

# ASYMPTOTICALLY LIBERATING SEQUENCES OF RANDOM UNITARY MATRICES

GREG W. ANDERSON AND BRENDAN FARRELL

ABSTRACT. A fundamental result of free probability theory due to Voiculescu and subsequently refined by many authors states that conjugation by independent Haar-distributed random unitary matrices delivers asymptotic freeness. In this paper we exhibit many other systems of random unitary matrices that, when used for conjugation, lead to freeness. We do so by first proving a general result asserting “asymptotic liberation” under quite mild conditions, and then we explain how to specialize these general results in a striking way by exploiting Hadamard matrices. In particular, we recover and generalize results of the second-named author and of Tulino-Caire-Shamai-Verdú.

## 1. INTRODUCTION

Of the results of Voiculescu [23, 24] providing the foundations for free probability theory, arguably the simplest and most familiar is the following. Let  $A^{(N)}$  and  $B^{(N)}$  be deterministic  $N$ -by- $N$  hermitian matrices with singular values bounded independently of  $N$  and having empirical distributions of eigenvalues tending in moments to limits  $\mu_A$  and  $\mu_B$ , respectively. Let  $U^{(N)}$  be an  $N$ -by- $N$  Haar-distributed random unitary matrix. Then the empirical distribution of eigenvalues of the sum  $A^{(N)} + U^{(N)}B^{(N)}U^{(N)*}$  tends in moments to the free additive convolution  $\mu_A \boxplus \mu_B$ . The question addressed here, very roughly speaking, is this: how much less random can we make  $U^{(N)}$  and still get free additive convolution in the limit? More generally, we ask: what sorts of random unitary matrices  $U^{(N)}$  make  $A_N$  and  $U^{(N)}B^{(N)}U^{(N)*}$  asymptotically free? Using the theory introduced here, we can show, for example, that the desired property of delivering asymptotic freeness is possessed by the random unitary matrix  $U^{(N)} = W^{(N)*} \frac{H^{(N)}}{\sqrt{N}} W^{(N)}$  where  $W^{(N)}$  is a uniformly distributed random  $N$ -by- $N$  signed permutation matrix and  $H^{(N)}$  is a deterministic  $N$ -by- $N$  complex Hadamard matrix. (See Corollary 3.5 below.)

One precedent for our line of research is the main result of [13] which calculates the limiting distribution of singular values of a randomly chosen submatrix of the  $N$ -by- $N$  discrete Fourier transform matrix  $\text{DFT}^{(N)}$ . We can recover this result using our theory. (See Corollary 3.9 below.)

Another and much farther-reaching precedent is [21, Lemma 1, p. 1194]. This result is part of a study applying free probabilistic methods to problems of signal

---

*Date:* January 7, 2014.

*2010 Mathematics Subject Classification.* 60B20, 42A61, 46L54, 15B52 .

*Key words and phrases.* Free probability, asymptotic liberation, random matrices, unitary matrices, Hadamard matrices.

G.W.A. is the corresponding author.

B.F. is partially supported by Joel A. Tropp under ONR awards N00014-08-1-0883 and N00014-11-1002 and a Sloan Research Fellowship.

processing. A sample application of the result is the following. Let  $X$  and  $Y$  be bounded classical real random variables. Let  $X^{(N)}$  and  $Y^{(N)}$  be independent  $N$ -by- $N$  diagonal matrices with diagonal entries that are i.i.d. copies of  $X$  and  $Y$ , respectively. Then  $X^{(N)}$  and  $\text{DFT}^{(N)}Y^{(N)}\text{DFT}^{(N)*}$  are asymptotically free. The latter result we can recover from our theory. (See Corollary 3.7 below.)

The notion of asymptotic freeness of Haar-distributed unitaries and other types of random or deterministic matrices has been extensively developed by many authors and in many directions. We just mention the papers [7], [8] and [16] as particularly important influences on our work. The reader may consult, say, [1, Chap. 5], [15] or [23] for background and further references.

To the extent we make progress in this paper we do so by side-stepping issues of asymptotic freeness almost entirely. Instead we focus on the notion of *asymptotic liberation* (see §2.2) which is far easier to define and manipulate than asymptotic freeness. We mention in passing that the operator-theoretic paper [9] helped to push us toward a point of view emphasizing operators conjugation by which create freeness, and in particular we learned the term “liberation” from that source.

Another influence on the paper comes from applied mathematics, specifically the analysis of high-dimensional data. See for example the paper [20], which in the applied setting makes use of Hadamard matrices randomized both through random choice of block and randomization of signs through multiplication by diagonal matrices with i.i.d. diagonal entries of  $\pm 1$ . We use a similar randomization to create arbitrarily large asymptotically liberating families from a single deterministic Hadamard matrix. (See Corollary 3.2 below.)

The paper [14] considers problems at least superficially similar to those considered here, the simplest of which also have the form  $A^{(N)} + U^{(N)*}B^{(N)}U^{(N)}$ , but the focus is not on finding special distributions for  $U^{(N)}$  to create freeness; rather, with  $U^{(N)}$  a uniformly distributed permutation matrix, the goal is to calculate the limit under various hypotheses on (possibly random) matrices  $A^{(N)}$  and  $B^{(N)}$ , using the theoretical framework of *traffics* it is ultimately the purpose of [14] to introduce. This theory has significant applications, for example, to limit laws for heavy-tailed Wigner matrices. The problems considered typically fall outside the domain of free probability theory proper. A point of view encompassing both that of [14] and of this paper would be very interesting to have.

Here is a brief outline of the paper. In §2 we introduce the notion of asymptotic liberation, we state a technical result (see Proposition 2.6 below) spelling out the relationship between asymptotic liberation and asymptotic freeness, and finally we state our main result (see Theorem 2.8 below). In §3 we work out a number of corollaries to the main result. In §4 we reduce the proof of Theorem 2.8 to a combinatorial statement (see Theorem 4.2 below). In §5 we complete the proof of Theorem 2.8 by proving Theorem 4.2. In the proof of the latter the references [25, Lemma 3.4] and [26] play an important role. In §6 we supply a brief review of the relevant notions of free probability and a proof of Proposition 2.6. In §7 we prove a concentration result along the lines of that obtained in [4] based not on concentration for Haar measure on the unitary group but rather on concentration for the group of signed permutation matrices. Finally, in §8 we make concluding remarks and mention several more important references the relevance of which is easier to explain after the main work of the paper has been carried out.

## 2. FORMULATION OF THE MAIN RESULT

**2.1. Matrix notation.** Let  $\text{Mat}_N$  denote the space of  $N$ -by- $N$  matrices with complex entries. We invariably denote the entry of a matrix  $A$  in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column by  $A(i, j)$ . Let  $\mathbf{I}_N \in \text{Mat}_N$  denote the  $N$ -by- $N$  identity matrix. For  $A \in \text{Mat}_N$ , let  $A^* \in \text{Mat}_N$  denote the transpose conjugate of  $A$ , let  $\|A\|$  denote the largest singular value of  $A$  and let  $\text{tr}A$  denote the trace of  $A$  not normalized by division by  $N$ , i.e.,  $\text{tr}A = \sum_{i=1}^N A(i, i)$ .

**2.2. Asymptotically liberating sequences.** Let  $I$  be a finite index set. For each positive integer  $N$  and index  $i \in I$  suppose one is given a random unitary matrix  $U_i^{(N)} \in \text{Mat}_N$  defined on a probability space depending only on  $N$ . We say that the sequence of families

$$\left\{ \left\{ U_i^{(N)} \right\}_{i \in I} \right\}_{N=1}^{\infty}$$

is *asymptotically liberating* if for  $i_1, \dots, i_\ell \in I$  satisfying

$$(1) \quad \ell \geq 2, \quad i_1 \neq i_2, \quad \dots, \quad i_{\ell-1} \neq i_\ell, \quad i_\ell \neq i_1,$$

there exists a constant  $c(i_1, \dots, i_\ell)$  such that

$$(2) \quad \left| \mathbf{E} \text{tr} \left( U_{i_1}^{(N)} A_1 U_{i_1}^{(N)*} \dots U_{i_\ell}^{(N)} A_\ell U_{i_\ell}^{(N)*} \right) \right| \leq c(i_1, \dots, i_\ell) \|A_1\| \dots \|A_\ell\|$$

for all positive integers  $N$  and constant matrices  $A_1, \dots, A_\ell \in \text{Mat}_N$  each of trace zero. We emphasize that  $c(i_1, \dots, i_\ell)$  is independent of  $N$  and  $A_1, \dots, A_\ell$ , but may depend on  $i_1, \dots, i_\ell$  (and thus, since  $I$  is finite, can be chosen to depend on  $\ell$  alone).

The interest of the notion of asymptotic liberation stems from its close relationship with the concept of asymptotic freeness. We will describe the relationship precisely in Proposition 2.6 below.

**2.3. Remark on constants.** We always use  $c$ ,  $C$  or  $K$  (perhaps with subscripts or function arguments) to denote constants in estimates. The numerical values of these constants may vary from context to context. When we need to recall some particular constant we will take care to reference the line on which it was introduced.

Since readers familiar with free probability might find the condition  $i_\ell \neq i_1$  appearing on line (1) jarring, we hasten to make the following simple observation.

**Proposition 2.4.** *As above, let  $\left\{ \left\{ U_i^{(N)} \right\}_{i \in I} \right\}_{N=1}^{\infty}$  be asymptotically liberating.*

*Then for  $i_1, \dots, i_\ell \in I$  satisfying*

$$(3) \quad \ell \geq 2, \quad i_1 \neq i_2, \quad \dots, \quad i_{\ell-1} \neq i_\ell, \quad (\text{but equality } i_\ell = i_1 \text{ not a priori ruled out}),$$

*there exists a constant  $c(i_1, \dots, i_\ell)$  satisfying (2) for all positive integers  $N$  and matrices  $A_1, \dots, A_\ell \in \text{Mat}_N$  each of trace zero.*

*Proof.* For any  $i_1, \dots, i_\ell \in I$  and  $A_1, \dots, A_\ell \in \text{Mat}_N$  we temporarily write

$$\phi_{i_1 \dots i_\ell}^{(N)}(A_1, \dots, A_\ell) = \mathbf{E} \text{tr} \left( U_{i_1}^{(N)} A_1 U_{i_1}^{(N)*} \dots U_{i_\ell}^{(N)} A_\ell U_{i_\ell}^{(N)*} \right).$$

Now suppose that  $i_1, \dots, i_\ell \in I$  satisfy (3) and  $A_1, \dots, A_\ell$  are of trace zero. To prove existence of  $c(i_1, \dots, i_\ell)$  we may assume that  $\ell \geq 3$  and  $i_\ell = i_1$  because otherwise

there is nothing to prove. We then have a relation

$$\begin{aligned} \phi_{i_1 \dots i_\ell}^{(N)}(A_1, \dots, A_\ell) &= \phi_{i_1 \dots i_{\ell-1}}^{(N)} \left( A_\ell A_1 - \frac{1}{N} \text{tr}(A_\ell A_1) \mathbf{I}_N, A_2, \dots, A_{\ell-1} \right) \\ &\quad + \frac{1}{N} \text{tr}(A_\ell A_1) \phi_{i_2 \dots i_{\ell-1}}^{(N)}(A_2, \dots, A_{\ell-1}). \end{aligned}$$

Thus we can take, say,

$$c(i_1, \dots, i_\ell) = 2c(i_1, \dots, i_{\ell-1}) + \begin{cases} 0 & \text{if } \ell = 3, \\ c(i_2, \dots, i_{\ell-1}) & \text{if } \ell > 3 \end{cases}$$

by induction on  $\ell$ . □

**2.5. The relationship between asymptotic freeness and asymptotically liberating sequences.** (See §6.1 below for a brief review of relevant notions from noncommutative and free probability theory. Terms not otherwise defined here are recalled there. We work in a setup fairly close to that of [1, Chap. 5].) Let  $I$  be a finite index set. For each  $i \in I$ , let  $J_i$  be a further finite index set. Suppose that for each positive integer  $N$  and index  $i \in I$  one is given a family  $\{T_{ij}^{(N)}\}_{j \in J_i}$  of deterministic  $N$ -by- $N$  matrices with complex entries and an  $N$ -by- $N$  random unitary matrix  $U_i^{(N)}$  on a probability space depending only on  $N$ . These data in hand, we may then speak for each  $i \in I$  of the joint law

$$\tau_i^{(N)} : \mathbb{C}\langle \{\mathbf{X}_{ij}\}_{j \in J_i} \rangle \rightarrow \mathbb{C}$$

of

$$\left\{ T_{ij}^{(N)} \right\}_{j \in J_i}$$

viewed as a family of noncommutative random variables and of the joint law

$$\mu^{(N)} : \mathbb{C}\langle \{\{\mathbf{X}_{ij}\}_{j \in J_i}\}_{i \in I} \rangle \rightarrow \mathbb{C}$$

of

$$\left\{ \left\{ U_i^{(N)} T_{ij}^{(N)} U_i^{(N)*} \right\}_{j \in J_i} \right\}_{i \in I}$$

viewed as a triangular array of noncommutative random variables. For the purpose of defining these joint laws for a given  $N$  it is understood that the ambient noncommutative probability space is the pair consisting of

- the algebra of  $N$ -by- $N$  random matrices with essentially bounded complex entries defined on the same probability space as the family  $\{U_i^{(N)}\}_{i \in I}$  and
- the tracial state  $A \mapsto \frac{1}{N} \mathbf{E} \text{tr} A$ .

