

# Strong Convergence: A Short Survey\*

Ramon van Handel<sup>†</sup>

**Abstract.** A family of random matrices is said to converge strongly to a limiting family of operators if the operator norm of every noncommutative polynomial of the matrices converges to that of the limiting operators. Recent developments surrounding the strong convergence phenomenon have led to new progress on important problems in random graphs, geometry, operator algebras, and applied mathematics. We review classical and recent results in this area, and their applications to various areas of mathematics.

**1 Introduction.** Throughout this survey, we denote by  $\mathbb{C}^*\langle x_1, \dots, x_r \rangle$  the  $*$ -algebra of noncommutative polynomials  $P$  in the free variables  $x_1, \dots, x_r$  and their adjoints; for example,

$$P(x, y, z) = 2xy^*x + (1+i)z - \pi z^3 x^*y.$$

For simplicity, we refer to any such polynomial as a  $*$ -polynomial. A  $*$ -polynomial  $P(x_1, \dots, x_r)$  defines a bounded operator whenever bounded operators are substituted for  $x_1, \dots, x_r$ .

**DEFINITION 1.1** (Strong convergence). *Let  $\mathbf{X}^N = (X_1^N, \dots, X_r^N)$  be a family of random matrices for every  $N \geq 1$ , and let  $\mathbf{x} = (x_1, \dots, x_r)$  be a family of bounded operators on a Hilbert space. If*

$$\lim_{N \rightarrow \infty} \|P(\mathbf{X}^N)\| = \|P(\mathbf{x})\| \quad \text{in probability}$$

*for every  $*$ -polynomial  $P$ , then  $\mathbf{X}^N$  is said to converge strongly to  $\mathbf{x}$ .*

This innocent looking definition belies the fact that it is an extremely strong property of random matrices, since it must hold for *every*  $*$ -polynomial  $P$ . It was observed by Voiculescu in 1993 [93] that the existence of any model (deterministic or random) that strongly converges to a free limiting model would resolve a long-standing conjecture in the theory of  $C^*$ -algebras; see Section 4.3.1 below. It was a major breakthrough when Haagerup and Thorbjørnsen proved for the first time, more than a decade later [47], that such a random matrix model exists. The title of their 2005 paper, “*A new application of random matrices ...*” foreshadowed a series of unexpected and wide-ranging developments that are the subject of this survey.

In recent years, the notion of strong convergence has led to significant progress on important problems in several different areas of mathematics, including random graphs, hyperbolic surfaces, minimal surfaces, operator algebras, and applied mathematics. These new applications of strong convergence have gone hand in hand with the development of new methods of random matrix theory, which made it possible to establish strong convergence in challenging situations that remained well out of reach until very recently.

The aim of this survey is to review these and related developments surrounding strong convergence. We begin in Section 2 by providing an overview of random matrix models that have been shown to converge strongly, and of the main methods of proof that are used for this purpose. These results are concerned with concrete random matrix models that converge asymptotically to a limiting set of operators as in Definition 1.1. In Section 3, we discuss a surprising nonasymptotic complement to such results: under mild conditions, “almost any” random matrix behaves like a suitable limiting operator for strong convergence, even if it does not arise as in Definition 1.1. This is especially useful in applied mathematics, where it is often necessary to consider random matrices that have an arbitrary structure. Section 4 discusses a wide variety of applications of strong convergence to random graphs, geometry, operator algebras, and more. Finally, Section 5 discusses in more detail a new technique, the polynomial method, which has been instrumental in several recent developments.

---

\*Contribution to the Proceedings of the International Congress of Mathematicians, 2026.

<sup>†</sup>Department of Mathematics, Princeton University, Princeton, NJ 08544, USA (rvan@math.princeton.edu).

The focus of this short survey is twofold: we aim to convey the breadth of the subject, and to highlight some recent developments in this area that arise from work of the author and coauthors (especially in Sections 3 and 5). A more extensive mathematical introduction, and a more detailed treatment of some applications and open problems, is given in [92]. The survey of Magee [65], which is focused on the interactions between strong convergence, representation theory, and geometry, is highly recommended for a complementary perspective.

## 2 Strong convergence.

**2.1 Limiting models.** Before we can discuss what is known about strong convergence, we must first introduce the limiting models that random matrices converge to. Essentially all known strong convergence results may be viewed as arising, directly or indirectly, from the following classical construction.

Let  $\mathbf{G}$  be a finitely generated group. For every  $g \in \mathbf{G}$ , define the operator

$$\lambda(g)\delta_w = \delta_{gw}$$

on  $l^2(\mathbf{G})$ , where  $\delta_w$  denotes the standard basis vector associated to  $w \in \mathbf{G}$  (that is, the function in  $l^2(\mathbf{G})$  that equals one at  $w$  and zero elsewhere). Then  $\lambda : \mathbf{G} \rightarrow B(l^2(\mathbf{G}))$  defines the *regular representation* of  $\mathbf{G}$ . Note that, by construction,  $\lambda(g)^* = \lambda(g^{-1})$  and that  $\lambda(g)$  is a unitary operator.

DEFINITION 2.1. Let  $\mathbf{F}_r$  be the free group with free generators  $g_1, \dots, g_r$ , and define

$$u_k = \lambda(g_k).$$

Then the operators  $u_1, \dots, u_r$  on  $l^2(\mathbf{F}_r)$  are called free Haar unitaries.

By construction, free Haar unitaries are algebraically free, i.e., they satisfy no algebraic relations. One may therefore expect such operators to arise as the limiting model of “generic” families of random unitary matrices, since such random matrices are increasingly unlikely to satisfy any relation of fixed length as their dimension goes to infinity. As we will shortly see, this is indeed the case.

To motivate the analogous limit model for self-adjoint random matrices, we recall a ubiquitous observation in random matrix theory: the spectral properties of many self-adjoint random matrices behave as those of the classical gaussian ensembles. This suggests we should aim to define a free limiting model that captures the properties of gaussian distributions. This idea is made precise by Voiculescu’s free probability theory, where the free analogue of independent gaussian random variables is provided by a *free semicircular family*  $s_1, \dots, s_r$ . For a precise definition, we refer to the excellent text [77]. Such families can be constructed in several ways: for example, they can be obtained as  $s_k = \Phi(u_k + u_k^*)$ , where  $u_1, \dots, u_r$  are free Haar unitaries and  $\Phi$  is a suitably chosen continuous function (see, e.g., the proof of [44, Theorem 2.4]); an often more useful construction arises from the creation and annihilation operators on the free Fock space, see [77, pp. 102–108].

An important feature of free Haar unitaries and free semicircular families is not only that they describe the limiting behavior of many random matrix models, but also that free probability theory provides a powerful toolbox for explicitly computing the spectra of polynomials of such matrices. A simple example

$$(2.1) \quad \|u_1 + u_1^* + \dots + u_r + u_r^*\| = 2\sqrt{2r-1}$$

for free Haar unitaries  $u_1, \dots, u_r$  is a classical result of Kesten [58], since the operator  $u_1 + u_1^* + \dots + u_r + u_r^*$  may be recognized as the adjacency operator of the infinite  $2r$ -regular tree. However, in principle  $\|P(u_1, \dots, u_r)\|$  can be computed for any  $*$ -polynomial  $P$  by means of a variational principle due to Lehner [61]. Analogous computations can be done for free semicircular families as well (see Section 3).

While free Haar unitaries and free semicircular families are based on the free group  $\mathbf{F}_r$ , models based on non-free discrete groups  $\mathbf{G}$  are of great interest; see Section 2.3 below.

**2.2 Strong asymptotic freeness.** From a probabilistic perspective, the most natural way to choose a family of random matrices is to sample them *independently* from a given ensemble. As long as the ensemble is “sufficiently random”, it is highly unlikely that independent random matrices will satisfy any fixed relation, and one therefore expects such models to behave freely. This is indeed the case.

The first results in this direction were obtained by Haagerup and Thorbjørnsen [47] and by Schultz [89] for the classical gaussian ensembles: that is, the GUE/GOE/GSE models of  $N \times N$  self-adjoint gaussian random matrices whose law is invariant under unitary/orthogonal/symplectic conjugation.

