015) 327–350.
- [26] B. D. McKay, The expected eigenvalue distribution of a large regular graph, *Linear Algebra Appl.* 40 (1981) 203–216.
- [27] D. G. Mixon, Unit norm tight frames in finite-dimensional spaces, in: *Finite Frame Theory: A Complete Introduction to Overcompleteness*, K. A. Okoudjou (ed.), *Proceedings of Symposia in Applied Mathematics*, 2016, pp. 68–93.
- [28] D. G. Mixon, C. J. Quinn, N. Kiyavash, M. Fickus, Fingerprinting with equiangular tight frames, *IEEE Trans. Inf. Theory* 59 (2013) 1855–1865.
- [29] A. Nica, R. Speicher, *Lectures on the combinatorics of free probability*, vol. 13, Cambridge University Press, 2006.
- [30] J. M. Renes, R. Blume-Kohout, A. J. Scott, C. M. Caves, Symmetric informationally complete quantum measurements, *J. Math. Phys.* 45 (2004) 2171–2180.
- [31] I. Soloveychik, Y. Xiang, V. Tarokh, Pseudo-Wigner matrices, *IEEE Trans. Inform. Theory* 64 (2018) 3170–3178.
- [32] I. Soloveychik, Y. Xiang, V. Tarokh, Symmetric pseudo-random matrices, *IEEE Trans. Inform. Theory* 64 (2018) 3179–3196.

- [33] T. Strohmer, R. W. Heath, Grassmannian frames with applications to coding and communication, *Appl. Comput. Harmon. Anal.* 14 (2003) 257–275.
- [34] T. Tao, *Topics in random matrix theory*, Graduate Studies in Mathematics 132 (2011).
- [35] J. A. Tropp, On the conditioning of random subdictionaries, *Appl. Comput. Harmon. Anal.* 25 (2008) 1–24.
- [36] A. M. Tulino, G. Caire, S. Shamai, S. Verdú, Capacity of channels with frequency-selective and time-selective fading, *IEEE Trans. Inform. Theory* 56 (2010) 1187–1215.
- [37] K. W. Wachter, The strong limits of random matrix spectra for sample matrices of independent elements, *Ann. Probab.* 6 (1978) 1–18.
- [38] N. Weaver, The Kadison–Singer problem in discrepancy theory, *Discrete Math.* 278 (2004) 227–239.