(For lack of need we do not name this space.) We make the following assumptions.

$$(4) \quad \sup_N \max_{i \in I} \max_{j \in J_i} \left[ \left[ T_{ij}^{(N)} \right] \right] < \infty.$$

$$(5) \quad \text{For each } i \in I, \text{ the limit } \tau_i = \lim_{N \rightarrow \infty} \tau_i^{(N)} \text{ in moments exists.}$$

$$(6) \quad \text{The sequence of families } \left\{ \left\{ U_i^{(N)} \right\}_{i \in I} \right\}_{N=1}^\infty \text{ is asymptotically liberating.}$$

The following technical result explains the relationship between asymptotic liberation and asymptotic freeness, thus setting the stage for our main result.

**Proposition 2.6.** *Notation and assumptions are as described above. Then the limit  $\mu = \lim_{N \rightarrow \infty} \mu^{(N)}$  in moments exists and is tracial. Furthermore, with respect to  $\mu$ , the rows of the triangular array  $\{\{\mathbf{X}_{ij}\}_{j \in J_i}\}_{i \in I}$  are free of each other, and for each  $i \in I$  the joint law of the row  $\{\mathbf{X}_{ij}\}_{j \in J_i}$  is  $\tau_i$ .*

We give the straightforward proof in §6 below. We emphasize that the limit  $\mu$  is uniquely determined by its marginals  $\tau_i$  in the presence of the asserted freeness.

**2.7. Further notation and terminology.** A matrix  $W \in \text{Mat}_N$  with entries  $W(i, j) = \epsilon_i \delta_{i, \sigma(j)}$  for some signs  $\epsilon_1, \dots, \epsilon_N \in \{\pm 1\}$  and a permutation  $\sigma \in S_N$  will be called a *signed permutation matrix*. For a  $\mathbb{C}$ -valued random variable  $Z$  let  $\|Z\|_p = (\mathbf{E}|Z|^p)^{1/p}$  for exponents  $p \in [1, \infty)$ . Also let  $\|Z\|_\infty$  denote the essential supremum of  $|Z|$ . For random vectors  $X$  and  $Y$  we write  $X \stackrel{d}{=} Y$  to denote equality in distribution.

Here is the main result of the paper.

**Theorem 2.8.** *Let  $I$  be a finite index set. For each positive integer  $N$  and index  $i \in I$  let there be given a random unitary matrix  $U_i^{(N)} \in \text{Mat}_N$  defined on a probability space depending only on  $N$ . For distinct  $i, i' \in I$  put  $U_{ii'}^{(N)} = U_i^{(N)*} U_{i'}^{(N)}$  to abbreviate notation. Make the following assumptions.*

(7) *For each positive integer  $N$  and deterministic signed permutation matrix  $W \in \text{Mat}_N$ , one has*

$$\left\{ W^* U_{ii'}^{(N)} W \right\}_{\substack{i, i' \in I \\ \text{s.t. } i \neq i'}} \stackrel{d}{=} \left\{ U_{ii'}^{(N)} \right\}_{\substack{i, i' \in I \\ \text{s.t. } i \neq i'}}.$$

(8) *For each positive integer  $\ell$ , one has*

$$\sup_{N=1}^{\infty} \max_{\substack{i, i' \in I \\ \text{s.t. } i \neq i'}} \max_{\alpha, \beta=1}^N \sqrt{N} \left\| \left( U_{ii'}^{(N)} \right) (\alpha, \beta) \right\|_\ell < \infty.$$

*Then the sequence of families  $\left\{ \left\{ U_i^{(N)} \right\}_{i \in I} \right\}_{N=1}^{\infty}$  is asymptotically liberating.*

We will derive Theorem 2.8 from a combinatorial estimate having nothing *a priori* to do with free probability, namely Theorem 4.2 below.

## 2.9. Remarks and notes on the theorem.

2.9.1. No claim is made that Theorem 2.8 exhausts the class of asymptotically liberating sequences. It remains to find other useful general classes of examples. It is likely the recently introduced theory of traffics [14] can provide many examples.

2.9.2. It should be possible to use many other groups besides the group of signed permutation matrices to state variants of hypothesis (7). The group of signed permutation matrices was emphasized here because of its convenience for applications.

2.9.3. In hypothesis (7) we merely require the law of the random vector  $\left\{ U_{ii'}^{(N)} \right\}_{\substack{i, i' \in I \\ \text{s.t. } i \neq i'}}$  to exhibit invariance under a certain action of the group of  $N$ -by- $N$  signed permutation matrices. Dependence of the family  $\{U_i^{(N)}\}_{i \in I}$  is not ruled out. This is in contrast to the usual run of free independence results.

2.9.4. Nothing is said in Theorem 2.8 about the asymptotic freeness of the families  $\{U_i^{(N)}\}_{i \in I}$  themselves. The theorem only concerns the possibility of making other families of noncommutative random variables asymptotically free by conjugation.

2.9.5. If each random matrix  $U_i^{(N)}$  is Haar-distributed in the  $N$ -by- $N$  unitary (resp., orthogonal) group and each family  $\{U_i^{(N)}\}_{i \in I}$  is independent, then

$\left\{ \left\{ U_i^{(N)} \right\}_{i \in I} \right\}_{N=1}^{\infty}$  satisfies the hypotheses of Theorem 2.8, as is easy to check.

Indeed, hypothesis (7) holds by definition of Haar measure and hypothesis (8) holds because each entry of a Haar-distributed  $N$ -by- $N$  unitary (resp., orthogonal) random matrix is approximately a centered complex (resp., real) normal random variable of variance  $1/N$ .

2.9.6. Let  $A^{(N)}$  and  $B^{(N)}$  be deterministic  $N$ -by- $N$  hermitian matrices which are  $\llbracket \cdot \rrbracket$ -bounded in  $N$  and have empirical distributions of eigenvalues tending in moments to limits  $\mu_A$  and  $\mu_B$ , respectively. Let  $U^{(N)}$  be an  $N$ -by- $N$  Haar-distributed random unitary. An often-quoted result of Voiculescu [24] asserts that the empirical distribution of eigenvalues of  $A^{(N)} + U^{(N)}B^{(N)}U^{(N)*}$  tends in moments to the free additive convolution  $\mu_A \boxplus \mu_B$ . We will see below in Corollary 3.5 below that Theorem 2.8 permits a much-less-random unitary matrix  $U^{(N)}$  to be used to the same end.

### 3. APPLICATION TO FREE CONVOLUTION OF CLASSICAL RANDOM VARIABLES

We now specialize Theorem 2.8 so as to recover and generalize results of [13].

3.1. **Hadamard matrices.** We pause to review a special class of matrices. One says that  $H \in \text{Mat}_N$  is a *Hadamard matrix* (resp., *complex Hadamard matrix*) if  $\frac{H}{\sqrt{N}}$  is orthogonal (resp., unitary) and  $|H(i, j)| = 1$  for  $i, j = 1, \dots, N$ . The  $k$ -fold Kronecker product

$$\left[ \begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array} \right]^{\otimes k} \in \text{Mat}_N \quad (N = 2^k)$$

is the simplest example of a Hadamard matrix when  $N$  is a power of 2. More generally the *discrete Fourier transform matrix*  $\text{DFT}^{(N)} \in \text{Mat}_N$  with entries

$$(9) \quad \text{DFT}^{(N)}(i, j) = \frac{1}{\sqrt{N}} \exp \left( -2\pi i \frac{(i-1)(j-1)}{N} \right)$$

is an  $N$ -by- $N$  complex Hadamard matrix divided by  $\sqrt{N}$ , for every integer  $N > 0$ .

We now have the following specialization of Theorem 2.8, the proof of which relies on the Khinchin inequality.

**Corollary 3.2.** *Let  $I$  be a finite index set. For each positive integer  $N$ , let  $H^{(N)} \in \text{Mat}_N$  be a deterministic complex Hadamard matrix, let  $W^{(N)} \in \text{Mat}_N$  be a uniformly distributed random signed permutation matrix and let  $\{D_i^{(N)}\}_{i \in I}$  be an i.i.d. family of uniformly distributed random  $N$ -by- $N$  diagonal signed permutation matrices independent of  $W^{(N)}$ . Then*

$$(10) \quad \left\{ \left\{ W^{(N)} \right\} \cup \left\{ \frac{H^{(N)}}{\sqrt{N}} W^{(N)} \right\} \cup \left\{ D_i^{(N)} \frac{H^{(N)}}{\sqrt{N}} W^{(N)} \right\}_{i \in I} \right\}_{N=1}^{\infty}$$

*is an asymptotically liberating sequence of families of random unitary matrices.*

*Proof.* Let  $\tilde{I} \supset I$  be a set such that  $\tilde{I} \setminus I$  consists of two elements, say  $a$  and  $b$ . Put  $U_a^{(N)} = W^{(N)}$  and  $D_b^{(N)} = \mathbf{I}_N$ . For  $i \in I \cup \{b\}$ , put  $U_i^{(N)} = D_i^{(N)} \frac{H^{(N)}}{\sqrt{N}} W^{(N)}$ . Our goal is to show that  $\left\{ \left\{ U_i^{(N)} \right\}_{i \in \tilde{I}} \right\}_{N=1}^{\infty}$  is asymptotically liberating. In any case, it clearly satisfies hypothesis (7) of Theorem 2.8. Thus we have only to check hypothesis (8). To that end, arbitrarily fix positive integers  $\ell$  and  $N$ , indices  $\alpha, \beta \in \{1, \dots, N\}$  and distinct indices  $i, i' \in \tilde{I}$ . As in Theorem 2.8, put  $U_{ii'}^{(N)} = U_i^{(N)*} U_{i'}^{(N)}$ . Our task is to bound the quantity  $\sqrt{N} \left\| U_{ii'}^{(N)}(\alpha, \beta) \right\|_{\ell}$  independently of  $N$ ,  $\alpha$ ,  $\beta$ ,  $i$  and  $i'$ . (However, the bound may depend on  $\ell$ .) Now if  $a \in \{i, i'\}$ , then there is nothing to prove, because the bound is obviously 1. Thus we may assume that  $i, i' \in I \cup \{b\}$ . Then the quantity to be bounded can be written

$$(11) \quad \frac{1}{\sqrt{N}} \left\| \left( W^{(N)*} H^{(N)*} D_{i'}^{(N)} D_i^{(N)} H^{(N)} W^{(N)} \right) (\alpha, \beta) \right\|_{\ell}.$$

Note that the diagonal entries of the diagonal matrix  $D_{i'}^{(N)} D_i^{(N)}$  are i.i.d. uniformly distributed in  $\{\pm 1\}$ . Now the *Khinchin inequality* (see, e.g., [6, Chap. 10, Sec. 3] for a textbook treatment) says that for constants  $a_1, \dots, a_N \in \mathbb{C}$ , i.i.d. random variables  $\epsilon_1, \dots, \epsilon_N$  uniformly distributed in  $\{\pm 1\}$  and  $p \in [1, \infty)$ , one has

$$\left\| \sum_{i=1}^N a_i \epsilon_i \right\|_p \leq K_p \left( \sum_{i=1}^N |a_i|^2 \right)^{1/2}$$

where the constant  $K_p$  depends only on  $p$ , NOT on  $N$  or  $a_1, \dots, a_N$ . Clearly, the Khinchin constant  $K_{\ell}$  bounds the quantity (11).  $\square$

**3.3. Question.** Could the random complex Hadamard matrices  $D_i^{(N)} H^{(N)}$  appearing in (10) be replaced by deterministic complex Hadamard matrices  $H_i^{(N)}$  in sufficiently general position so as still to get asymptotic liberation? More generally we wonder to what extent randomness could be reduced in the construction described by Corollary 3.2.

**3.4. Fake Haar unitaries.** Let  $N$  run through the positive integers. Let  $W^{(N)}$  be a uniformly distributed  $N$ -by- $N$  random signed permutation matrix. Let  $H^{(N)}$  be an  $N$ -by- $N$  deterministic complex Hadamard matrix. Let

$$U^{(N)} = \frac{1}{N} W^{(N)*} H^{(N)} W^{(N)}.$$

Let  $A^{(N)}$  and  $B^{(N)}$  be deterministic  $N$ -by- $N$  hermitian matrices such that

$$(12) \quad \sup_{N=1}^{\infty} \left[ \left[ A^{(N)} \right] \right] < \infty, \quad \sup_{N=1}^{\infty} \left[ \left[ B^{(N)} \right] \right] < \infty,$$

$$(13) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \operatorname{tr}(A^{(N)})^k = \mathbf{E}A^k \quad \text{and} \quad \lim_{N \rightarrow \infty} \frac{1}{N} \operatorname{tr}(B^{(N)})^k = \mathbf{E}B^k$$

for integers  $k > 0$ , where  $A$  and  $B$  are bounded real random variables with laws  $\mu_A$  and  $\mu_B$ , respectively. We regard  $A^{(N)}$ ,  $B^{(N)}$  and  $U^{(N)}$  as noncommutative random variables. In this connection the ambient noncommutative probability space is the pair consisting of (i) the algebra of  $N$ -by- $N$  matrices with essentially bounded entries defined on the probability space where  $W^{(N)}$  is defined and (ii) the tracial state  $Z \mapsto \frac{1}{N} \mathbf{E} \operatorname{tr} Z$ .

Let  $\boxplus$  and  $\boxtimes$  denote, respectively, additive and multiplicative free convolution.

**Corollary 3.5.** *Assumptions and notation are as above. The noncommutative random variables*

$$A^{(N)} \text{ and } U^{(N)} B^{(N)} U^{(N)*}$$

*are asymptotically free. In particular, the law of*

$$A^{(N)} + U^{(N)} B^{(N)} U^{(N)*}$$

*converges in moments to*

$$\mu_A \boxplus \mu_B$$

*and (supposing  $A^{(N)}$  is nonnegative definite for every  $N$ ) the law of*

$$(A^{(N)})^{1/2} U^{(N)} B^{(N)} U^{(N)*} (A^{(N)})^{1/2}$$

*converges in moments to*

$$\mu_A \boxtimes \mu_B.$$

In view of the remark in §2.9.6 above, we venture to call  $U^{(N)}$  a *fake Haar unitary*. It is an interesting problem to find further examples of such “fakes,” the less random the better.