**THEOREM 2.2** (Haagerup–Thorbjørnsen; Schultz). *Let  $\mathbf{X}^N = (X_1^N, \dots, X_r^N)$  be i.i.d. GUE/GOE/GSE matrices of dimension  $N$ , and  $\mathbf{s} = (s_1, \dots, s_r)$  be a free semicircular family. Then  $\mathbf{X}^N$  converges strongly to  $\mathbf{s}$ .*

Theorem 2.2 was subsequently extended to much more general random matrix models:

- Building on the methods developed in [47, 89], it was shown by Anderson [3] that the same conclusion holds if  $X_1^N, \dots, X_r^N$  are independent Wigner matrices, that is, self-adjoint random matrices that have arbitrary (non-gaussian) i.i.d. entries with bounded fourth moment on and above the diagonal.
- Using new methods discussed in Section 3, Bandeira, Boedihardjo, and the author [7, Theorem 2.10] showed that the same conclusion holds for *any* independent  $N \times N$  self-adjoint random matrices  $X_1^N, \dots, X_r^N$  with jointly gaussian entries, assuming only that  $\|\mathbf{E}[X_k^N]\| = o(1)$ ,  $\|\mathbf{E}[(X_k^N)^2] - \mathbf{1}\| = o(1)$ , and  $\|\text{Cov}(X_k^N)\| = o((\log N)^{-3/2})$  (where  $\text{Cov}(X)$  denotes the covariance matrix of the entries of  $X$ ). These mild assumptions are satisfied even by nonhomogeneous and dependent models, such as random band matrices with polylogarithmic band width. An extension to many non-gaussian models appears in [21].

Further extensions include strong convergence of random matrices interacting through a potential [43]; strong convergence to operator-valued semicircular families [57]; joint strong convergence of deterministic and self-adjoint random matrices [71, 14]; and strong quantitative forms of Theorem 2.2 [32, 81, 82, 27].

We now turn to strong convergence of random unitary matrices. It was observed by Haagerup and Thorbjørnsen [47, Lemma 8.1] that, since one can construct free Haar unitaries  $u_1, \dots, u_r$  as  $u_k = \Psi(s_k)$  where  $s_1, \dots, s_r$  is a free semicircular family and  $\Psi$  is a suitably defined continuous function, one can obtain a model of random unitary matrices that strongly converges to free Haar unitaries by applying  $\Psi$  to a family of independent GUE matrices. This suffices for certain applications, but yields a random matrix model with some unusual properties [47, Remark 8.3]. The following result, which was subsequently obtained by Collins and Male [33], may be viewed as the natural counterpart of Theorem 2.2 for random unitary matrices.

**THEOREM 2.3** (Collins–Male). *Let  $\mathbf{U}^N = (U_1^N, \dots, U_r^N)$  be i.i.d. Haar-distributed random matrices in the groups  $\text{U}(N)/\text{O}(N)/\text{Sp}(N)$ , and  $\mathbf{u} = (u_1, \dots, u_r)$  be free Haar unitaries. Then  $\mathbf{U}^N$  converges strongly to  $\mathbf{u}$ .*

To prove this theorem, Collins and Male introduce a simple construction that makes it possible to deduce Theorem 2.3 from Theorem 2.2. Subsequent works have developed new techniques that can analyze Haar-distributed random matrices directly, which have led to strong quantitative results [79, 80, 18, 27]. The recent work of Austin [5] presents a new perspective on Theorem 2.3 through an associated large deviations theorem. Theorem 2.3 has also been extended to certain unitary Brownian motions, cf. [31, 10].

All the results discussed so far are concerned with models that are amenable to analytic methods, such as integration by parts and Poincaré inequalities. This stands in contrast to the following breakthrough result of Bordenave and Collins [17], which has a more combinatorial flavor.

**THEOREM 2.4** (Bordenave–Collins). *Let  $\Pi^N = (\Pi_1^N, \dots, \Pi_r^N)$  be i.i.d. uniformly distributed  $N \times N$  random permutation matrices, and let  $\mathbf{u} = (u_1, \dots, u_r)$  be free Haar unitaries. Denote by  $U_k^N = \Pi_k^N|_{1^\perp}$  the restriction of the permutation matrix  $\Pi_k^N$  to the orthogonal complement of the vector  $\mathbf{1}$  (the vector with unit entries, which is fixed by every permutation matrix). Then  $\mathbf{U}^N = (U_1^N, \dots, U_r^N)$  converges strongly to  $\mathbf{u}$ .*

To give a first hint of the strength of Theorem 2.4, note that

$$A^N = \Pi_1^N + \Pi_1^{N*} + \dots + \Pi_r^N + \Pi_r^{N*}$$

may be viewed as the adjacency matrix of a random  $2r$ -regular graph with  $N$  vertices. By the Perron-Frobenius theorem, every  $2r$ -regular graph has a trivial largest eigenvalue  $2r$  with eigenvector  $\mathbf{1}$ . Theorem 2.4 and (2.1) imply that the nontrivial eigenvalues of a random  $2r$ -regular graph satisfy

$$\max_{i=2, \dots, N} |\lambda_i(A^N)| = \|A^N|_{1^\perp}\| \xrightarrow{N \rightarrow \infty} 2\sqrt{2r-1}.$$

This is one of the deepest results in the spectral theory of random graphs, due to Friedman [38]. It is recovered here as one very special case of strong convergence of random permutation matrices. But Theorem 2.4 is a much stronger result that paves the way for new applications of strong convergence (cf. Section 4). New proofs of Theorem 2.4 that yield much stronger quantitative information were obtained in [18, 26].

We now describe a different perspective on Theorems 2.3 and 2.4 that has recently led to far-reaching generalizations of these results. Let  $\mathbf{S}_N$  be the symmetric group on  $N$  letters, and denote by

$$\text{std}_N : \mathbf{S}_N \rightarrow \text{M}_{N-1}(\mathbb{C})$$

the map that associates to each permutation  $\sigma \in \mathbf{S}_N$  the restriction of the corresponding  $N \times N$  permutation matrix to  $1^\perp$ . Then  $\text{std}_N$  is an irreducible representation of  $\mathbf{S}_N$ , called the *standard representation*. The random matrices that appear in Theorem 2.4 are therefore defined by  $U_k^N = \text{std}_N(\sigma_k)$ , where  $\sigma_1, \dots, \sigma_r$  are i.i.d. uniformly distributed random elements of  $\mathbf{S}_N$ . The random matrices in Theorem 2.3 may similarly be viewed as arising from the defining representation of the classical Lie groups  $\text{U}(N)/\text{O}(N)/\text{Sp}(N)$ .

One may now ask what happens if we consider other irreducible representations of these groups. In a series of recent papers [19, 26, 67, 27, 25], it has been shown that strong convergence remains valid for a remarkably large range of representations. For sake of illustration, we state one of the strongest results to date in this direction due to Cassidy [25] (see [92, Theorem 5.8] for this formulation).

**THEOREM 2.5** (Cassidy). *Let  $\sigma_1^N, \dots, \sigma_r^N$  be i.i.d. uniform random elements of  $\mathbf{S}_N$ , and  $\pi_N : \mathbf{S}_N \rightarrow \text{U}(D_N)$  be any irreducible unitary representation of  $\mathbf{S}_N$  of dimension  $1 < D_N \leq \exp(N^{1/21})$ . Define  $\mathbf{U}^N = (U_1^N, \dots, U_r^N)$  by  $U_k^N = \pi_N(\sigma_k^N)$ , and let  $\mathbf{u} = (u_1, \dots, u_r)$  be free Haar unitaries. Then  $\mathbf{U}^N$  converges strongly to  $\mathbf{u}$ .*

If  $\pi_N = \text{std}_N$  and  $D_N = N - 1$  this result recovers Theorem 2.4. However, the spirit of Theorem 2.5 is that it requires much less randomness: it can produce strongly convergent random matrices of dimension  $D$  using only  $(\log D)^{22}$  random bits, while Theorem 2.4 requires of order  $D \log D$  random bits to achieve the same conclusion. Analogous results for the unitary group may be found in [19, 67, 27].

Results such as Theorem 2.5 make one wonder how much randomness is really needed to achieve strong convergence. Could it be that Theorem 2.5 remains valid for *any* choice of representations  $\pi_N$  with  $D_N > 1$ ? Could one hope to achieve strong convergence in a situation where the group itself is fixed, such as  $\text{SU}(2)$ , and only the dimension of the representations  $\pi_N$  grows? Could one hope to achieve strong convergence with no randomness at all, using number-theoretic constructions such as those that have been used to obtain regular graphs with optimal spectral properties [64]? These tantalizing questions remain very much open.<sup>1</sup>

**2.3 Beyond freeness.** All results discussed so far are concerned with families of i.i.d. random matrices, whose limiting objects are free. Whether strong convergence can also hold outside the setting of free groups is however of major interest, particularly for applications to geometry where the relevant group is the fundamental group of the underlying manifold (see Section 4.2.1). The study of such questions was pioneered by Magee, see the survey [65]. Here we briefly describe some results in this direction.

Let  $\mathbf{G}$  be a finitely generated group with generators  $g_1, \dots, g_r$ . The question is whether there is a sequence

$$\rho_N : \mathbf{G} \rightarrow \text{U}(D_N)$$

of random unitary representations of  $\mathbf{G}$  that converge strongly to the regular representation  $\lambda_{\mathbf{G}}$ , in the sense that

$$\lim_{N \rightarrow \infty} \|\rho_N(x)\| = \|\lambda_{\mathbf{G}}(x)\| \quad \text{in probability}$$

for every  $x \in \mathbb{C}[\mathbf{G}]$ . This question can be rephrased as a special instance of Definition 1.1: we aim to find random unitary matrices  $\mathbf{U}^N = (U_1^N, \dots, U_r^N)$  that converge strongly to  $\mathbf{u} = (u_1, \dots, u_r)$  defined by  $u_k = \lambda_{\mathbf{G}}(g_k)$ , and such that any relation of  $\mathbf{u}$  is also satisfied by  $\mathbf{U}^N$ . If it is case that  $U_k^N = \Pi_k^N|_{1^\perp}$  for some random permutation matrices  $\Pi_k^N$ , then  $\rho_N$  are called random permutation representations.

Let us illustrate this question in two concrete examples. When  $\mathbf{G} = \mathbf{F}_r$  is a free group, since there are no relations, the existence of random unitary or permutation representations is simply a reformulation of Theorems 2.3 and 2.4, respectively. On the other hand, suppose that  $\mathbf{G} = \Gamma_2$  is

$$\Gamma_2 = \langle g_1, g_2, g_3, g_4 : [g_1, g_2][g_3, g_4] = \mathbf{1} \rangle,$$

which is the fundamental group of a surface of genus two (here  $[a, b] = aba^{-1}b^{-1}$  and  $\mathbf{1}$  is the identity). Then the question is to find random unitary matrices  $\mathbf{U}^N = (U_1^N, U_2^N, U_3^N, U_4^N)$  that converge strongly to

<sup>1</sup>These questions are folklore, see, e.g., [93, 19], and [24, 41, 87] for closely related questions.

$\mathbf{u} = (u_1, u_2, u_3, u_4)$  with  $u_k = \lambda_{\mathbf{r}_2}(g_k)$ , with the additional requirement that  $[U_1^N, U_2^N][U_3^N, U_4^N] = \mathbf{1}$  for every  $N$ . The latter constraint leads to complicated random matrix models.

In first instance, one may attempt to reduce this question to the results of the previous section by embedding the non-free group  $\mathbf{G}$  in a free group  $\mathbf{F}_r$ . This is not strictly possible, since every subgroup of a free group is free. However, there is a class of groups, called *limit groups*, that have the following property: for every  $N$ , one can associate to each generator  $g_i$  of  $\mathbf{G}$  an element  $h_i$  in  $\mathbf{F}_r$  such that  $g_1, \dots, g_r$  and  $h_1, \dots, h_r$  have the same relations of length up to  $N$ . For such groups, Louder and Magee [62] prove the following.

**THEOREM 2.6** (Louder–Magee). *Any limit group  $\mathbf{G}$  admits a sequence of random permutation representations that converge strongly to the regular representation  $\lambda_{\mathbf{G}}$ .*

The proof exploits the encoding of  $\mathbf{G}$  in  $\mathbf{F}_r$  to reduce the problem to an instance of Theorem 2.4. In this model, each  $U_k^N$  is a word (that depends on  $N$ ) of independent random permutation matrices.

An important example of limit groups are the surface groups  $\Gamma_g$  of genus  $g$ , and thus Theorem 2.6 provides a strongly convergent random matrix model for surface groups. However, the distribution of these random matrices is highly nonuniform. Motivated by geometric applications (see Section 4.2.1), one may ask whether a *typical* permutation representation of  $\Gamma_g$  converges strongly. For example, for genus two, this question asks whether sampling permutation matrices uniformly at random from the set

$$\{(\Pi_1^N, \Pi_2^N, \Pi_3^N, \Pi_4^N) : \Pi_1^N, \Pi_2^N, \Pi_3^N, \Pi_4^N \text{ are } N \times N \text{ permutation matrices such that } [\Pi_1^N, \Pi_2^N][\Pi_3^N, \Pi_4^N] = \mathbf{1}\}$$

and defining  $U_k^N = \Pi_k^N|_{1^\perp}$  yields a strongly convergent model for  $\Gamma_2$ . That this is indeed the case was proved by Magee, Puder, and the author [69] using new methods of random matrix theory (see Section 5).

**THEOREM 2.7** (Magee–Puder–van Handel). *For any  $g \geq 2$ , uniform random permutation representations of  $\Gamma_g$  converge strongly to the regular representation  $\lambda_{\Gamma_g}$ .*

We now turn to another class of non-free groups which may be viewed as a mixture of free and abelian groups. Let  $G = ([r], E)$  be a finite simple graph with  $r$  vertices. The right-angled Artin group  $\mathbf{A}_G$  has one generator for each vertex of  $G$ , where a pair of generators commutes if and only if there is an edge between them:

$$\mathbf{A}_G = \langle g_1, \dots, g_r : [g_i, g_j] = \mathbf{1} \text{ for every } \{i, j\} \in E \rangle.$$

The following was proved by Magee and Thomas [70].

**THEOREM 2.8** (Magee–Thomas). *Every right-angled Artin group  $\mathbf{A}_G$  admits a sequence of random unitary representations that converge strongly to the regular representation  $\lambda_{\mathbf{A}_G}$ .*

A natural candidate random matrix model for  $\mathbf{A}_G$  is obtained by choosing  $U_k^N$  to be independent Haar-distributed random unitary matrices of dimension  $N^2$  that act on pairs of factors of a tensor product  $(\mathbb{C}^N)^{\otimes K}$ , chosen so that  $U_i^N$  and  $U_j^N$  act on disjoint tensor factors if and only if  $\{i, j\} \in E$ . This model was conjectured to converge strongly in [70]. The model used in the proof of Theorem 2.8 is a more complicated variant of this construction; the above conjecture was subsequently resolved in [27, §9.4].

The importance of Theorem 2.8 is that many interesting groups virtually embed in a right-angled Artin group, so that Theorem 2.8 provides strongly convergent random unitary representations for any such group. This includes, notably, the fundamental group of any closed hyperbolic 3-manifold. It should be emphasized, however, that Theorem 2.8 provides only random *unitary* representations and not random *permutation* representations. In fact, there are right-angled Artin groups for which the latter *cannot* exist, see [65, Proposition 2.7]. The situation is even worse for some other groups: it was shown by Magee and de la Salle [66] that the group  $\mathrm{SL}_4(\mathbb{Z})$  does not even admit a strongly convergent sequence of unitary representations. Beyond the results discussed above, the question of which groups admit strongly convergent representations remains largely open.

**2.4 The main approaches to strong convergence.** We now aim to briefly survey, without details, the methods of random matrix theory that have been used to prove strong convergence in different models. Roughly speaking, this has been achieved using four distinct approaches.

1. The original approach of Haagerup and Thorbjørnsen [47] uses a variant of the Schwinger-Dyson equations of classical random matrix theory (see, e.g., [42]) to obtain approximate “master equations” for the expected resolvents of the random matrices in question.

2. The approach developed by Bordenave and Collins [17, 18] uses sophisticated forms of the moment method of classical random matrix theory, relying in particular on matrix-valued extensions of nonbacktracking methods that were previously used for the study of random graphs.
3. The interpolation method, variants of which were developed independently by Collins, Guionnet and Parraud [32] and by Bandeira, Boedihardjo and the author [7], is based on the idea of constructing a continuous interpolation  $(\mathbf{X}_t^N)_{t \in [0,1]}$  between the random matrices  $\mathbf{X}_1^N = \mathbf{X}^N$  and the limiting operators  $\mathbf{X}_0^N = \mathbf{x}$  that appear in Definition 1.1, and bounding the derivative of spectral statistics with respect to  $t$ .
4. The polynomial method, which was introduced by Chen, Garza-Vargas, Tropp, and the author [26] and refined in several further works, is based on the observation that the spectral statistics of many random matrix models of dimension  $N$  are regular functions of  $\frac{1}{N}$ . The method provides a way of interpolating between the random matrix and limiting models by “differentiating with respect to  $\frac{1}{N}$ ”.