*Proof.* It is equivalent to show that the noncommutative random variables

$$W^{(N)} A^{(N)} W^{(N)*} \text{ and } \frac{1}{N} H^{(N)} W^{(N)} B^{(N)} W^{(N)*} H^{(N)*}$$

are asymptotically free. Hypotheses (4), (5) and (6) of Proposition 2.6 with

$$\begin{aligned} \left\{ \left\{ T_{ij}^{(N)} \right\}_{j \in J_i} \right\}_{i \in I} & \text{ specialized to } \left\{ A^{(N)}, B^{(N)} \right\} \text{ and} \\ \left\{ U_i^{(N)} \right\}_{i \in I} & \text{ specialized to } \left\{ W^{(N)}, \frac{H^{(N)}}{\sqrt{N}} W^{(N)} \right\} \end{aligned}$$

are fulfilled in view of (12), (13) and Corollary 3.2. The result follows.  $\square$

**3.6. Free convolution without “extra” randomness.** Let  $X$  and  $Y$  be bounded real random variables defined on the same probability space with marginal laws  $\nu_X$  and  $\nu_Y$ , respectively. (We need not assume that  $X$  and  $Y$  are independent and we may for example take  $X = Y$ .) Let  $\{(X(i), Y(i))\}_{i=1}^\infty$  be an i.i.d. family of copies of the pair  $(X, Y)$ . Let  $N$  run through positive integers. Let  $X^{(N)}$  (resp.  $Y^{(N)}$ ) be the  $N$ -by- $N$  diagonal random matrix with diagonal entries  $X(i)$  (resp.,  $Y(i)$ ). Let  $H^{(N)}$  be an  $N$ -by- $N$  complex Hadamard matrix. We view  $X^{(N)}$ ,  $Y^{(N)}$  and  $H^{(N)}$  as noncommutative random variables. In this connection the ambient noncommutative probability space is the pair consisting of (i) the algebra of  $N$ -by- $N$  random matrices with essentially bounded complex entries defined on the probability space where the sequence  $\{(X(i), Y(i))\}_{i=1}^\infty$  is defined and (ii) the tracial state  $A \mapsto \frac{1}{N} \mathbf{E} \operatorname{tr} A$ .

When  $X$  and  $Y$  are independent and  $H^{(N)} = \sqrt{N} \operatorname{DFT}^{(N)}$ , the next result is a special case of [21, Lemma 1, p. 1194].

**Corollary 3.7.** *Assumptions and notation are as above. The noncommutative random variables*

$$(14) \quad X^{(N)} \text{ and } \frac{1}{N} H^{(N)} Y^{(N)} H^{(N)*}$$

are asymptotically free. In particular, the law of

$$X^{(N)} + \frac{1}{N} H^{(N)} Y^{(N)} H^{(N)*}$$

converges in moments to

$$\nu_X \boxplus \nu_Y$$

and (assuming  $X^{(N)}$  has nonnegative entries for all  $N$ ) the law of

$$\frac{1}{N} (X^{(N)})^{1/2} H^{(N)} Y^{(N)} H^{(N)*} (X^{(N)})^{1/2}$$

converges in moments to

$$\nu_X \boxtimes \nu_Y.$$

We emphasize that the only randomness in these random matrices arises by taking i.i.d. samples from the distribution of  $(X, Y)$ . The unitary matrix  $\frac{H^{(N)}}{\sqrt{N}}$  by which we conjugate is deterministic.

*Proof.* The overall strategy is to pass via dominated convergence from a “quenched” result readily obtainable from Proposition 2.6 to the desired “annealed” result.

Let

$$X_1^{(N)} = X^{(N)} \quad \text{and} \quad X_2^{(N)} = Y^{(N)}.$$

We begin by observing that

$$(15) \quad \sup_{N=1}^{\infty} \left[ \left[ X_i^{(N)} \right] \right] \leq \|X_i\|_{\infty} < \infty \quad \text{a.s.}$$

for  $i = 1, 2$  since  $X$  and  $Y$  are assumed bounded. Furthermore, by the strong law of large numbers, we have

$$(16) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr}(X_i^{(N)})^k = \mathbf{E} X_i^k \quad \text{a.s.}$$

for  $i = 1, 2$  and for every positive integer  $k$ .

Let

$$V_1^{(N)} = \mathbf{I}_N \quad \text{and} \quad V_2^{(N)} = \frac{H^{(N)}}{\sqrt{N}}.$$

Let  $W^{(N)}$  be an  $N$ -by- $N$  uniformly distributed random signed permutation matrix. (The probability space on which  $W^{(N)}$  is defined is irrelevant.) By Corollary 3.2:

$$(17) \quad \text{The sequence } \left\{ V_1^{(N)} W^{(N)}, V_2^{(N)} W^{(N)} \right\}_{N=1}^{\infty} \text{ is asymptotically liberating.}$$

For  $i_1, \dots, i_{\ell} \in \{1, 2\}$  and  $A_1, \dots, A_{\ell} \in \text{Mat}_N$  let

$$\begin{aligned} & \psi_{i_1, \dots, i_{\ell}}^{(N)}(A_1, \dots, A_{\ell}) \\ &= \frac{1}{N} \mathbf{E} \text{tr} \left( V_{i_1}^{(N)} W^{(N)} A_1 W^{(N)*} V_{i_1}^{(N)*} \dots V_{i_{\ell}}^{(N)} W^{(N)} A_{\ell} W^{(N)*} V_{i_{\ell}}^{(N)*} \right) \\ &= \frac{1}{N \cdot 2^N \cdot N!} \sum_{\substack{N\text{-by-}N \text{ signed} \\ \text{permutation matrices } w}} \text{tr} \left( V_{i_1}^{(N)} w A_1 w^* V_{i_1}^{(N)*} \dots V_{i_{\ell}}^{(N)} w A_{\ell} w^* V_{i_{\ell}}^{(N)*} \right). \end{aligned}$$

In turn, we define a random variable

$$\theta_{i_1, \dots, i_{\ell}}^{(N)} = \psi_{i_1, \dots, i_{\ell}}^{(N)}(X_{i_1}^{(N)}, \dots, X_{i_{\ell}}^{(N)})$$

for which we have an evident bound

$$(18) \quad |\theta_{i_1, \dots, i_{\ell}}^{(N)}| \leq (\|X\|_{\infty} + \|Y\|_{\infty})^{\ell} \quad \text{a.s.}$$

uniform in  $N$ . Let

$$\mu^{(N)} : \mathbb{C}\langle \mathbf{X}_1, \mathbf{X}_2 \rangle \rightarrow \mathbb{C}$$

be the joint law of the noncommutative random variables (14). We then have

$$(19) \quad \begin{aligned} & \mu^{(N)}(\mathbf{X}_{i_1} \cdots \mathbf{X}_{i_\ell}) \\ &= \frac{1}{N} \mathbf{E} \operatorname{tr} \left( V_{i_1}^{(N)} X_{i_1}^{(N)} V_{i_1}^{(N)*} \cdots V_{i_\ell}^{(N)} X_{i_\ell}^{(N)} V_{i_\ell}^{(N)*} \right) = \mathbf{E} \theta_{i_1, \dots, i_\ell}^{(N)} \end{aligned}$$

because

$$(X^{(N)}, Y^{(N)}) \stackrel{d}{=} (wX^{(N)}w^*, wY^{(N)}w^*)$$

for any  $N$ -by- $N$  deterministic signed permutation matrix  $w$ .

Let  $\Omega_0$  be the event of probability 1 on which (15) and (16) hold exactly. For every sample point  $\omega \in \Omega_0$  the hypotheses (4), (5) and (6) of Proposition 2.6 with

$$\begin{aligned} \left\{ \left\{ T_{ij}^{(N)} \right\}_{j \in J_i} \right\}_{i \in I} & \text{specialized to } \left\{ X_1^{(N)}(\omega), X_2^{(N)}(\omega) \right\} \text{ and} \\ \left\{ U_i^{(N)} \right\}_{i \in I} & \text{specialized to } \left\{ V_1^{(N)} W^{(N)}, V_2^{(N)} W^{(N)} \right\} \end{aligned}$$

are fulfilled in view of (15), (16) and (17) above, respectively. Thus there exists for each  $\omega \in \Omega_0$  a state  $\mu_\omega$  on  $\mathbb{C}\langle \mathbf{X}_1, \mathbf{X}_2 \rangle$  with respect to which  $\mathbf{X}_1$  and  $\mathbf{X}_2$  are free with marginals  $\nu_X$  and  $\nu_Y$ , respectively, such that for all indices  $i_1, \dots, i_\ell \in \{1, 2\}$  we have

$$(20) \quad \mu_\omega(\mathbf{X}_{i_1} \cdots \mathbf{X}_{i_\ell}) = \lim_{N \rightarrow \infty} \theta_{i_1 \dots i_\ell}^{(N)}(\omega).$$

But then  $\mu_\omega$  is independent of  $\omega$  since it is uniquely determined by its marginals in the presence of freeness. Thus we write  $\mu = \mu_\omega$  hereafter. Finally, by (18), (19), (20) and dominated convergence,  $\mu^{(N)}$  tends in moments to  $\mu$ , thus completing the proof of asymptotic freeness.  $\square$

**3.8. Compressions and Discrete Fourier Transform Matrices.** Asymptotic freeness can be used to address the product of random projections. Consider independent random matrices  $P_\alpha, P_\beta \in \operatorname{Mat}_N$  that project orthogonally onto uniformly distributed subspaces of dimensions  $[\alpha N]$  and  $[\beta N]$ , respectively, where  $[\cdot]$  denotes the integer part. Recall that the *empirical distribution function* of a Hermitian matrix  $H \in \operatorname{Mat}_N$ , which we denote by  $F_H$ , is

$$F_H(x) = \frac{1}{N} \#\{i : \lambda_i(H) \leq x\},$$

where  $\lambda_1(H), \dots, \lambda_N(H)$  denote the eigenvalues of  $H$ . Then the expected empirical distribution function of  $P_\alpha P_\beta P_\alpha$  converges in moments to the distribution with density given in (22) below. Corollary 3.9 below shows that the same behavior occurs when the subspaces are the spans of appropriate random sets of standard basis vectors and columns of the discrete Fourier transform matrix  $\operatorname{DFT}^{(N)}$ . (See (9) above to be reminded of the definition of  $\operatorname{DFT}^{(N)}$ .)

We next explain in more detail the setup involving the discrete Fourier transform matrix  $\operatorname{DFT}^{(N)}$ . Let  $R_1, R_2 \in \operatorname{Mat}_N$  be two coordinate projection matrices, i.e. diagonal matrices with all diagonal entries equal to 0 or 1. Then  $R_2 \operatorname{DFT}^{(N)} R_1$  is the matrix that first projects onto a set of indices, performs the Fourier transform, and then projects onto another set of indices in the Fourier domain. In turn,  $\operatorname{DFT}^{(N)} R_2 \operatorname{DFT}^{(N)*} R_1$  is the composition of the projection onto the span of a set of

standard basis vectors followed by the projection onto the span of a set of columns of the discrete Fourier transform matrix. To make this operator self-adjoint, we compose it with its adjoint. Thus, the eigenvalues of

$$R_1 \text{DFT}^{(N)} R_2 \text{DFT}^{(N)*} R_1$$

describe the relationship between a subspace spanned by standard basis vectors and a subspace spanned by Fourier vectors. These considerations and the trivial observation that  $\sqrt{N} \text{DFT}^{(N)}$  belongs to the class of complex Hadamard matrices motivate the following result.

**Corollary 3.9.** *Let  $\{H^{(N)}\}_{N=1}^{\infty}$  be a sequence of  $N$ -by- $N$  complex Hadamard matrices. Define independent random variables  $X$  and  $Y$  by*

$$X = \begin{cases} 1 & \text{with probability } \alpha \\ 0 & \text{with probability } 1 - \alpha \end{cases} \quad \text{and} \quad Y = \begin{cases} 1 & \text{with probability } \beta \\ 0 & \text{with probability } 1 - \beta. \end{cases}$$

*In turn, define  $X^{(N)}$  and  $Y^{(N)}$  as in Corollary 3.7. Then the expected empirical distribution function of*

$$(21) \quad \frac{1}{N} X^{(N)} H^{(N)} Y^{(N)} H^{(N)*} X^{(N)}$$

*converges in moments to the distribution with density*

$$(22) \quad (1 - \min(\alpha, \beta))\delta_0(x) + \max(\alpha + \beta - 1, 0)\delta_1(x) + \frac{\sqrt{(\lambda_+ - x)(x - \lambda_-)}}{2\pi x(1 - x)} I_{[\lambda_-, \lambda_+]}(x),$$

*where  $I_{[\lambda_-, \lambda_+]}$  denotes the indicator function on  $[\lambda_-, \lambda_+]$  and*

$$\lambda_{\pm} = \alpha + \beta - 2\alpha\beta \pm \sqrt{4\alpha\beta(1 - \alpha)(1 - \beta)}.$$

In particular, we may take

$$\frac{H^{(N)}}{\sqrt{N}} = \text{DFT}^{(N)}$$

in the above statement, thus recovering the main result of [13]. Similar but considerably more general results involving  $\text{DFT}^{(N)}$  and bearing on signal processing were proved in [21].