A major difficulty in establishing strong convergence is that one must understand the behavior of arbitrary  $*$ -polynomials of the underlying matrices; these can have a very complicated structure, and their spectral statistics are not described by tractable equations. An influential idea that was introduced by Haagerup and Thorbjørnsen (based on earlier work of Pisier [84, 86] in operator space theory) is the *linearization trick*: to prove that

$$\lim_{N \rightarrow \infty} \|P(\mathbf{X}^N)\| = \|P(\mathbf{x})\|$$

for every  $*$ -polynomial  $P$ , it suffices to prove convergence of the spectrum of *linear* self-adjoint  $*$ -polynomials with *matrix* coefficients, that is, expressions of the form

$$Q(x_1, \dots, x_r) = A_0 \otimes \mathbf{1} + \sum_{k=1}^r (A_k \otimes x_k + A_k^* \otimes x_k^*)$$

where  $A_0, \dots, A_r$  are matrices of any fixed dimension  $D$  and  $A_0$  is self-adjoint. This reduction is crucial for obtaining tractable equations: for example, the matrix Stieltjes transform of  $Q(\mathbf{s})$ , where  $\mathbf{s}$  is a free semicircular family, satisfies an explicit quadratic equation called the matrix Dyson equation [47, Eq. (1.5)].

Approaches 1. and 2. to strong convergence described above rely strongly on linearization. However, the interpolation and polynomial methods 3. and 4. can be applied directly to arbitrary  $*$ -polynomials, since they work by interpolating between the spectral statistics of the matrix and limiting models rather than by analyzing equations satisfied by their spectral statistics. For this reason, the latter two methods also tend to be more robust and have been successfully applied to a broader range of models.

A different distinction between these methods is that approaches 1. and 3. rely strongly on analytic tools, such as integration by parts and Poincaré inequalities, which are not available for many discrete models. In contrast, methods 2. and 4. are ultimately based only on moment computations which are accessible for a broad class of random matrix models. The latter approaches have therefore proved to be essential for the study of questions such as strong convergence of random permutations. However, the study of certain highly irregular models, such as joint strong convergence of random and deterministic matrices [71, 32] or the intrinsic freeness phenomenon discussed in Section 3 below, has so far been accomplished only in analytic settings.

Among the methods described above, the recently introduced polynomial method has proved to be especially powerful both in the range of models whose analysis it enables and in the strength of the quantitative results that can be obtained from it. We will discuss this method further in Section 5.

**3 Intrinsic freeness.** The aim of this section is to describe a surprising cousin of the strong convergence phenomenon, the *intrinsic freeness principle*, developed by Bandeira, Boedihardjo, and the author [7]. While strong convergence in the sense of Definition 1.1 states that the spectrum of a sequence of random matrices behaves asymptotically as that of a limiting operator, the upshot of this section is that—in a certain sense—the spectrum of “almost any” gaussian random matrix behaves, nonasymptotically, like that of an associated deterministic operator. This opens the door to studying essentially arbitrarily structured random matrices of the kind that appear, for example, in many problems of applied mathematics.

To motivate this development, let us begin by revisiting the classical strong convergence theorem of Haagerup and Thorbjørnsen. Let  $G_1^N, \dots, G_r^N$  be independent GUE matrices of dimension  $N$ , and let  $s_1, \dots, s_r$  be a free

semicircular family. Define the  $DN$ -dimensional random matrix

$$X^N = A_0 \otimes \mathbf{1} + \sum_{i=1}^r A_i \otimes G_i^N$$

and the associated limiting operator

$$X_{\text{free}} = A_0 \otimes \mathbf{1} + \sum_{i=1}^r A_i \otimes s_i,$$

where  $A_0, \dots, A_r$  is an arbitrary family of nonrandom self-adjoint matrices of dimension  $D$  that are independent of  $N$ . The main result of the paper of Haagerup and Thorbjørnsen [47] is the following.

**THEOREM 3.1** (Haagerup–Thorbjørnsen). *For any  $X^N$  and  $X_{\text{free}}$  as above,*

$$\lim_{N \rightarrow \infty} d_H(\text{sp}(X^N), \text{sp}(X_{\text{free}})) = 0 \quad a.s.$$

Here  $\text{sp}(X)$  denotes the spectrum of  $X$  and  $d_H$  is the Hausdorff distance.

Even though it is formulated in a different manner, this statement is in fact equivalent to strong convergence of  $\mathbf{G}^N = (G_1^N, \dots, G_r^N)$  to  $\mathbf{s} = (s_1, \dots, s_r)$  by the linearization trick described in Section 2.4.

We now take a different perspective, however, and note that the above model is of significant interest in its own right *especially* in the case  $N = 1$ : in this case, the random matrix  $X = X^1$ , that is,

$$X = A_0 + \sum_{i=1}^r A_i g_i$$

where  $g_1, \dots, g_r$  are i.i.d. standard gaussian (scalar) variables, defines an *arbitrary*  $D$ -dimensional self-adjoint random matrix with jointly gaussian entries by a suitable choice of the matrix coefficients. If the spectrum of such matrices could be understood at this level of generality, one would have understood the behavior of completely arbitrary gaussian random matrices. But there is of course no reason to expect that strong convergence as  $N \rightarrow \infty$ , as in Theorem 3.1, sheds any light on the behavior of the model for  $N = 1$ .

The intrinsic freeness principle states that spectrum of  $X$  is nonetheless captured by that of  $X_{\text{free}}$  in surprising generality. This phenomenon is not quantified by the dimension, but rather by an “intrinsic” parameter

$$v(X) = \|\text{Cov}(X)\|^{1/2}$$

where  $\text{Cov}(X)$  denotes the  $D^2 \times D^2$  covariance matrix of the entries of  $X$ . Among various such results, we highlight the following theorem (stated here in slightly simplified form) that is a combination of a result of Bandeira, Boedihardjo, and the author [7] and of Bandeira, Cipolloni, Schröder, and the author [8].

**THEOREM 3.2** (Bandeira–Boedihardjo–Cipolloni–Schröder–van Handel). *For any  $X$  and  $X_{\text{free}}$  as above,*

$$\mathbf{P}[d_H(\text{sp}(X), \text{sp}(X_{\text{free}})) > Cv(X)^{1/2} \|X_{\text{free}}\|^{1/2} ((\log D)^{3/4} + t)] \leq e^{-t^2}$$

for all  $t \geq 0$ . Here  $C$  is a universal constant.

The utility of this result comes from two directions. On the one hand, the parameter  $v(X)$  turns out to small under surprisingly mild assumptions, even when the random matrix  $X$  is very sparse or has significant dependence between its entries. To give just one simple example, a random band matrix  $X$  has  $v(X) = o((\log D)^{3/2})$  as soon as its band width is polylogarithmic in the dimension  $D$ , which is nearly optimal up to the power on the logarithm.<sup>2</sup> Moreover, since Theorem 3.2 imposes no structural assumptions on the random matrix  $X$ , it is readily applicable to all kinds of messy random matrices that appear in applications.

On the other hand, the spectrum of  $X_{\text{free}}$  is amenable to analysis using tools of free probability. For example, the upper edge of the spectrum  $\lambda_{\max}(X_{\text{free}}) = \sup \text{sp}(X_{\text{free}})$  is given by a variational principle

$$(3.1) \quad \lambda_{\max}(X_{\text{free}}) = \inf_{M > 0} \lambda_{\max} \left( A_0 + M^{-1} + \sum_{i=1}^r A_i M A_i \right)$$

---

<sup>2</sup>It is easily seen in this case that  $d_H(\text{sp}(X), \text{sp}(X_{\text{free}})) \not\rightarrow 0$  when the band width is  $o(\log D)$ .

due to Lehner [61]. When combined with Theorem 3.2, this formula enables a precise analysis of various complex random matrix models; see, for example, [8] and Section 4.4.3 below.