*Proof.* By applying Corollary 3.7 we only need to determine the free multiplicative convolution of the laws of  $X$  and  $Y$ . This is the standard example of multiplicative free convolution and can be found in [23, Section 3.6]. The resulting distribution function has the density claimed.  $\square$

### 3.10. Remarks on the corollaries.

3.10.1. In all the corollaries mentioned above it is not really necessary to have matrices  $H^{(N)}$  which satisfy the definition of a complex Hadamard matrix exactly. We stuck to that case only to make all the statements above simple. The reader can easily verify that in all the corollaries it suffices merely, say, to have  $\frac{H^{(N)}}{\sqrt{N}} \in \text{Mat}_N$  unitary for every  $N$  and to have  $\max_{i,j=1}^N |H^{(N)}(i,j)|$  be a bounded function of  $N$ .

3.10.2. Corollaries 3.5 and 3.7 have been stated for pairs of variables for the sake of simplicity. Using Corollary 3.2 these statements may with evident modifications be generalized to handle finite collections of variables.

3.10.3. If in Corollary 3.5 one restricts attention to diagonal matrices  $A^{(N)}$  and  $B^{(N)}$  it takes even less randomness to create a fake Haar unitary. More precisely, one need only conjugate  $H^{(N)}$  by a uniformly distributed unsigned  $N$ -by- $N$  permutation matrix to achieve asymptotic liberation.

3.10.4. An analogue of Corollary 3.9 can be derived from Corollary 3.5. This analogue concerns the singular values of a randomly chosen block of  $\text{DFT}^{(N)}$  the numbers of rows and columns of which are fixed deterministic fractions of  $N$ .

3.10.5. The assumption in Corollary 3.9 of independence of  $X$  and  $Y$  can be dropped since it is not needed to apply Corollary 3.7. One can for example take  $X = Y$  and thus  $X^{(N)} = Y^{(N)}$  for all  $N$ . The latter generalization concerns the singular values of randomly chosen principal submatrices of  $\text{DFT}^{(N)}$ .

#### 4. REDUCTION OF THEOREM 2.8 TO A COMBINATORIAL ESTIMATE

##### 4.1. Formulation of the combinatorial estimate.

4.1.1. *Further notation.* Let  $|S|$  denote the cardinality of a finite set  $S$ . Let  $\mathbb{N}$  denote the set of positive integers. For  $N \in \mathbb{N}$ , let  $\langle N \rangle = \{1, \dots, N\}$ . For  $A \in \text{Mat}_N$ , let  $\|A\|_2 = (\text{tr}AA^*)^{1/2}$  denote the *Hilbert-Schmidt norm* of  $A$ .

4.1.2. *Functions of  $\chi$ -class.* Fix  $\ell, N \in \mathbb{N}$ . Let  $F : \langle N \rangle^\ell \rightarrow \mathbb{C}$  be a function. We say that  $F$  belongs to the  $\chi$ -class (the letter  $\chi$  is an allusion to Khinchin) if for all

$$i_1, \dots, i_{\ell-1} \in \langle N \rangle, \quad j \in \langle N \rangle \setminus \{i_1, \dots, i_{\ell-1}\} \quad \text{and} \quad \lambda \in \langle \ell \rangle$$

one has

$$F(i_1, \dots, i_{\lambda-1}, j, i_\lambda, \dots, i_{\ell-1}) = 0.$$

4.1.3. *Functions of  $\chi\chi$ -class.* Fix  $\ell, N \in \mathbb{N}$ . Let  $F : \langle N \rangle^{2\ell} \rightarrow \mathbb{C}$  be a function. We say that  $F$  belongs to the  $\chi\chi$ -class if  $F$  belongs to the  $\chi$ -class and for all

$$i_1, \dots, i_{2\ell-2} \in \langle N \rangle, \quad j, j' \in \langle N \rangle \setminus \{i_1, \dots, i_{2\ell-2}\} \quad \text{and} \quad \lambda \in \langle \ell \rangle$$

one has

$$F(i_1, \dots, i_{2\lambda-2}, j, j', i_{2\lambda-1}, \dots, i_{2\ell-2}) = F(i_1, \dots, i_{2\lambda-2}, j', j', i_{2\lambda-1}, \dots, i_{2\ell-2}).$$

Lemma 5.2.4 below explains the structure of functions of  $\chi\chi$ -class in more detail.

We have the following estimate implying Theorem 2.8. This is the main technical result underpinning the whole paper.

**Theorem 4.2.** *Fix  $\ell, N \in \mathbb{N}$ . Let  $F : \langle N \rangle^{2\ell} \rightarrow \mathbb{C}$  be a function of  $\chi\chi$ -class. Let  $A_1, \dots, A_\ell \in \text{Mat}_N$  be matrices each of trace zero and put*

$$A(\mathbf{i}) = \prod_{\lambda=1}^{\ell} A_\lambda(i_{2\lambda-1}, i_{2\lambda}) \quad \text{for } \mathbf{i} = (i_1, \dots, i_{2\ell}) \in \langle N \rangle^{2\ell}.$$

Then we have

$$(23) \quad \left| \sum_{\mathbf{i} \in \langle N \rangle^{2\ell}} F(\mathbf{i}) A(\mathbf{i}) \right| \leq C_\ell \max_{\mathbf{i} \in \langle N \rangle^{2\ell}} |F(\mathbf{i})| \prod_{\lambda=1}^{\ell} \|A_\lambda\|_2$$

for a constant  $C_\ell > 0$  depending only on  $\ell$ , NOT on  $N$ ,  $A_1, \dots, A_\ell$  or  $F$ .

We will prove Theorem 4.2 in §5. In the course of the proof we will highlight close connections with Yin's lemma [25, Lemma 3.4] and Whittle's results [26]. We emphasize that the Hilbert-Schmidt norm  $\llbracket \cdot \rrbracket_2$  intervenes here rather than the largest-singular-value norm  $\llbracket \cdot \rrbracket$ .

**4.3. Remark.** In the setup of Theorem 4.2 note that we have

$$\sum_{\mathbf{i}} |A(\mathbf{i})|^2 = \prod_{\lambda} \llbracket A_{\lambda} \rrbracket_2^2.$$

Thus, by Cauchy-Schwarz, (23) holds with  $C_{\ell} = N^{\ell}$ . The theorem beats the trivial bound  $N^{\ell}$  by a wide margin.

**4.4. Derivation of Theorem 2.8 from Theorem 4.2.** Fix an integer  $\ell \geq 2$  and a sequence of indices  $\alpha_1, \dots, \alpha_{\ell} \in I$  such that  $\alpha_1 \neq \alpha_2, \dots, \alpha_{\ell-1} \neq \alpha_{\ell}$  and  $\alpha_{\ell} \neq \alpha_1$ . For each  $N$  and index  $\lambda = 1, \dots, \ell$  we define a random unitary matrix

$$V_{\lambda}^{(N)} = \begin{cases} U_{\alpha_{\ell}}^{(N)*} U_{\alpha_1}^{(N)} & \text{if } \lambda = 1, \\ U_{\alpha_{\lambda-1}}^{(N)*} U_{\alpha_{\lambda}}^{(N)} & \text{if } \lambda > 1. \end{cases}$$

In turn, we define a (deterministic) function  $F^{(N)} : \langle N \rangle^{2\ell} \rightarrow \mathbb{C}$  by the formula

$$F^{(N)}(\mathbf{i}) = \mathbf{E} \prod_{\lambda=1}^{\ell} V_{\lambda}^{(N)}(i_{2\lambda-1}, i_{2\lambda}) \quad \text{for } \mathbf{i} = (i_2, i_3, \dots, i_{2\ell}, i_1) \in \langle N \rangle^{2\ell}$$

which on account of the ‘‘twist’’ in the definition satisfies

$$\begin{aligned} (24) \quad & \sum_{\mathbf{i} \in \langle N \rangle^{2\ell}} F^{(N)}(\mathbf{i}) A(\mathbf{i}) \\ &= \sum_{(i_2, i_3, \dots, i_{2\ell}, i_1) \in \langle N \rangle^{2\ell}} \mathbf{E}(V_1^{(N)}(i_1, i_2) A_1(i_2, i_3) \cdots V_{\ell}^{(N)}(i_{2\ell-1}, i_{2\ell}) A_{\ell}(i_{2\ell}, i_1)) \\ &= \mathbf{E} \text{tr}(V_1^{(N)} A_1 \cdots V_{\ell}^{(N)} A_{\ell}). \end{aligned}$$

Arbitrarily fix matrices  $A_1, \dots, A_{\ell} \in \text{Mat}_N$  of trace zero. It will be enough to show that

$$(25) \quad \left| \sum_{\mathbf{i} \in \langle N \rangle^{2\ell}} F^{(N)}(\mathbf{i}) A(\mathbf{i}) \right| \leq c(\alpha_1, \dots, \alpha_{\ell}) \prod_{\lambda=1}^{\ell} \llbracket A_{\lambda} \rrbracket$$

where  $c(\alpha_1, \dots, \alpha_{\ell})$  is a constant independent of  $N$  and  $A_1, \dots, A_{\ell}$ .

Fix  $\mathbf{i} = (i_1, \dots, i_{2\ell}) \in \langle N \rangle^{2\ell}$  arbitrarily. By hypothesis (7) used firstly with diagonal signed permutation matrices, and then secondly with unsigned permutation matrices, we have relations

$$(26) \quad F^{(N)}(\mathbf{i}) = \left( \prod_{\lambda=1}^{2\ell} \epsilon_{i_{\lambda}} \right) F^{(N)}(\mathbf{i}) \quad \text{for } \epsilon_1, \dots, \epsilon_N \in \{\pm 1\} \text{ and}$$

$$(27) \quad F^{(N)}(\sigma(i_1), \dots, \sigma(i_{2\ell})) = F^{(N)}(\mathbf{i}) \quad \text{for } \sigma \in S_N.$$

Formula (26) implies that  $F^{(N)}$  is of  $\chi$ -class and in turn formula (27) implies that  $F^{(N)}$  is of  $\chi\chi$ -class. By hypothesis (8) and the Hölder inequality we have

$$\sup_{N=1}^{\infty} N^{\ell/2} \max_{\mathbf{i} \in \langle N \rangle^{2\ell}} |F^{(N)}(\mathbf{i})| < \infty.$$

The bound (25) now follows from (23) and the bound

$$\llbracket A \rrbracket_2 \leq \llbracket \mathbf{I}_N \rrbracket_2 \llbracket A \rrbracket = \sqrt{N} \llbracket A \rrbracket$$

for  $A \in \text{Mat}_N$ . The proof of Theorem 2.8 modulo Theorem 4.2 is complete.  $\square$

## 5. PROOF OF THEOREM 4.2

**5.1. Background on set partitions.** We are strongly influenced by the paper [18] and recommend it as an introduction to the circle of combinatorial ideas being exploited here. (But we do not follow the notation of [18] very closely.) We also remark that while set partitions (especially noncrossing ones) have a large role to play in the combinatorial side of free probability theory, our use here of set partitions seems to be in a different spirit.

**5.1.1. Basic definitions.** Let  $\ell \in \mathbb{N}$  be given. A *partition* of  $\langle \ell \rangle$  is a family  $\Pi$  of subsets of  $\langle \ell \rangle$  such that  $\emptyset \notin \Pi$ ,  $\bigcup \Pi = \langle \ell \rangle$  and for all  $B_1, B_2 \in \Pi$ , if  $B_1 \cap B_2 \neq \emptyset$ , then  $B_1 = B_2$ . Members of a partition are called its *blocks*. We denote the family of partitions of  $\langle \ell \rangle$  by  $\text{Part}(\ell)$ . Given an  $\ell$ -tuple  $\mathbf{i} = (i_1, \dots, i_\ell) \in \langle N \rangle^\ell$ , let

$$\Pi(\mathbf{i}) = \{ \{ \lambda \in \langle \ell \rangle \mid i_\lambda = i \} \mid i \in \langle N \rangle \} \setminus \emptyset \in \text{Part}(\ell),$$

which we call the set partition *generated* by the  $\ell$ -tuple  $\mathbf{i}$ . For example, we have

$$\Pi(1, 2, 3, 3, 2, 2, 2, 4) = \{ \{1\}, \{2, 5, 6, 7\}, \{3, 4\}, \{8\} \} \in \text{Part}(8).$$

Given  $\Pi_1, \Pi_2 \in \text{Part}(\ell)$  we write  $\Pi_1 \leq \Pi_2$  and we say that  $\Pi_1$  *refines*  $\Pi_2$  if for every  $B_1 \in \Pi_1$  there exists  $B_2 \in \Pi_2$  such that  $B_1 \subset B_2$ . (In other words,  $\Pi_2$  arises from  $\Pi_1$  by coalescing some blocks of the latter.) In this way  $\text{Part}(\ell)$  becomes a partially ordered set. Let  $\mathbf{0}_\ell = \{ \{ \lambda \} \mid \lambda \in \langle \ell \rangle \} \in \text{Part}(\ell)$ , which is the minimal set partition (the one with the greatest number of blocks).