Theorem 3.2 is one of several results that capture the intrinsic freeness phenomenon. While our focus here is on the spectrum itself, analogous results for the spectral distribution may be found in [7]. In another direction, Brailovskaya and the author [21] extend these results to a large class of non-gaussian random matrices. The papers [7, 8, 21, 9] further illustrate the utility of these results in a diverse range of applications.

The above developments were motivated by the work of Haagerup and Thorbjørnsen, as well as by a paper of Tropp [91] which suggested the idea of capturing free behavior in the context of matrix concentration inequalities and developed some initial tools for this purpose. The key new ingredients developed in [7] are the correct formulation of the intrinsic freeness principle and the associated interpolation method which is essential to the proof. The role of the parameter  $v(X)$ , and the reason that it quantifies the degree to which  $X$  behaves “freely”, is not obvious at first sight; a discussion of how it arises may be found in [92, §4].

## 4 Applications.

### 4.1 Random graphs.

**4.1.1 Random lifts of graphs.** Let  $G^N$  be a  $d$ -regular graph with  $N$  vertices and adjacency matrix  $A^N$ . Such a graph always has largest eigenvalue  $\lambda_1(A^N) = d$  with eigenvector  $\mathbf{1}$ , and the remaining eigenvalues are bounded by  $\|A^N|_{1^\perp}\|$ ; the latter quantity controls the rate at which a random walk on  $G^N$  mixes. The following classical result shows that random walks on  $d$ -regular graphs cannot mix arbitrarily quickly [55, §5.2].

LEMMA 4.1 (Alon–Boppana). *For any sequence of  $d$ -regular graphs  $G^N$  with  $N$  vertices,*

$$\|A^N|_{1^\perp}\| \geq 2\sqrt{d-1} - o(1) \quad \text{as } N \rightarrow \infty.$$

The existence of a universal lower bound raises the question whether there exist sequences of graphs that achieve this bound; random walks on such graphs mix at the fastest possible rate. That this is indeed the case was already discussed in Section 2.2: Friedman’s theorem, which may be viewed as a very special case of Theorem 2.4, states that the adjacency matrix  $A^N$  of a *random*  $d$ -regular graph satisfies

$$\|A^N|_{1^\perp}\| \leq 2\sqrt{d-1} + o(1) \quad \text{as } N \rightarrow \infty$$

in probability. However, strong convergence of random permutation matrices yields a far more general understanding of such questions, as we will presently explain.

The Alon–Boppana bound is one instance of a general phenomenon: many geometric objects admit a universal bound on their nontrivial eigenvalues in terms of the spectrum of their universal covering space. This explains the form of Lemma 4.1, since the universal covering space of any  $d$ -regular graph is the infinite  $d$ -regular tree which has spectral radius  $2\sqrt{d-1}$ . An analogous result for hyperbolic surfaces appears in Section 4.2.1 below. It is less obvious how to formulate such a result for non-regular graphs, however: the universal cover of a non-regular graph is still a tree, but different graphs give rise to different universal covers.

To construct a sequence  $G^N$  of non-regular graphs with the same universal cover, it is natural to fix a base graph  $G$  and choose  $G^N$  to be an  $N$ -fold cover of  $G$ . As every eigenfunction of  $G$  lifts to an eigenfunction of  $G^N$ , the nontrivial eigenvalues in this setting are the *new* eigenvalues of  $G^N$  relative to  $G$ . The analogue of Lemma 4.1 then states [37, §4] that  $\|A^N|_{\text{new}}\|$  is asymptotically lower bounded by the spectral radius  $\rho$  of the universal cover. It was conjectured by Friedman [37] that this lower bound is achieved by *random  $N$ -lifts*, that is, for  $G^N$  chosen uniformly at random from all  $N$ -fold covers of  $G$ . This was proved by Bordenave and Collins [17].

THEOREM 4.2 (Bordenave–Collins). *For any fixed base graph  $G$ , its random  $N$ -lift  $G^N$  satisfies*

$$\lim_{N \rightarrow \infty} \|A^N|_{\text{new}}\| = \rho \quad \text{in probability,}$$

where  $\rho$  denotes the spectral radius of the universal cover of  $G$ .

The proof of Theorem 4.2 is in fact an easy corollary of Theorem 2.4. The random graph  $G^N$  can be constructed explicitly by starting with  $N$  duplicates of the base graph  $G$ , and randomly permuting the endpoints of each duplicate edge among the duplicate vertices. The resulting adjacency matrix  $A^N$  can be expressed as a linear  $*$ -polynomial with matrix coefficients of independent random permutation matrices, and what remains is a straightforward application of strong convergence (cf. Section 2.4).



**4.1.2 Random Schreier graphs.** As was explained in Section 2.2, one can model a random  $2r$ -regular graph by choosing its adjacency matrix  $A^N$  to be the sum of  $r$  independent uniformly distributed  $N \times N$  permutation matrices and their adjoints. Combinatorially, this graph is defined by connecting each vertex  $x \in [N]$  to its  $2r$  neighbors  $\sigma_i(x)$  and  $\sigma_i^{-1}(x)$  for  $i = 1, \dots, r$ .

It is possible, however, to use a very similar construction to produce random  $2r$ -regular graphs that use much less randomness [39]. Denote by  $[N]_k$  the set of all  $k$ -tuples of distinct elements of  $[N]$ . We define the action  $\mathbf{S}_N \curvearrowright [N]_k$  by applying  $\sigma \in \mathbf{S}_N$  elementwise to each tuple  $(x_1, \dots, x_k) \in [N]_k$ , that is,

$$\sigma(x_1, \dots, x_k) = (\sigma(x_1), \dots, \sigma(x_k)).$$

We now define a random  $2r$ -regular graph whose vertex set is  $[N]_k$ , and where each vertex  $\tilde{x} \in [N]_k$  is connected to its  $2r$  neighbors  $\sigma_i(\tilde{x})$  and  $\sigma_i^{-1}(\tilde{x})$  for  $i = 1, \dots, r$ . When  $k = 1$ , this is the classical model discussed above. As  $k$  is increased, the same set of random permutations is used to construct  $2r$ -regular graphs with an increasingly large number of vertices. Do such graphs still have an optimal spectral gap?

Theorem 2.5 provides a striking answer to this question: such graphs do indeed have an optimal spectral gap even when  $k$  is allowed to grow polynomially with  $N$ . In this case, the number of random bits needed to construct the graphs is only polylogarithmic in the number of vertices, in contrast to the classical model of random regular graphs which requires a superlinear number of random bits.

**THEOREM 4.3 (Cassidy).** *Let  $A^{N,k}$  be the adjacency matrix of the random  $2r$ -regular graph with vertex set  $[N]_k$  and edges defined by the action  $\mathbf{S}_N \curvearrowright [N]_k$  of  $r$  independent uniform random permutations. Then*

$$\lim_{N \rightarrow \infty} \|A^{N,k_N}|_{1^\perp}\| = 2\sqrt{2r-1} \quad \text{in probability}$$

as long as  $k_N \leq N^{1/21}$ .

The point here is that the map  $\pi_{N,k} : \mathbf{S}_N \rightarrow \mathbf{S}_{(N)_k}$  that associates to each permutation of  $[N]$  the corresponding permutation of  $[N]_k$  that is defined by the action  $\mathbf{S}_N \curvearrowright [N]_k$  is a representation of  $\mathbf{S}_N$ . Since

$$A^{N,k} = \pi_{N,k}(\sigma_1) + \pi_{N,k}(\sigma_1)^* + \dots + \pi_{N,k}(\sigma_r) + \pi_{N,k}(\sigma_r)^*,$$

the conclusion of Theorem 4.3 follows readily from Theorem 2.5.

The model of random graphs discussed above may be viewed as a Schreier graph  $\text{Sch}(\mathbf{S}_N \curvearrowright [N]_k; \sigma_1, \dots, \sigma_r)$  of the symmetric group. When  $k = 1$ , it recovers the classical permutation model of random regular graphs. When  $k = N$ , it coincides with the random Cayley graph  $\text{Cay}(\mathbf{S}_N; \sigma_1, \dots, \sigma_r)$ . Whether random Cayley graphs of  $\mathbf{S}_N$  have a nonvanishing spectral gap at all—let alone an optimal one—is a long-standing open question. Theorem 2.5 settles a situation that is intermediate between these two extremes.

It could be argued that Theorem 4.3 does not really rely on strong convergence, since it is concerned only with one very special  $*$ -polynomial  $P(x_1, \dots, x_r) = x_1 + x_1^* + \dots + x_r + x_r^*$ . However, the connection with strong convergence is twofold. First, methods that were developed to establish strong convergence play a key role in the proof of Theorem 2.5. Second, the fact that Theorem 2.5 provides a full strong convergence statement yields direct analogues of Theorem 4.3 in many other situations: for example, an analogous modification of Theorem 4.2 yields a model of random  $N$ -lifts that uses only a polylogarithmic number of random bits.

## 4.2 Geometry.

**4.2.1 Hyperbolic surfaces.** As was discussed in Section 4.1.1, the phenomenon described by the Alon–Boppana bound appears in many other situations. A completely analogous result for hyperbolic surfaces was observed long ago by Huber [56] and (in more general form) by Cheng [29]. In the following, we denote by  $\Delta_X$  the Laplacian on  $X$ , and by  $0 = \lambda_0(X) < \lambda_1(X) \leq \lambda_2(X) \leq \dots$  the eigenvalues of  $\Delta_X$ .

**LEMMA 4.4 (Huber; Cheng).** *For any sequence of closed hyperbolic surfaces  $X^N$  with diverging diameter,*

$$\lambda_1(X^N) \leq \frac{1}{4} + o(1) \quad \text{as } N \rightarrow \infty.$$

This bound arises because the universal covering space of every hyperbolic surface is the hyperbolic plane  $\mathbb{H}^2$ , which has  $\lambda_1(\mathbb{H}^2) = \frac{1}{4}$ . As in the case of graphs, this universal upper bound raises the question whether there exist sequences of closed hyperbolic surfaces that achieve this bound. This long-standing conjecture was resolved in the affirmative in a breakthrough paper of Hide and Magee [53].

THEOREM 4.5 (Hide–Magee). *There exist closed hyperbolic surfaces  $X^N$  with diverging diameter such that*

$$\lambda_1(X^N) \geq \frac{1}{4} - o(1) \quad \text{as } N \rightarrow \infty.$$

Hide and Magee construct their surfaces analogously to the construction of random  $N$ -lifts of graphs: they fix a base surface  $X$ , and choose each  $X^N$  to be an  $N$ -fold cover of  $X$ . More explicitly, let  $X = \Gamma \backslash \mathbb{H}^2$  for a Fuchsian group  $\Gamma \simeq \pi_1(X)$  acting on  $\mathbb{H}^2$ . The fundamental domain  $F$  of this action is a hyperbolic polygon in  $\mathbb{H}^2$  whose edges are of the form  $F \cap g_k F$  or  $F \cap g_k^{-1} F$ , where  $g_1, \dots, g_r$  are generators of  $\Gamma$ ; one recovers  $X$  by gluing each pair of edges that are defined by the same generator. To construct an  $N$ -fold cover of  $X$ , we start with  $N$  duplicates of  $F$  and permute the edges that we glue together among the duplicate polygons. If these permutations are chosen randomly, we obtain a random  $N$ -fold cover of  $X$ .

There are two significant obstacles in the analysis of such models. First, in contrast to the case of random  $N$ -lifts, it is not obvious how to relate the spectral properties of the Laplacian  $\Delta_{X^N}$  to those of the permutation matrices that define the  $N$ -fold cover  $X^N$ . A key insight of Hide and Magee is that the resolvent of  $\Delta_{X^N}$  can be approximated by a (nonlinear)  $*$ -polynomial with matrix coefficients of the underlying permutation matrices, which yields a nontrivial reduction from the spectral properties of the Laplacian to a strong convergence problem. Several variants of this reduction are developed in [54, 51, 65].

Second, in contrast to the case of graphs, not every choice of permutation matrices defines a valid cover: if one glues the edges of the fundamental polygons without accounting for their corners, one may no longer obtain a closed surface. To obtain a valid cover, what is needed is precisely that the permutations are chosen to satisfy the same relations as the corresponding generators of  $\Gamma$  [48, pp. 68–70]. In the original paper of Hide and Magee [53], this issue was circumvented by working instead with a *noncompact* base surface  $X$  for which  $\Gamma \simeq \mathbf{F}_r$  is free, so that Theorem 2.4 could be applied; closed surfaces are then obtained a posteriori by a compactification procedure. Covers of a closed surface  $X$  were subsequently constructed in [62] using Theorem 2.6.

Even though the proof of Theorem 4.5 is based on a random construction, this random model is highly nonuniform. The construction therefore does not shed much light on what a *typical*  $N$ -fold cover of  $X$  looks like. This question was resolved by Magee, Puder, and the author [69] using Theorem 2.7; the following may be viewed as the exact analogue of Theorem 4.2 in the setting of hyperbolic surfaces.

THEOREM 4.6 (Magee–Puder–van Handel). *For any closed orientable hyperbolic surface  $X$ , a fraction  $1 - o(1)$  of all  $N$ -fold covers  $X^N$  has the property that all their new eigenvalues are greater than  $\frac{1}{4} - o(1)$  as  $N \rightarrow \infty$ .*

We mention in this context a closely related question: does a typical hyperbolic surface of genus  $g$  satisfy the conclusion of Theorem 4.5 as  $g \rightarrow \infty$ ? Theorem 4.6 does not answer this question, as most surfaces of genus  $g$  do not cover a surface of smaller genus. The natural notion of “typical” in this setting is with respect to the Weil-Petersson measure on the moduli space of surfaces of genus  $g$ , whose study was pioneered by Mirzakhani [72]. In an impressive tour-de-force, Anantharaman and Monk [2] provided an affirmative answer to this question using methods inspired by the original proof of Friedman’s theorem. While it does not appear that this question can be reduced to a strong convergence problem, Hide, Macera, and Thomas [52] gave a new proof of this result by directly applying the polynomial method (Section 5) to this problem. This approach notably provides a polynomial convergence rate, which was expected in view of deep conjectures on quantum chaos. A polynomial rate in Theorem 4.6 was achieved by the same authors in [51].

In contrast to the Weil-Petersson model, the random cover model remains meaningful beyond the setting of hyperbolic surfaces. For example, Hide–Moy–Naud [54, 76] establish results analogous to Theorems 4.5 and 4.6 for surfaces with variable negative curvature. Analogous questions for hyperbolic manifolds in higher dimension remain open. Both the universal upper bound as in Lemma 4.4 and the methods of Hide–Magee extend to this setting (see, e.g., [29, 6]); what is missing is that, at present, it is not known whether the fundamental group of any such manifold admits a strongly convergent sequence of permutation representations.

**4.2.2 Minimal surfaces.** We now discuss a very different application of strong convergence to the theory of minimal surfaces. Recall that a surface  $Y$  in a Riemannian manifold  $M$  is called a minimal surface if it is a critical point of the area functional under compact perturbations. A basic question in this context is how the geometry of  $M$  constrains the minimal surfaces that sit inside it. For example, it was shown by Bryant [23] that the Euclidean unit sphere  $S^N$ , which has constant positive curvature, cannot contain a minimal surface of

constant negative curvature. The following surprising result of Song [90] presents a very different picture than what the result of Bryant might lead one to expect.

**THEOREM 4.7 (Song).** *There exists a sequence of closed minimal surfaces  $Y^N$  in Euclidean unit spheres  $S^{D_N}$  such that the Gaussian curvature  $K^N$  of  $Y^N$  satisfies*

$$\lim_{N \rightarrow \infty} \frac{1}{\text{area}(Y^N)} \int_{Y^N} |K^N + 8| = 0.$$

In other words, high-dimensional spheres contain minimal surfaces that have nearly constant curvature  $-8$ . These unusual surfaces are obtained by a random construction that we briefly sketch.

The approach is based on a classical variational method that constructs minimal surfaces by minimizing the Dirichlet energy [74, Chapter 4]. By imposing a symmetry constraint in the variational problem, one can construct closed minimal surfaces  $Y^N$  in  $S^{2N-1}$  that are  $\rho_N$ -equivariant, where  $\rho_N : \mathbf{F}_2 \rightarrow \mathbf{U}(N)$  is a unitary representation with finite range and where we identify  $S^{2N-1}$  with the unit sphere in  $\mathbb{C}^N$ . By a compactness argument, a subsequence of these surfaces converges to a minimal surface  $Y^\infty$  in the unit sphere  $S^\infty$  of an infinite-dimensional Hilbert space  $H$  that is equivariant with respect to some unitary representation  $\rho_\infty : \mathbf{F}_2 \rightarrow \mathbf{U}(H)$ .