**5.1.2. The sum-product formula.** Fix  $\ell, N \in \mathbb{N}$ . Let  $f_1, \dots, f_\ell : \langle N \rangle \rightarrow \mathbb{C}$  be functions. Write

$$f(\mathbf{i}) = f_1(i_1) \cdots f_\ell(i_\ell) \text{ for } \mathbf{i} = (i_1, \dots, i_\ell) \in \langle N \rangle^\ell.$$

Then one has the trivial but important observation that

$$(28) \quad \sum_{\substack{\mathbf{i} \in \langle N \rangle^\ell \\ \text{s.t. } \Pi \leq \Pi(\mathbf{i})}} f(\mathbf{i}) = \prod_{B \in \Pi} \sum_{i \in \langle N \rangle} \prod_{b \in B} f_b(i) \text{ for } \Pi \in \text{Part}(\ell).$$

**5.1.3. Möbius inversion.** Consider the function

$$\zeta : \text{Part}(\ell) \times \text{Part}(\ell) \rightarrow \{0, 1\}, \quad \zeta(\Pi, \Theta) = \begin{cases} 1 & \text{if } \Pi \leq \Theta, \\ 0 & \text{otherwise,} \end{cases}$$

viewed as a matrix. It is by definition the *incidence matrix* of the poset  $\text{Part}(\ell)$ . Since  $\zeta$  is upper unitriangular, it is invertible and has an upper unitriangular inverse, denoted  $\mu : \text{Part}(\ell) \times \text{Part}(\ell) \rightarrow \mathbb{Z}$ , which is by definition the *Möbius function* of the poset  $\text{Part}(\ell)$ . By definition of  $\mu$  one has

$$(29) \quad \sum_{\Theta \leq \Pi} \mu(\mathbf{0}_\ell, \Theta) = \begin{cases} 1 & \text{if } \Pi = \mathbf{0}_\ell, \\ 0 & \text{otherwise,} \end{cases} \text{ for } \Pi \in \text{Part}(\ell),$$

which is a special case of the *Möbius inversion formula* for the poset  $\text{Part}(\ell)$ . In fact it is the only case needed here.

5.1.4. *Crude bounds.* It is well-known how to evaluate  $\mu$  explicitly. One has in particular

$$\mu(\mathbf{0}_\ell, \Pi) = \prod_{B \in \Pi} (-1)^{|B|-1} (|B| - 1)! \text{ for } \Pi \in \text{Part}(\ell),$$

and hence crudely

$$(30) \quad \max_{\Pi \in \text{Part}(\ell)} |\mu(\mathbf{0}_\ell, \Pi)| \leq \ell^\ell.$$

Furthermore, we have

$$(31) \quad |\text{Part}(\ell)| \leq \ell^\ell$$

since there can be no more elements of  $\text{Part}(\ell)$  than functions  $\langle \ell \rangle \rightarrow \langle \ell \rangle$ .

## 5.2. Set partitions and the structure of functions of $\chi$ - and $\chi\chi$ -classes.

5.2.1. *Special classes of set partitions.* For  $\ell \in \mathbb{N}$  put

$$\begin{aligned} \text{Part}_\chi(\ell) &= \{\Pi \in \text{Part}(\ell) \mid \Pi \cap \{\{\lambda\} \mid \lambda \in \langle \ell \rangle\} = \emptyset\}, \\ \text{Part}_{\chi\chi}(2\ell) &= \{\Pi \in \text{Part}_\chi(2\ell) \mid \Pi \cap \{\{2\lambda - 1, 2\lambda\} \mid \lambda \in \langle \ell \rangle\} = \emptyset\}. \end{aligned}$$

In other words, (i)  $\text{Part}_\chi(\ell)$  consists of partitions of  $\langle \ell \rangle$  lacking singleton blocks and (ii)  $\text{Part}_{\chi\chi}(2\ell)$  consists of partitions of  $\langle 2\ell \rangle$  lacking singleton blocks and also lacking certain *forbidden doubleton blocks*, namely  $\{1, 2\}, \dots, \{2\ell - 1, 2\ell\}$ .

5.2.2. *Structure of functions of  $\chi$ -class.* A function  $f : \langle N \rangle^\ell \rightarrow \mathbb{C}$  is of  $\chi$ -class if and only if for all  $\mathbf{i} \in \langle N \rangle^\ell$  one has  $f(\mathbf{i}) = 0$  unless  $\Pi(\mathbf{i}) \in \text{Part}_\chi(\ell)$ . A similar if more complicated remark holding for functions of the  $\chi\chi$ -class will be formulated below as Lemma 5.2.4.

5.2.3. *Clump-equivalence.* We introduce some *ad hoc* and admittedly very ugly terminology. Fix  $\ell, N \in \mathbb{N}$ . Given  $\Pi \in \text{Part}(2\ell)$ , let

$$\text{Clump}(\Pi) = \bigcup_{B \in \Pi \setminus \{\{2\lambda - 1, 2\lambda\} \mid \lambda \in \langle \ell \rangle\}} B \subset \langle 2\ell \rangle.$$

Note that

$$\Pi \in \text{Part}_{\chi\chi}(2\ell) \Leftrightarrow (\Pi \in \text{Part}_\chi(2\ell) \ \& \ \text{Clump}(\Pi) = \langle 2\ell \rangle).$$

Given  $\mathbf{i} = (i_1, \dots, i_{2\ell}) \in \langle N \rangle^{2\ell}$  and  $\mathbf{j} = (j_1, \dots, j_{2\ell}) \in \langle N \rangle^{2\ell}$ , we say that  $\mathbf{i}$  and  $\mathbf{j}$  are *clump-equivalent* and write  $\mathbf{i} \sim \mathbf{j}$  if  $\Pi(\mathbf{i}) = \Pi(\mathbf{j})$  and  $i_\lambda = j_\lambda$  for  $\lambda \in \text{Clump}(\Pi(\mathbf{i}))$ .

**Lemma 5.2.4.** *Fix  $\ell, N \in \mathbb{N}$  such that  $N \geq 6\ell$ . Let  $F : \langle N \rangle^{2\ell} \rightarrow \mathbb{C}$  be a function of the  $\chi\chi$ -class. Then: (i) For  $\mathbf{i} \in \langle N \rangle^{2\ell}$  we have  $F(\mathbf{i}) = 0$  unless  $\Pi(\mathbf{i}) \in \text{Part}_\chi(2\ell)$ . (ii) Furthermore,  $F$  is constant on clump-equivalence classes.*

*Proof.* Statement (i) follows directly from the definition of the  $\chi$ -class. We turn now to the proof of statement (ii) which is tedious but only slightly more difficult. Fix  $\mathbf{i} = (i_1, \dots, i_{2\ell}) \in \langle N \rangle^{2\ell}$  arbitrarily and let  $\Pi = \Pi(\mathbf{i})$ . We have to show that  $F(\mathbf{i}) = F(\mathbf{i}')$  for  $\mathbf{i}' \in \langle N \rangle^{2\ell}$  such that  $\mathbf{i} \sim \mathbf{i}'$ . By statement (i) we may assume that  $\Pi \in \text{Part}_\chi(2\ell)$ , and since otherwise there is nothing to prove, we may assume  $\Pi \notin \text{Part}_{\chi\chi}(2\ell)$ . By symmetry we may assume that

$$\Pi = \Theta \cup \{\{2s + 1, 2s + 2\}, \dots, \{2\ell - 1, 2\ell\}\}$$

for some integer  $0 \leq s < \ell$  and set  $\Theta$  which is empty for  $s = 0$  and otherwise belongs to  $\text{Part}_{\chi\chi}(2s)$ . Let  $t = \ell - s$ . We have

$$(32) \quad \{\mathbf{i}' \in \langle N \rangle^{2\ell} \mid \mathbf{i} \sim \mathbf{i}'\} = \{(i_1, \dots, i_{2s}, k_1, k_1, \dots, k_t, k_t) \in \langle N \rangle^{2\ell} \\ \mid \text{distinct } k_1, \dots, k_t \in \langle N \rangle \setminus \{i_1, \dots, i_{2s}\}\}.$$

To finish the proof we have to show that  $F$  is constant on the set (32). Given

$$\mathbf{i} = (i_1, \dots, i_{2s}, k_1, k_1, \dots, k_t, k_t) \text{ and } \mathbf{i}' = (i_1, \dots, i_{2s}, k'_1, k'_1, \dots, k'_t, k'_t)$$

belonging to the set (32), since  $N \geq 6\ell$ , we can find a third element

$$\mathbf{i}'' = (i_1, \dots, i_{2s}, k''_1, k''_1, \dots, k''_t, k''_t)$$

of the set (32) such that

$$\{k''_1, \dots, k''_t\} \cap \{k_1, \dots, k_t, k'_1, \dots, k'_t\} = \emptyset.$$

We then have  $F(\mathbf{i}) = F(\mathbf{i}'') = F(\mathbf{i}')$  by  $2t$  applications of the definition of the  $\chi\chi$ -class.  $\square$

### 5.3. The Fibonacci-Whittle inequality.

5.3.1. *Fibonacci random variables.* We say that a real random variable  $\phi$  has the *Fibonacci distribution* if  $\phi^2 = \phi + 1$  almost surely and  $\mathbf{E}\phi = 0$ , in which case  $\phi$  necessarily has variance 1. (Equivalently:  $\Pr\left(\phi = \frac{1 \pm \sqrt{5}}{2}\right) = \frac{\sqrt{5} \mp 1}{2} / \sqrt{5}$ .) It follows also that

$$(33) \quad \|\phi\|_\infty = \frac{1 + \sqrt{5}}{2} \text{ and } \{\mathbf{E}\phi^k\}_{k=1}^\infty = (0, 1, 1, 2, 3, 5, 8, \dots).$$

The latter observation about moments justifies our reference to Fibonacci. We emphasize that while  $\mathbf{E}\phi = 0$ , we have  $\mathbf{E}\phi^k \geq 1$  for  $k \geq 2$ .

**Lemma 5.3.2.** *Fix  $\ell, N \in \mathbb{N}$ . Let  $\phi_1, \dots, \phi_N$  be i.i.d. Fibonacci random variables. For  $\mathbf{i} = (i_1, \dots, i_{2\ell}) \in \langle N \rangle^{2\ell}$  we have*

$$(34) \quad \mathbf{E} \prod_{\lambda=1}^{\ell} (\phi_{i_{2\lambda-1}} \phi_{i_{2\lambda}} - \delta_{i_{2\lambda-1}, i_{2\lambda}}) \geq \begin{cases} 1 & \text{if } \Pi(\mathbf{i}) \in \text{Part}_{\chi\chi}(2\ell), \\ 0 & \text{otherwise.} \end{cases}$$

*Proof.* Consider the product

$$\Phi = \prod_{\lambda=1}^{\ell} \begin{cases} \phi_{i_{2\lambda-1}} \phi_{i_{2\lambda}} & \text{if } i_{2\lambda-1} \neq i_{2\lambda}, \\ \phi_{i_{2\lambda}} & \text{if } i_{2\lambda-1} = i_{2\lambda}. \end{cases}$$

The left side of (34) equals  $\mathbf{E}\Phi$  by definition of a Fibonacci random variable. By (33) and independence of the factors  $\phi_i$  the quantity  $\mathbf{E}\Phi$  is a nonnegative integer. Since otherwise there is nothing to prove, assume for the rest of the proof that  $\Pi(\mathbf{i}) \in \text{Part}_{\chi\chi}(2\ell)$ . For each block  $A \in \Pi(\mathbf{i})$ , let  $i_A$  be the common value of  $i_\lambda$  for  $\lambda \in A$  and let  $m_A$  be the number of forbidden doubleton blocks contained in  $A$ . Then we have relations

$$\min_{A \in \Pi(\mathbf{i})} (|A| - m_A) \geq 2 \text{ and } \Phi = \prod_{A \in \Pi(\mathbf{i})} \phi_{i_A}^{|A| - m_A},$$

the former by definition of  $\text{Part}_{\chi\chi}(2\ell)$  and the latter by definition of  $\Phi$ . Since the family of random variables  $\{\phi_{i_A}\}_{A \in \Pi(\mathbf{i})}$  is i.i.d. Fibonacci, the desired inequality follows from (33).  $\square$

**Lemma 5.3.3.** *Let  $A \in \text{Mat}_N$  be a matrix. Let  $p \in [2, \infty)$  be an exponent. Let  $\phi_1, \dots, \phi_N$  be i.i.d. Fibonacci random variables. Then we have*

$$(35) \quad \left\| \sum_{i,j=1}^N A(i,j)(\phi_i\phi_j - \delta_{ij}) \right\|_p \leq K_p \|A\|_2$$

for a constant  $K_p$  depending only on  $p$ .

We refer to this inequality as the *Fibonacci-Whittle inequality*.

*Proof.* Let  $X_1, \dots, X_N$  be independent real random variables such that  $\|X_i\|_{2p} < \infty$  and  $\mathbf{E}X_i = 0$  for  $i = 1, \dots, N$ . The main result of the short paper [26], which is derived from the Khinchin inequality by a short elegant argument, gives a bound

$$(36) \quad \left\| \sum_{i,j=1}^N A(i,j)(X_iX_j - \mathbf{E}(X_iX_j)) \right\|_p \leq c_p \left( \sum_{i,j=1}^N |A(i,j)|^2 \|X_i\|_{2p}^2 \|X_j\|_{2p}^2 \right)^{1/2}$$

for a constant  $c_p$  depending only on  $p$ . (We note also that estimates similar in form to (36) are in common use in RMT. See for example [3, Lemma 2.7] which is more sophisticated but in the same spirit.) Upon specializing this result to the case  $X_i = \phi_i$  for  $i = 1, \dots, N$ , we obtain a bound of the desired form.  $\square$

We now take a long step toward completing the proof of Theorem 4.2.

**Proposition 5.4.** *Fix  $\ell, N \in \mathbb{N}$  along with matrices  $A_1, \dots, A_\ell \in \text{Mat}_N$ . (It is not necessary at this stage to assume that the matrices are of trace zero.) Put*

$$A(\mathbf{i}) = \prod_{\lambda=1}^{\ell} A_\lambda(i_{2\lambda-1}, i_{2\lambda}) \quad \text{for } \mathbf{i} = (i_1, \dots, i_{2\ell}) \in \langle N \rangle^{2\ell}.$$

We have

$$(37) \quad \sum_{\substack{\mathbf{i} \in \langle N \rangle^{2\ell} \text{ s.t.} \\ \Pi(\mathbf{i}) \in \text{Part}_{\chi\chi}(2\ell)}} |A(\mathbf{i})| \leq K_\ell^\ell \prod_{\lambda=1}^{\ell} \|A_\lambda\|_2$$

where  $K_\ell$  is the constant from line (35).