In the absence of further assumptions, it is not clear what this limiting surface might look like. However, [90] proves a remarkable rigidity property: any  $\rho_\infty$ -equivariant minimal surface in  $S^\infty$  such that  $\rho_\infty$  is weakly equivalent to the regular representation  $\lambda$  of  $\mathbf{F}_2$  in the sense that (cf. [11, Appendix F])

$$\|\rho_\infty(x)\| = \|\lambda(x)\| \quad \text{for all } x \in \mathbb{C}[\mathbf{F}_2],$$

has constant curvature  $-8$  (note that such surfaces cannot exist in finite dimension due to the result of Bryant). Thus to complete the proof, it suffices to choose the finite dimensional representations  $\rho_N$  such that

$$\lim_{N \rightarrow \infty} \|\rho_N(x)\| = \|\lambda(x)\|,$$

which is nothing other than strong convergence (see Section 2.3). Theorem 4.7 follows by choosing  $\rho_N$  to be the random permutation representations of  $\mathbf{F}_2$  that converge strongly by Theorem 2.4.

### 4.3 Operator algebras.

**4.3.1  $\text{Ext}(C_{\text{red}}^*(\mathbf{F}_2))$  is not a group.** For our purposes, a  $C^*$ -algebra may be defined as a  $*$ -algebra of bounded operators on a Hilbert space that is closed under the operator norm. For example, for any finitely generated group  $\mathbf{G}$  with generators  $g_1, \dots, g_r$  and regular representation  $\lambda$ , the norm-closure of all  $*$ -polynomials in  $\lambda(g_1), \dots, \lambda(g_r)$  defines a  $C^*$ -algebra  $C_{\text{red}}^*(\mathbf{G})$ , called the reduced  $C^*$ -algebra of  $\mathbf{G}$ .

That a family of bounded operators  $x_1, \dots, x_r$  admits a strongly convergent (random) matrix model as in Definition 1.1 implies, in a particular sense, that the  $C^*$ -algebra generated by  $x_1, \dots, x_r$  admits a sequence of finite-dimensional approximations. This places strong constraints on the structure of such a  $C^*$ -algebra; for example, it implies that it is an MF-algebra in the sense of Blackadar and Kirchberg [16]. The initial development of the strong convergence phenomenon was strongly motivated by open problems in the theory of  $C^*$ -algebras. We briefly sketch one important problem of this kind, whose resolution was a major aim of the original work on strong convergence due to Haagerup and Thorbjørnsen [47].

To motivate this problem, recall that the spectrum of a bounded self-adjoint operator  $X$  on an infinite-dimensional separable Hilbert space  $H$  can be decomposed into the discrete spectrum and the essential spectrum. The Weyl-von Neumann theorem characterizes the essential spectrum as the part of the spectrum that is invariant under compact perturbations of  $X$ ; in other words, it is the spectrum of the image of  $X \in B(H)$  in the Calkin algebra  $B(H)/K(H)$ , where  $K(H)$  denotes the ideal of compact operators in  $B(H)$ .

Motivated by analogous questions for non-self-adjoint operators, Brown, Douglas, and Fillmore [22] proposed to investigate properties of  $C^*$ -algebras  $\mathcal{A}$  up to compact perturbations. A central role in this program is played by  $\text{Ext}(\mathcal{A})$ , which is defined as the set of  $*$ -homomorphisms  $\pi : \mathcal{A} \rightarrow B(H)/K(H)$  modulo unitary conjugation. The invariant  $\text{Ext}(\mathcal{A})$  may naturally be viewed as a semigroup with respect to the addition  $(\pi_1, \pi_2) \mapsto \pi_1 \oplus \pi_2$ . Rather surprisingly, there are many  $C^*$ -algebras in which every element of  $\text{Ext}(\mathcal{A})$  has an inverse, so that it is in fact a group. This suggests that perhaps  $\text{Ext}(\mathcal{A})$  might *always* be a group, but this is not the case: a counterexample was provided by Anderson [4]. It remained unclear, however, how to understand such questions in specific situations. In particular, the following result [47] had long remained open.

THEOREM 4.8 (Haagerup–Thorbjørnsen).  $\text{Ext}(C_{\text{red}}^*(\mathbf{F}_2))$  is not a group.

The key observation behind Theorem 4.8, due to Voiculescu [93, §5.14] (see also [47, Remark 8.6]), is that the existence of a strongly convergent sequence of finite-dimensional unitary representations  $\rho_N$  of  $\mathbf{F}_2$  presents an obstruction to  $\text{Ext}(C_{\text{red}}^*(\mathbf{F}_2))$  being a group, since it can be shown that the  $*$ -homomorphism defined by  $\bigoplus_{N=1}^{\infty} \rho_N$  is not invertible. This led Voiculescu to ask whether such a strongly convergent sequence of representations exists. This was established for the first time by Haagerup and Thorbjørnsen, settling the problem.

Strong convergence has subsequently been applied to various other problems in the theory of  $C^*$ -algebras. The original paper of Haagerup and Thorbjørnsen [47] also settles a question that arises from the work of Junge and Pisier on tensor products of  $B(H)$ . Haagerup, Schultz, and Thorbjørnsen [45] used strong convergence to give a random matrix theory proof of the fact that  $C_{\text{red}}^*(\mathbf{F}_2)$  has no nontrivial projections, which had previously been established using K-theory. Other applications include work of Voiculescu on topological free entropy [94], as well as various applications of the fact that  $C_{\text{red}}^*(\mathbf{F}_2)$  is an MF-algebra.

**4.3.2 The Peterson-Thom conjecture.** A *von Neumann algebra* is a  $*$ -algebra of bounded operators on a Hilbert space that is closed in the strong operator topology. For example, for any finitely generated group  $\mathbf{G}$  with generators  $g_1, \dots, g_r$  and regular representation  $\lambda$ , the closure of all  $*$ -polynomials in  $\lambda(g_1), \dots, \lambda(g_r)$  with respect to the strong operator topology defines a von Neumann algebra  $L(\mathbf{G})$ , called the group von Neumann algebra of  $\mathbf{G}$ . Since the strong operator topology is much weaker than the norm topology, the von Neumann algebra  $L(\mathbf{G})$  is much bigger than its  $C^*$ -counterpart  $C_{\text{red}}^*(\mathbf{G})$  and is in some ways more poorly understood. For example, it is not even known whether or not  $L(\mathbf{F}_r)$  and  $L(\mathbf{F}_s)$  are isomorphic for  $r \neq s$ .

One way to gain insight into an operator algebra is to investigate the structure of the subalgebras that sit inside it. The following theorem of Hayes [49] in this spirit settled a long-standing conjecture about amenable von Neumann subalgebras of  $L(\mathbf{F}_r)$  due to Peterson and Thom [83].

THEOREM 4.9 (Hayes). *Any diffuse amenable von Neumann subalgebra of  $L(\mathbf{F}_r)$  is contained in a unique maximal amenable von Neumann subalgebra of  $L(\mathbf{F}_r)$ .*

Hayes actually proves a stronger result that provides an entropic characterization of amenable subalgebras of  $L(\mathbf{F}_r)$ , of which Theorem 4.9 is a corollary. The methods of [49] have subsequently led to various related developments in the theory of von Neumann algebras [50]. We omit further discussion of the precise meaning and significance of the statement of Theorem 4.9, which is outside the scope of this survey.

The central insight of Hayes was that the Peterson-Thom conjecture can be reduced to a certain question of strong convergence of tensor products of GUE matrices: the main result of [49] states that the conclusion of Theorem 4.9 would follow if it can be shown that the family of  $N^2$ -dimensional random matrices

$$G_1^N \otimes \mathbf{1}, \dots, G_r^N \otimes \mathbf{1}, \mathbf{1} \otimes H_1^N, \dots, \mathbf{1} \otimes H_r^N$$

converges strongly to

$$s_1 \otimes \mathbf{1}, \dots, s_r \otimes \mathbf{1}, \mathbf{1} \otimes s_1, \dots, \mathbf{1} \otimes s_r,$$

where  $G_1^N, \dots, G_r^N, H_1^N, \dots, H_r^N$  are i.i.d. GUE matrices,  $s_1, \dots, s_r$  is a free semicircular family, and all tensor products are minimal. This strong convergence problem was beyond the reach of the methods that were available at the time that [49] was written, and Theorem 4.9 therefore appears in [49] as a conditional statement. Hayes' work drew much attention on the random matrix side, and the strong convergence result that is needed as input to [49] was subsequently proved by several different approaches [12, 18, 67, 82, 27].