We hasten to point out that this result is not new. It can in principle (albeit with a different constant) be deduced as a corollary to Yin's lemma [25, Lemma 3.4]. (And for that matter, the Fibonacci-Whittle inequality above could also be deduced as a corollary to Yin's lemma.) We omit the derivation of (37) by way of Yin's lemma because the following derivation seems easier to us.

*Proof.* Since  $\text{Part}_{\chi\chi}(2) = \emptyset$ , we may assume that  $\ell \geq 2$ . Let  $\phi_1, \dots, \phi_N$  be i.i.d. Fibonacci random variables. We have

$$\begin{aligned} & \sum_{\substack{\mathbf{i} \in \langle N \rangle^{2\ell} \text{ s.t.} \\ \Pi(\mathbf{i}) \in \text{Part}_{\chi\chi}(2\ell)}} |A(\mathbf{i})| \leq \sum_{\mathbf{i}=(i_1, \dots, i_{2\ell}) \in \langle N \rangle^{2\ell}} |A(\mathbf{i})| \mathbf{E} \prod_{\lambda=1}^{\ell} (\phi_{i_{2\lambda-1}} \phi_{i_{2\lambda}} - \delta_{i_{2\lambda-1}, i_{2\lambda}}) \\ &= \mathbf{E} \prod_{\lambda=1}^{\ell} \left( \sum_{i \in \langle N \rangle} \sum_{j \in \langle N \rangle} |A_\lambda(i,j)| (\phi_i \phi_j - \delta_{ij}) \right) \leq K_\ell^\ell \prod_{\lambda=1}^{\ell} \|A_\lambda\|_2 \end{aligned}$$

at the first step by Lemma 5.3.2, at the second step by algebraic manipulation and at the last step by Lemma 5.3.3 and the Hölder inequality.  $\square$

One last technical but quite elementary observation is required before the endgame.

**Proposition 5.5.** *Fix  $\ell, N \in \mathbb{N}$ ,  $J \subset \langle N \rangle$  and  $f_1, \dots, f_\ell : \langle N \rangle \rightarrow \mathbb{C}$  such that*

$$\sum_{i \in \langle N \rangle} f_\lambda(i) = 0 \text{ for } \lambda \in \langle \ell \rangle.$$

Put

$$f(\mathbf{i}) = f_1(i_1) \cdots f_\ell(i_\ell) \text{ for } \mathbf{i} = (i_1, \dots, i_\ell) \in \langle N \rangle^\ell.$$

Then we have

$$(38) \quad \left| \sum_{\substack{\text{distinct} \\ i_1, \dots, i_\ell \in \langle N \rangle \setminus J}} f(i_1, \dots, i_\ell) \right| \leq \ell^{2\ell} \sum_{\mathbf{i} \in \langle N \rangle_{\chi, J}^\ell} |f(\mathbf{i})|,$$

where

$$\langle N \rangle_{\chi, J}^\ell = \{\mathbf{i} = (i_1, \dots, i_\ell) \in \langle N \rangle^\ell \mid \{\lambda\} \in \Pi(\mathbf{i}) \Rightarrow i_\lambda \in J \text{ for } \lambda \in \langle \ell \rangle\}.$$

In other words,  $\mathbf{i} = (i_1, \dots, i_\ell)$  belongs to  $\langle N \rangle_{\chi, J}^\ell$  provided that for every  $i \in \langle N \rangle$ , if  $i$  appears exactly once in  $\mathbf{i}$  then  $i \in J$ . In particular,  $\mathbf{i} \in \langle N \rangle_{\chi, \emptyset}^\ell$  if and only if  $\Pi(\mathbf{i}) \in \text{Part}_\chi(\ell)$ .

*Proof.* Using successively the Möbius inversion formula (29), the sum-product formula (28) and the hypothesis  $\sum_i f_\lambda(i) = 0$ , we obtain an algebraic identity

$$\begin{aligned} & \sum_{\substack{\text{distinct} \\ i_1, \dots, i_\ell \in \langle N \rangle \setminus J}} f(i_1, \dots, i_\ell) = \sum_{\substack{\mathbf{i} \in (\langle N \rangle \setminus J)^\ell \\ \text{s.t. } \Pi(\mathbf{i}) = \mathbf{0}_\ell}} f(\mathbf{i}) \\ &= \sum_{\Pi \in \text{Part}(\ell)} \mu(\mathbf{0}_\ell, \Pi) \sum_{\substack{\mathbf{i} \in (\langle N \rangle \setminus J)^\ell \\ \text{s.t. } \Pi \leq \Pi(\mathbf{i})}} f(\mathbf{i}) = \sum_{\Pi \in \text{Part}(\ell)} \mu(\mathbf{0}_\ell, \Pi) \prod_{B \in \Pi} \left( \sum_{i \in \langle N \rangle \setminus J} \prod_{b \in B} f_b(i) \right) \\ &= \sum_{\Pi \in \text{Part}(\ell)} \mu(\mathbf{0}_\ell, \Pi) \prod_{B \in \Pi} \begin{cases} \sum_{i \in \langle N \rangle \setminus J} \prod_{b \in B} f_b(i) & \text{if } B \text{ is not a singleton,} \\ - \sum_{i \in J} f_b(i) & \text{if } B = \{b\} \text{ is a singleton.} \end{cases} \end{aligned}$$

The desired estimate follows easily, via the crude bounds (30) and (31).  $\square$

**5.6. Completion of the proof of Theorem 4.2.** Let us recall the setup. We fix  $\ell, N \in \mathbb{N}$  and  $F : \langle N \rangle^{2\ell} \rightarrow \mathbb{C}$  of  $\chi\chi$ -class. We fix matrices  $A_1, \dots, A_\ell \in \text{Mat}_N$  of trace zero and put  $A(\mathbf{i}) = \prod_{\lambda=1}^\ell A_\lambda(i_{2\lambda-1}, i_{2\lambda})$  for  $\mathbf{i} = (i_1, \dots, i_{2\ell}) \in \langle N \rangle^{2\ell}$ . Without loss of generality, since (23) anyhow holds with  $C_\ell = (6\ell)^\ell$  for  $N < 6\ell$  (see Remark 4.3 above) we may assume that  $N \geq 6\ell$ . We now also fix a set partition  $\Pi \in \text{Part}_\chi(2\ell)$ . By Proposition 5.4 it will be enough to show that

$$(39) \quad \left| \sum_{\substack{\mathbf{i} \in \langle N \rangle^{2\ell} \text{ s.t.} \\ \Pi(\mathbf{i}) = \Pi}} F(\mathbf{i}) A(\mathbf{i}) \right| \leq \ell^{2\ell} \max_{\mathbf{i} \in \langle N \rangle^{2\ell}} |F(\mathbf{i})| \cdot \sum_{\substack{\mathbf{i} \in \langle N \rangle^{2\ell} \text{ s.t.} \\ \Pi(\mathbf{i}) \in \text{Part}_{\chi\chi}(2\ell)}} |A(\mathbf{i})|.$$

Since otherwise there is nothing to prove, we may assume that  $\Pi \in \text{Part}_\chi(2\ell) \setminus \text{Part}_{\chi\chi}(2\ell)$ . Finally, we may assume by symmetry that for some integers  $s \geq 0$  and  $t > 0$  summing to  $\ell$  and a set  $\Theta$  which is empty for  $s = 0$  and otherwise belongs to  $\text{Part}_{\chi\chi}(2s)$  we have

$$\Theta \cup \{\{2s+1, 2s+2\}, \dots, \{2\ell-1, 2\ell\}\} = \Pi.$$

We thus put ourselves once again in the situation contemplated in the proof of Lemma 5.2.4. We will consider the cases  $s = 0$  and  $s > 0$  separately below. The simple case  $s = 0$  is worth treating separately because it already reveals the mechanism at work in the general case.

Consider the case  $s = 0$ . We clearly have

$$(40) \quad \sum_{\substack{\mathbf{i} \in \langle N \rangle^{2\ell} \text{ s.t.} \\ \Pi(\mathbf{i}) = \Pi}} F(\mathbf{i})A(\mathbf{i}) = \sum_{\substack{\mathbf{k} \in \langle N \rangle^\ell \\ \text{s.t. } \Pi(\mathbf{k}) = \mathbf{0}_\ell}} F(\mathbf{k}^{(2)})A(\mathbf{k}^{(2)})$$

where for  $\mathbf{k} = (k_1, \dots, k_\ell) \in \langle N \rangle^\ell$  we let  $\mathbf{k}^{(2)} = (k_1, k_1, \dots, k_\ell, k_\ell) \in \langle N \rangle^{2\ell}$ . By Lemma 5.2.4 the value  $F(\mathbf{k}^{(2)})$  for  $\mathbf{k}$  as on the right side of (40) is a constant which we will denote by  $\overline{F}$ . We now calculate.

$$\begin{aligned} \left| \sum_{\substack{\mathbf{i} \in \langle N \rangle^{2\ell} \text{ s.t.} \\ \Pi(\mathbf{i}) = \Pi}} F(\mathbf{i})A(\mathbf{i}) \right| &= \left| \overline{F} \sum_{\substack{\mathbf{k} \in \langle N \rangle^\ell \\ \text{s.t. } \Pi(\mathbf{k}) = \mathbf{0}_\ell}} A(\mathbf{k}^{(2)}) \right| \\ &\leq t^{2t} \max_{\mathbf{i} \in \langle N \rangle^{2\ell}} |F(\mathbf{i})| \cdot \sum_{\mathbf{k} \in \langle N \rangle_{\chi, \emptyset}^\ell} |A(\mathbf{k}^{(2)})| \\ &\leq \ell^{2\ell} \max_{\mathbf{i} \in \langle N \rangle^{2\ell}} |F(\mathbf{i})| \cdot \sum_{\substack{\mathbf{i} \in \langle N \rangle^{2\ell} \text{ s.t.} \\ \Pi(\mathbf{i}) \in \text{Part}_{\chi\chi}(2\ell)}} |A(\mathbf{i})|. \end{aligned}$$

The first step is routine, the second step is an application of Proposition 5.5 and the last follows from the definitions of  $\langle N \rangle_{\chi, \emptyset}^\ell$  and  $\text{Part}_{\chi\chi}(2\ell)$ . Indeed, for  $\mathbf{k} \in \langle N \rangle_{\chi, \emptyset}^\ell$  one automatically has  $\Pi(\mathbf{k}) \in \text{Part}_\chi(\ell)$  and hence  $\Pi(\mathbf{k}^{(2)}) \in \text{Part}_{\chi\chi}(2\ell)$ .

Consider finally the case  $s > 0$ . Clearly we have

$$(41) \quad \sum_{\substack{\mathbf{i} \in \langle N \rangle^{2\ell} \text{ s.t.} \\ \Pi(\mathbf{i}) = \Pi}} F(\mathbf{i})A(\mathbf{i}) = \sum_{\substack{\mathbf{j} \in \langle N \rangle^{2s} \\ \text{s.t. } \Pi(\mathbf{j}) = \Theta}} \sum_{\substack{\mathbf{k} \in \langle N \rangle \setminus \{\mathbf{j}\}^t \\ \text{s.t. } \Pi(\mathbf{k}) = \mathbf{0}_t}} F(\mathbf{j}, \mathbf{k}^{(2)})A(\mathbf{j}, \mathbf{k}^{(2)})$$

where for  $\mathbf{j} = (j_1, \dots, j_{2s}) \in \langle N \rangle^{2s}$  we define  $\{\mathbf{j}\} = \{j_1, \dots, j_{2s}\}$  to abbreviate notation. Note that by Lemma 5.2.4, for  $(\mathbf{j}, \mathbf{k}^{(2)})$  as appearing on the right side of (41), the value  $F(\mathbf{j}, \mathbf{k}^{(2)})$  depends only on  $\mathbf{j}$  and hence we may write simply  $F(\mathbf{j}, \mathbf{k}^{(2)}) = \overline{F}(\mathbf{j})$ . We now calculate as before albeit with more indices to manage.

$$\begin{aligned}
\left| \sum_{\substack{\mathbf{i} \in \langle N \rangle^{2\ell} \\ \Pi(\mathbf{i}) = \Pi}} F(\mathbf{i}) A(\mathbf{i}) \right| &= \left| \sum_{\substack{\mathbf{j} \in \langle N \rangle^{2s} \\ \text{s.t. } \Pi(\mathbf{j}) = \Theta}} \bar{F}(\mathbf{j}) \sum_{\substack{\mathbf{k} \in (\langle N \rangle \setminus \{\mathbf{j}\})^t \\ \text{s.t. } \Pi(\mathbf{k}) = \mathbf{0}_t}} A(\mathbf{j}, \mathbf{k}^{(2)}) \right| \\
&\leq \max_{\mathbf{i} \in \langle N \rangle^{2\ell}} |F(\mathbf{i})| \cdot \sum_{\substack{\mathbf{j} \in \langle N \rangle^{2s} \\ \text{s.t. } \Pi(\mathbf{j}) = \Theta}} \left| \sum_{\substack{\mathbf{k} \in (\langle N \rangle \setminus \{\mathbf{j}\})^t \\ \text{s.t. } \Pi(\mathbf{k}) = \mathbf{0}_t}} A(\mathbf{j}, \mathbf{k}^{(2)}) \right| \\
&\leq t^{2t} \max_{\mathbf{i} \in \langle N \rangle^{2\ell}} |F(\mathbf{i})| \cdot \sum_{\substack{\mathbf{j} \in \langle N \rangle^{2s} \\ \text{s.t. } \Pi(\mathbf{j}) = \Theta}} \sum_{\mathbf{k} \in \langle N \rangle_{\chi, \{\mathbf{j}\}}^t} |A(\mathbf{j}, \mathbf{k}^{(2)})| \\
&\leq \ell^{2\ell} \max_{\mathbf{i} \in \langle N \rangle^{2\ell}} |F(\mathbf{i})| \cdot \sum_{\substack{\mathbf{i} \in \langle N \rangle^{2\ell} \\ \text{s.t.} \\ \Pi(\mathbf{i}) \in \text{Part}_{\chi\chi}(2\ell)}} |A(\mathbf{i})|.
\end{aligned}$$

The first two steps are routine. At the third step we apply Proposition 5.5. The fourth step is justified by the crucial observation that for  $\mathbf{j} \in \langle N \rangle^{2s}$  satisfying  $\Pi(\mathbf{j}) = \Theta$  and  $\mathbf{k} \in \langle N \rangle_{\chi, \{\mathbf{j}\}}^t$  we automatically have  $\Pi(\mathbf{j}, \mathbf{k}^{(2)}) \in \text{Part}_{\chi\chi}(2\ell)$ . The proof of (39) is complete and with it the proof of Theorem 4.2.  $\square$

## 6. PROOF OF PROPOSITION 2.6

**6.1. Brief recollection of notions of free probability.** For the reader's convenience, we review key definitions. For background see, e.g., [1, Chap. 5], [15] or [23]. Let  $I$  be a finite index set, and for each  $i \in I$  let  $J_i$  be a further finite index set.

**6.1.1. Noncommutative probability spaces.** A *noncommutative probability space* is a pair  $(\mathcal{A}, \phi)$  consisting of a unital algebra  $\mathcal{A}$  (invariably having  $\mathbb{C}$  as scalar field) and a linear functional  $\phi : \mathcal{A} \rightarrow \mathbb{C}$  called a *state* satisfying  $\phi(1_{\mathcal{A}}) = 1$ . The state  $\phi$  is called *tracial* if  $\phi(ab) = \phi(ba)$  for all  $a, b \in \mathcal{A}$ , in which case by abuse of language the space  $(\mathcal{A}, \phi)$  is also called *tracial*. (Often one assumes  $\mathcal{A}$  is a  $C^*$ -algebra and that  $\phi$  is a positive linear functional but we will not need such refinements.)

**6.1.2. Unital subalgebras and bracket notation.** Given a unital algebra  $\mathcal{A}$ , a subspace  $1_{\mathcal{A}} \in \mathcal{A}_0 \subset \mathcal{A}$  closed under multiplication is called a *unital subalgebra*. Given a unital algebra  $\mathcal{A}$  and a family of elements  $\{a_i\}_{i \in I}$  of  $\mathcal{A}$ , let  $\mathbb{C}\langle\{a_i\}_{i \in I}\rangle$  denote the unital subalgebra of  $\mathcal{A}$  generated by the family  $\{a_i\}_{i \in I}$ .

**6.1.3. Noncommutative polynomial algebras.** Let  $\mathbb{C}\langle\{\mathbf{X}_i\}_{i \in I}\rangle$  denote the noncommutative polynomial algebra generated by a family  $\{\mathbf{X}_i\}_{i \in I}$  of independent noncommutative variables  $\mathbf{X}_i$ . By definition the family of all monomials

$$(42) \quad \mathbf{X}_{i_1} \cdots \mathbf{X}_{i_r} \quad (r \in \{0, 1, 2, \dots\}, i_1, \dots, i_r \in I)$$

is a Hamel basis for  $\mathbb{C}\langle\{\mathbf{X}_i\}_{i \in I}\rangle$ . Multiplication in  $\mathbb{C}\langle\{\mathbf{X}_i\}_{i \in I}\rangle$  on the basis acts by juxtaposition and the empty monomial is identified with  $1_{\mathbb{C}\langle\{\mathbf{X}_i\}_{i \in I}\rangle}$ . Given a family  $\{a_i\}_{i \in I}$  of elements of a unital algebra  $\mathcal{A}$  and  $f = f(\{\mathbf{X}_i\}_{i \in I}) \in \mathbb{C}\langle\{\mathbf{X}_i\}_{i \in I}\rangle$ , the evaluation  $f(\{a_i\}_{i \in I}) \in \mathcal{A}$  is defined by substituting  $\mathbf{X}_i = a_i$  for all  $i \in I$ .

6.1.4. *Freeness.* Let  $(\mathcal{A}, \phi)$  be a noncommutative probability space. Let  $\{\mathcal{A}_i\}_{i \in I}$  be a family of unital subalgebras of  $\mathcal{A}$ . One says that the family  $\{\mathcal{A}_i\}_{i \in I}$  is *free* (with respect to  $\phi$ ) if the following statement holds for each integer  $\ell \geq 2$ .

$$(43) \quad \text{For all } i_1, \dots, i_\ell \in I \text{ such that } i_1 \neq i_2, \dots, i_{\ell-1} \neq i_\ell \text{ and all } a_1, \dots, a_\ell \in \mathcal{A} \\ \text{such that } a_j \in \mathcal{A}_{i_j} \text{ and } \phi(a_j) = 0 \text{ for } j = 1, \dots, \ell, \text{ one has } \phi(a_1 \cdots a_\ell) = 0.$$

A family  $\{a_i\}_{i \in I}$  of elements of  $\mathcal{A}$  is called *free* if the family of subalgebras  $\{\mathbb{C}\langle a_i \rangle\}_{i \in I}$  of  $\mathcal{A}$  is free in the sense just defined. More generally, a triangular array  $\{\{a_{ij}\}_{j \in J_i}\}_{i \in I}$  is said to be free (more precisely, to have rows free of each other) if the family of algebras  $\{\mathbb{C}\langle \{a_{ij}\}_{j \in J_i} \rangle\}_{i \in I}$  is free in the sense just defined.

6.1.5. *Joint laws.* Let  $(\mathcal{A}, \phi)$  be a noncommutative probability space and let  $\{a_i\}_{i \in I}$  be a family of elements of  $\mathcal{A}$ . The *joint law*  $\mu_{\{a_i\}_{i \in I}}$  of  $\{a_i\}_{i \in I}$  is the state on  $\mathbb{C}\langle \{\mathbf{X}_i\}_{i \in I} \rangle$  induced by the push-forward rule  $\mu_{\{a_i\}_{i \in I}}(f) = \phi(f(\{a_i\}_{i \in I}))$  for all  $f \in \mathbb{C}\langle \{\mathbf{X}_i\}_{i \in I} \rangle$ . Recall that in the presence of freeness, the joint law of a family  $\{a_i\}_{i \in I}$  of noncommutative random variables is uniquely determined by the laws of the individual variables  $a_i$ . More generally, if a triangular array  $\{\{a_{ij}\}_{j \in J_i}\}_{i \in I}$  has rows freely independent of each other, then the joint law of the array is uniquely determined by the joint laws of the rows  $\{a_{ij}\}_{j \in J_i}$  for  $i \in I$ .

## 6.2. Generalities concerning asymptotic freeness.

6.2.1. *Convergence in moments.* Given a collection  $\{\mu\} \cup \{\mu^{(N)}\}_{N=1}^\infty$  of states on the unital algebra  $\mathbb{C}\langle \{\mathbf{X}_i\}_{i \in I} \rangle$ , we write  $\lim \mu^{(N)} = \mu$  if  $\lim_{N \rightarrow \infty} \mu^{(N)}(f) = \mu(f)$  for all  $f \in \mathbb{C}\langle \{\mathbf{X}_i\}_{i \in I} \rangle$  and we say that  $\{\mu^{(N)}\}$  converges *in moments* to  $\mu$ . (Some authors on free probability call such convergence *weak* but we refrain from doing so here since the analogue in classical probability is more nearly convergence in moments.) Note that if  $\lim_{N \rightarrow \infty} \mu^{(N)} = \mu$  and all the  $\mu^{(N)}$  are tracial, then so is  $\mu$ .

6.2.2. *Asymptotic freeness.* Let  $\{(\mathcal{A}^{(N)}, \phi^{(N)})\}_{N=1}^\infty$  be a sequence of noncommutative probability spaces. For each  $N$  let there be given a family  $\{a_i^{(N)}\}_{i \in I}$  of elements of  $\mathcal{A}^{(N)}$ . We say that the sequence  $\{\{a_i^{(N)}\}_{i \in I}\}_{N=1}^\infty$  is *convergent in law* if the sequence of joint laws  $\mu_{\{a_i^{(N)}\}_{i \in I}}$  is convergent in moments. We furthermore say  $\{\{a_i^{(N)}\}_{i \in I}\}_{N=1}^\infty$  is *asymptotically free* if convergent in law and the family  $\{\mathbf{X}_i\}_{i \in I}$  is free with respect to the limit law on  $\mathbb{C}\langle \{\mathbf{X}_i\}_{i \in I} \rangle$ . The preceding definitions can be generalized to triangular arrays in evident fashion.

**Lemma 6.2.3.** *For each positive integer  $N$  let*

$$\mu^{(N)} : \mathbb{C}\langle \{\{\mathbf{X}_{ij}\}_{j \in J_i}\}_{i \in I} \rangle \rightarrow \mathbb{C}$$

*be a tracial state with marginal states*

$$\tau_i^{(N)} = \mu^{(N)}|_{\mathbb{C}\langle \{\mathbf{X}_{ij}\}_{j \in J_i} \rangle} : \mathbb{C}\langle \{\mathbf{X}_{ij}\}_{j \in J_i} \rangle \rightarrow \mathbb{C} \text{ for } i \in I.$$

*Assume the following statement holds.*

$$(44) \quad \text{The limit } \tau_i = \lim_{N \rightarrow \infty} \tau_i^{(N)} \text{ in moments exists for each } i \in I.$$

Assume furthermore that for each integer  $\ell \geq 2$  the following statement holds.

$$(45) \quad \begin{aligned} & \text{For all } i_1, \dots, i_\ell \in I \text{ such that } i_1 \neq i_2, i_2 \neq i_3, \dots, i_{\ell-1} \neq i_\ell, \\ & \text{and for all } f_1 \in \mathbb{C}\langle \{\mathbf{X}_{i_1, j}\}_{j \in J_{i_1}} \rangle, \dots, f_\ell \in \mathbb{C}\langle \{\mathbf{X}_{i_\ell, j}\}_{j \in J_{i_\ell}} \rangle, \text{ one has} \\ & \lim_{N \rightarrow \infty} \mu^{(N)}((f_1 - \mu^{(N)}(f_1)) \cdots (f_\ell - \mu^{(N)}(f_\ell))) = 0. \end{aligned}$$

Then the limit  $\mu = \lim_{N \rightarrow \infty} \mu^{(N)}$  in moments exists and is tracial. Furthermore, with respect to  $\mu$ , the rows of the triangular array  $\{\{\mathbf{X}_{ij}\}_{j \in J_i}\}_{i \in I}$  are free of each other and for each  $i \in I$  the joint law of each row  $\{\mathbf{X}_{ij}\}_{j \in J_i}$  is  $\tau_i$ .

*Proof.* Consider the following statement for integers  $\ell \geq 2$ .

$$(46) \quad \begin{aligned} & \text{For all } i_1, \dots, i_\ell \in I \text{ s.t. } i_1 \neq i_2, \dots, i_{\ell-1} \neq i_\ell \\ & \text{and } f_1 \in \mathbb{C}\langle \{\mathbf{X}_{i_1, j}\}_{j \in J_{i_1}} \rangle, \dots, f_\ell \in \mathbb{C}\langle \{\mathbf{X}_{i_\ell, j}\}_{j \in J_{i_\ell}} \rangle, \\ & \text{the sequence } \{\mu^{(N)}(f_1 \cdots f_\ell)\} \text{ is convergent and moreover} \\ & \text{converges to 0 if } \lim_{N \rightarrow \infty} \mu^{(N)}(f_j) = 0 \text{ for } j = 1, \dots, \ell. \end{aligned}$$

In view of (44) and the last remark of §6.2.1, it is enough to verify (46) for each  $\ell \geq 2$ . For  $I \subset \langle \ell \rangle$  let  $f_I$  be the product of  $f_1, \dots, f_\ell$  after striking from the latter list the entries indexed by  $i \in I$  and let  $\pi_I^{(N)} = \prod_{i \in I} \mu^{(N)}(f_i)$ . We then have an algebraic identity

$$\begin{aligned} & \mu^{(N)}(f_1 \cdots f_\ell) \\ &= \mu^{(N)}((f_1 - \mu^{(N)}(f_1)) \cdots (f_\ell - \mu^{(N)}(f_\ell))) - \sum_{\emptyset \neq I \subset \langle \ell \rangle} (-1)^{|I|} \pi_I^{(N)} \mu^{(N)}(f_I). \end{aligned}$$

Hypotheses (44) and (45), the identity above and induction on  $\ell$  when  $\ell \geq 3$  imply that (46) holds.  $\square$

**6.3. Completion of the proof of Proposition 2.6.** It will be enough to show that the sequence  $\{\mu^{(N)}\}_{N=1}^\infty$  satisfies the hypotheses of Lemma 6.2.3. At any rate, it is clear that each state  $\mu^{(N)}$  is tracial and that hypothesis (44) holds for the sequence  $\{\mu^{(N)}\}$  by hypothesis (5) of Proposition 2.6. Only hypothesis (45) need be checked. To that end, fix  $i_1, \dots, i_\ell$  and  $f_1, \dots, f_\ell$  as they appear in (45). For  $\lambda = 1, \dots, \ell$  put

$$A_\lambda^{(N)} = f_\lambda(\{T_{i_\lambda, j}^{(N)}\}_{j \in J_{i_\lambda}}) - \frac{1}{N} \text{tr} f_\lambda(\{T_{i_\lambda, j}^{(N)}\}_{j \in J_{i_\lambda}}) \mathbf{I}_N.$$

We have

$$\text{tr} A_\lambda^{(N)} = 0 \quad \text{and} \quad \sup_N \max_\lambda \left[ \left[ A_\lambda^{(N)} \right] \right] < \infty,$$

the former assertion being clear and the latter following from hypothesis (4) of Proposition 2.6. We have an algebraic identity

$$N \mu^{(N)}((f_1 - \mu^{(N)}(f_1)) \cdots (f_\ell - \mu^{(N)}(f_\ell))) = \mathbf{E} \text{tr}(U_{i_1}^{(N)} A_1^{(N)} U_{i_1}^{(N)*} \cdots U_{i_\ell}^{(N)} A_\ell^{(N)} U_{i_\ell}^{(N)*})$$

following immediately from the definitions. Finally, by hypothesis (6) of Proposition 2.6 along with Proposition 2.4, one has a bound

$$\left| \mathbf{E} \text{tr}(U_{i_1}^{(N)} A_1^{(N)} U_{i_1}^{(N)*} \cdots U_{i_\ell}^{(N)} A_\ell^{(N)} U_{i_\ell}^{(N)*}) \right| \leq c(i_1, \dots, i_\ell) \prod_{\lambda=1}^{\ell} \left[ \left[ A_\lambda^{(N)} \right] \right].$$

Thus (45) indeed holds. The proof of Proposition 2.6 is complete.  $\square$

## 7. CONCENTRATION

While the asymptotic liberating property of random unitary matrices  $U^{(N)}$  that are invariant under conjugation by a signed permutation matrix allows us to determine the limiting expected eigenvalue distribution functions of matrices of the form

$$A_N + U^{(N)}B_NU^{(N)*} \quad \text{and} \quad A_NU^{(N)}B_NU^{(N)*}A_N^*$$

for suitable sequences of matrices  $A_N, B_N \in \text{Mat}_N$ , the concentration of the empirical eigenvalue distribution of such matrices is also of interest. Chatterjee obtained the first results of this kind when  $U^{(N)}$  are Haar distributed [4, 5], and below we show that the analogous results hold for the random unitary matrices addressed here. In particular, Theorem 7.1 is analogous to Theorem 1.1 in [4].

**Theorem 7.1.** *Assume that  $U^{(N)} \in \text{Mat}_N$  is a random unitary matrix and that  $WU^{(N)}W^* \stackrel{d}{=} U^{(N)}$  for any deterministic signed permutation matrix  $W \in \text{Mat}_N$ . Let  $A, B \in \text{Mat}_N$  be arbitrary Hermitian matrices and set*

$$H_+ = A + U^{(N)}BU^{(N)*}$$

and

$$H_\times = AU^{(N)}BU^{(N)*}A^*.$$

*There exists a constant  $c > 0$ , independent of  $A, B$  and  $N$ , such that for all  $x \in \mathbb{R}$ ,  $\text{Var}(F_{H_+}(x)) \leq 64 \frac{\log N + c}{N}$ . Further,*

$$\Pr(|F_{H_+}(x) - \mathbf{E}F_{H_+}(x)| > t) \leq 2 \exp\left(\frac{-Nt^2}{128(\log N + c)}\right)$$

*for all  $t > 0$  for the same constant  $c$ . The analogous two bounds hold as well for  $H_\times$ .*

Recall that if  $X$  and  $Y$  are two random variables taking values in the same space  $\mathcal{X}$ , then the *total variation distance*  $d_{TV}(X, Y)$  is defined by

$$d_{TV}(X, Y) = \sup_{B \in \mathcal{X}; B \text{ is Borel}} |\Pr(X \in B) - \Pr(Y \in B)|.$$

*Proof.* We construct a *random signed transposition matrix* in the following way. Let  $J$  and  $K$  be independent and uniformly distributed on  $\{1, \dots, N\}$ , let  $\{e_1, \dots, e_N\}$  denote the standard basis vectors for  $\mathbb{C}^N$  and let  $\epsilon_1, \epsilon_2$  be independent random variables taking the values  $\{\pm 1\}$  with equal probability. Define  $T \in \text{Mat}_N$  by

$$T = \begin{cases} \epsilon_1 e_J e_J^* + \sum_{i \neq J} e_i e_i^* & J = K \\ \epsilon_1 e_J e_K^* + \epsilon_2 e_K e_J^* + \sum_{i \neq J, i \neq K} e_i e_i^* & J \neq K. \end{cases}$$

That is,  $T$  either does not transpose any entries and multiplies a random entry by  $\pm 1$  or does transpose two entries and multiplies them by  $\pm 1$ .

We let  $W_r \in \text{Mat}_N$  denote the random matrix with uniform distribution on the set of signed permutation matrices in  $\text{Mat}_N$ . Set

$$H'_+ = A + W_r U^{(N)} W_r^* B W_r U^{(N)} W_r^*,$$

so that  $H_+ \stackrel{d}{=} H'_+$ , which implies that it suffices to prove the theorem for  $H'_+$  rather than  $H_+$ . Now set

$$H''_+ = A + T W_r U^{(N)} W_r^* T B T W_r U^{(N)} W_r^* T,$$

and note that  $H_+'' \stackrel{d}{=} H_+''$ . Set  $D = TW_r - W_r$ , so that

$$\begin{aligned} H_+' - H_+'' &= W_r U^{(N)} W_r^* B W_r U^{(N)} W_r^* - T W_r U^{(N)} W_r^* T B T W_r U^{(N)} W_r^* T \\ &= (D + W_r) U^{(N)} (D + W_r^*) B (D + W_r) U^{(N)} D \\ &\quad + (D + W_r) U^{(N)} (D + W_r^*) B D U^{(N)} W_r^* \\ &\quad + (D + W_r) U^{(N)} D B W_r U^{(N)} W_r^* + D U^{(N)} D B D U^{(N)} W_r^*. \end{aligned}$$

Since  $D$  has rank at most 2, each term in the sum above is of rank at most 2. Thus,  $\text{rank}(H_+' - H_+'' ) \leq 8$ . Following [4], we apply Lemma 2.2 of [2] to obtain

$$\|F_{H_+'} - F_{H_+''}\|_\infty \leq \frac{8}{N}.$$

For any fixed  $x \in \mathbb{R}$ , we define the function  $f$  which maps  $W_r$  to  $F_{H_+'}(x)$ . By (7),

$$|f(W) - f(TW)| \leq \frac{8}{N}$$

for any realizations of  $T, W$  and  $U^{(N)}$ ; in particular,

$$(47) \quad (\mathbf{E}(f(W) - f(TW))^2)^{1/2} \leq \frac{8}{N}$$

for all signed permutation matrices  $W \in \text{Mat}_N$ . In order to refer to the main theorem of [4] we require a bound on  $d_{TV}(T_1 \cdots T_k, W_r)$ . Here we apply Theorem 3.1.3 of [19] (which builds on Theorem 1 of [11]) to obtain

$$(48) \quad d_{TV}(T_1 \cdots T_k, W_r) \leq a n e^{-2k/N}$$

for all  $k \geq 1$  for a constant  $a > 0$ . (By choosing  $a$  large enough we satisfy the condition on  $k$  stated in [19].) Now the claim for  $H_+$  follows by using (7), (47) and (48) in Theorem 1.2 of [4].

The proof for  $H_\times$  follows in exactly the same way since  $H_\times' - H_\times''$ , defined analogously to  $H_+'$  and  $H_+''$ , also has rank at most 8.  $\square$

## 8. CONCLUDING REMARKS

If  $A^{(N)}, B^{(N)} \in \text{Mat}_N$  are a sequence of self-adjoint matrices whose spectral distribution converges in any of several ways to a measure, then the empirical spectral distribution of the sum and conjugated matrix defined in Theorem 7.1 also converges in a corresponding way. It remains to provide a bound of some form on the distance between the spectral distribution of the resulting random matrix in dimension  $N$  and the limiting distribution. For other forms of random matrices one may use equations for the trace of the resolvent to prove such behavior. This can be seen in [12] for Wigner matrices, for example. The authors of [17, 22] prove the existence of implicit equations as the dimension of the matrices tends to infinity, yet for both the Haar case and the case presented here, these equations have, in general, not yet been developed for finite dimensions. The exception is [13], where the very special case of coordinate projection matrices was addressed. Coordinate projections, however, are far simpler than the general case.

It is possible that using random signed permutation matrices and complex Hadamard matrices offers a new approach to study the convergence of the empirical spectrum to its limiting behavior. In particular, one can embed an element of  $S_N$  into  $S_{N+1}$  by inserting the index  $N + 1$  at a random location. If  $F_N$  denotes the empirical spectral distribution of the random matrix model described above in dimension  $N$ , then one can look at  $\|\mathbf{E}F_N - \mathbf{E}F_{N+1}\|_\infty$ . If this quantity is sufficiently small, then a bound on the distance between the empirical distribution and the limiting distribution will follow from Theorem 7.1.

Further questions include the following. What other families of random unitary matrices are asymptotically liberating? Do there exist families of unitary matrices that are not asymptotically liberating, yet yield sums and products that behave as if they are free?

Finally, Corollaries 3.2 and 3.7 suggest the interesting problem of proving a ‘‘Hadamard analogue’’ of the strong convergence result of [10]. New tools may have to be developed to solve this problem since there is no obvious analogue of the slick method used in [10].

## REFERENCES

- [1] G. Anderson, A. Guionnet, O. Zeitouni, *An Introduction to Random Matrices*, Cambridge Studies in Advanced Mathematics vol. 118, Cambridge University Press, Cambridge 2010. MR2760897
- [2] Z. D. Bai. Methodologies in spectral analysis of large-dimensional random matrices, a review, *Statist. Sinica*, 9(3):611–677, 1999. MR1711663
- [3] Z. D. Bai and J. W. Silverstein, No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices, *Ann. Probab.* 26 (1998) 316–345. MR1617051
- [4] S. Chatterjee, Concentration of Haar measures, with an application to random matrices, *J. Funct. Anal.* 245 (2007) 379–389. MR2309833
- [5] S. Chatterjee, *Concentration Inequalities with Exchangeable Pairs*, PhD Thesis, Stanford University, 2008.
- [6] Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, third ed., Springer Texts in Statistics, Springer-Verlag, New York, 1997. MR1476912
- [7] B. Collins, Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability, *Int. Math. Res. Not.* 2003, no. 17, 953–982. MR1959915
- [8] B. Collins, Product of random projections, Jacobi ensembles and universality problems arising from free probability, *Probab. Theory Related Fields* 133 (2005) 315–344. MR2198015
- [9] B. Collins, T. Kemp, Liberation of Projections, [arXiv:1211.6037v1](https://arxiv.org/abs/1211.6037v1)
- [10] B. Collins, C. Male, The strong asymptotic freeness of Haar and deterministic matrices. [arXiv:1105.4345v1](https://arxiv.org/abs/1105.4345v1)
- [11] P. Diaconis and M. Shahshahani, Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, 57(1981), 159–179. MR0626813
- [12] L. Erdős, B. Schlein, H.-T. Yau, Local semicircle law and complete delocalization for Wigner random matrices, *Comm. Math. Phys.* 287(2009) 641–655. MR2481753
- [13] B. Farrell, Limiting empirical singular value distribution of restrictions of discrete Fourier transform matrices, *J. Fourier Anal. Appl.* 17(2011) 733–753. MR2819175
- [14] C. Male, The distributions of traffics and their free product: an asymptotic freeness theorem for random matrices and a central limit theorem, [arXiv:1111.4662v3](https://arxiv.org/abs/1111.4662v3)
- [15] A. Nica, R. Speicher, *Lectures on the combinatorics of free probability*, London Mathematical Society Lecture Note Series 335, Cambridge University Press, Cambridge, 2006. MR2266879
- [16] A. Nica, Asymptotically free families of random unitaries in symmetric groups, *Pacific J. Math.* 157 (1993) 295–310. MR1197059
- [17] L. Pastur and V. Vasilchuk, On the law of addition of random matrices, *Comm. Math. Phys.*, 214(2000), 249–286. MR1796022.

- [18] G.-C. Rota, T. Wallstrom, Stochastic integrals: a combinatorial approach. *Ann. Probab.* 25(1997) 1257–1283. MR1457619
- [19] C. H. Schoolfield, Jr., Random walks on wreath products of groups, *J. Theoret. Probab.*, 15(2002) 667–693. MR1922442
- [20] J. A. Tropp, Improved analysis of the subsampled randomized Hadamard transform, *Adv. Adapt. Data Anal.* 3(2011) 115–126. MR2835584.
- [21] 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), no. 3, 1187–1215. MR2723670
- [22] V. Vasilchuk, On the law of multiplication of random matrices, *Math. Phys. Anal. Geom.*, 4(2001) 1–36. MR1855781
- [23] D. V. Voiculescu, K. Dykema, A. Nica, Free random variables, A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups, CRM Monograph Series 1. American Mathematical Society, Providence, RI, 1992. MR1217253
- [24] D. Voiculescu, Limit laws for random matrices and free products, *Invent. Math.* 104(1991), 201–220. MR1094052
- [25] Y. Q. Yin, Limiting spectral distribution for a class of random matrices, *J. Multivariate Anal.* 20(1986) 50–68. MR0862241
- [26] P. Whittle, Bounds for the moments of linear and quadratic forms in independent variables, *Teor. Veroyatnost. i Primenen.* 5(1960) 331–335, transl. *Theory Probab. Appl.* 5(1960) 303–305. MR0133849

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN 55455, U.S.A.  
*E-mail address:* gwanders@umn.edu

COMPUTING AND MATHEMATICAL SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA 91125, U.S.A.  
*E-mail address:* farrell@cms.caltech.edu