The above strong convergence problem is closely connected with another question that arises from Pisier's work on subexponential operator spaces [85]. While Definition 1.1 defines strong convergence by requiring that  $\|P(\mathbf{X}^N)\| \rightarrow \|P(\mathbf{x})\|$  for all  $*$ -polynomials  $P \in \mathbb{C}^*\langle x_1, \dots, x_r \rangle$  with scalar coefficients, it is an elementary fact that this implies the same property also for  $*$ -polynomials  $P \in M_D(\mathbb{C}) \otimes \mathbb{C}^*\langle x_1, \dots, x_r \rangle$  with matrix coefficients; see, e.g., [92, Lemma 2.16]. Pisier asked whether it is still the case that

$$\|P_N(\mathbf{X}^N)\| = (1 + o(1))\|P_N(\mathbf{x})\|$$

when  $P_N \in M_{D_N}(\mathbb{C}) \otimes \mathbb{C}^*\langle x_1, \dots, x_r \rangle$  are  $*$ -polynomials with matrix coefficients of growing dimension, and if so how rapidly  $D_N$  can grow with  $N$ . The connection between this question and Hayes' strong convergence problem is that as  $\mathbf{G}^N = (G_1^N, \dots, G_r^N)$  and  $\mathbf{H}^N = (H_1^N, \dots, H_r^N)$  are independent, we may interpret

$$P_N(\mathbf{G}^N) = P(\mathbf{G}^N \otimes \mathbf{1}, \mathbf{1} \otimes \mathbf{H}^N)$$

as a  $*$ -polynomial of  $\mathbf{G}^N$  with matrix coefficients of dimension  $D_N = N$  by conditioning on  $\mathbf{H}^N$ . This observation suffices (by using an additional property of the  $C^*$ -algebra generated by  $s_1, \dots, s_r$ , viz. exactness) to show that Hayes' question has an affirmative answer if Pisier's question has an affirmative answer with  $D_N = N$ . However, it was noted by Pisier [85, (0.9)] that the methods of Haagerup and Thorbjørnsen can establish such a property only for  $D_N = o(N^{1/4})$ , which does not suffice for the purpose of Theorem 4.9.

It is now known that Pisier's question has an affirmative answer for much higher-dimensional matrix coefficients [18, 67, 82, 27], which amply suffices for proving strong convergence of Hayes' model. The best result to date, due to Chen, Garza-Vargas, and the author [27], shows that strong convergence remains valid whenever  $D_N = e^{o(N)}$ . On the other hand, strong convergence is known to fail when  $D_N \geq e^{CN^2}$ . What happens in between these two regimes remains a tantalizing question.

The above discussion illustrates that it is sometimes possible to build more complicated strongly convergent models (such as Hayes' model) out of simpler building blocks (strong convergence with matrix coefficients). More sophisticated constructions in this spirit were used in [70] and in [27, §9.4] to obtain strongly convergent models where GUE matrices act on overlapping factors of a tensor product. We finally note that strong convergence of tensor products also arises naturally in other operator algebraic problems, see, e.g., [78, Theorem 4.1].

#### 4.4 Further applications.

**4.4.1 Cutoff of random walks.** As was noted in Section 4.1.1, the spectral gap of a graph determines the rate at which a simple random walk on that graph converges to equilibrium. However, some random walks are known to exhibit a more subtle and striking phenomenon: their total variation distance to equilibrium drops abruptly from nearly one to nearly zero on a time scale much shorter than the mixing time itself. This property, known as the cutoff phenomenon, has attracted much attention in probability theory. A general understanding of which random walks exhibit a cutoff remains far from complete [88].

It was shown by Lubetzky and Peres [63] that simple random walks on a sequence of  $d$ -regular graphs that have an asymptotically optimal spectral gap (in the sense that their nontrivial eigenvalues are asymptotically bounded by  $2\sqrt{d-1}$ , cf. Lemma 4.1) always exhibit a cutoff. It is natural to ask whether a similar phenomenon holds for non-regular graphs or non-simple random walks, but it is not even entirely clear how to formulate such a result precisely. In [20], Bordenave and Lacoïn provide an affirmative answer to this question for sequences of graphs that are defined by strongly convergent permutation representations of a discrete group. For example, their results show that random walks on random lifts of any base graph (cf. Theorem 4.2) exhibit a cutoff.

**4.4.2 Quantum information theory.** Random matrices play an important role in quantum information theory. We briefly list a few applications of strong convergence to this area, without providing any details. Strong convergence has been used in various studies of additivity violation of the minimum output entropy of quantum channels [13, 15, 30, 40]. The intrinsic freeness theory described in Section 3 has been used to construct a large class of quantum expanders [60, 59] and to provide lower bounds for quantum tomography [1]. We also note that models of random matrices that act on overlapping factors of a tensor product, which were previously discussed in Section 2.3 in the context of Theorem 2.8 and in Section 4.3.2, appeared independently in the physics literature as generic models of quantum spin systems that interact through an arbitrary dependency graph [75, 35].

**4.4.3 Applied mathematics.** The intrinsic freeness theory of Section 3 is especially useful in applied mathematics, since random matrices with a messy structure arise routinely in applications. The papers [7, 8, 21, 9] discuss a diverse range of applications, and many more have subsequently appeared in the literature. For sake of illustration, let us briefly sketch one example of such an application.

Fix  $\lambda > 0$ , a vector  $x \in \{-1, 1\}^n$ , and i.i.d. standard gaussian variables  $(Z_S)_{S \subseteq [n]: |S|=p}$ . We view the latter variables as the entries of a symmetric tensor  $Z$  of order  $p$  that is defined by  $Z_{i_1, \dots, i_p} = Z_{\{i_1, \dots, i_p\}}$ . We now consider a signal plus noise model of the form  $Y = \lambda x^{\otimes p} + Z$ , whose entries are defined by

$$Y_S = \lambda x_S + Z_S$$

where  $x_S = \prod_{i \in S} x_i$ . The tensor PCA problem asks under what condition on the signal strength  $\lambda$  it is (algorithmically) possible to detect the presence of the signal in the noisy observation  $Y$  [73, 95].

The following spectral method, which was motivated by ideas of statistical physics, was proposed in [95]. Let

$p \geq 4$  be even and  $\ell \in [p/2, n - p/2]$ . We define the  $\binom{n}{\ell} \times \binom{n}{\ell}$  Kikuchi matrix  $X = (X_{S,T})_{S,T \subseteq [n]: |S|=|T|=\ell}$  as

$$X_{S,T} = \begin{cases} Y_{S \Delta T} & \text{when } |S \Delta T| = p, \\ 0 & \text{otherwise,} \end{cases}$$

where  $\Delta$  denotes the symmetric difference. The signal is detected by the presence of an outlier eigenvalue in the spectrum of  $X$ . The question is how large  $\lambda$  must be for this outlier to appear.

In [95], matrix concentration inequalities were used to obtain the correct order of magnitude of  $\lambda$  up to logarithmic factors. In contrast, Theorem 3.2 makes it possible to locate the *exact* threshold at which the outlier appears in a nontrivial range of the design parameter  $\ell$ , cf. [8, Theorem 3.7].

**THEOREM 4.10** (Bandeira–Cipolloni–Schröder–van Handel). *Let  $p/2 \leq \ell < 3p/4$  and  $d = \binom{\ell}{p/2} \binom{n-\ell}{p/2}$ . Then*

$$\lambda_{\max}(d^{-1/2}X) = \begin{cases} 2 + o(1) & \text{if } \lambda \leq d^{-1/2}, \\ \lambda d^{1/2} + \frac{1}{\lambda d^{1/2}} + o(1) & \text{if } \lambda > d^{-1/2} \end{cases}$$

with probability  $1 - o(1)$  as  $n \rightarrow \infty$ .

This example illustrates a typical situation where intrinsic freeness is useful. The random matrix  $X$  has a complicated structure with a combinatorial entry pattern and many dependent entries. Nonetheless, Theorem 3.2 readily reduces the problem of understanding  $\lambda_{\max}(X)$  to an instance of Lehner’s formula (3.1). The deterministic problem of analyzing the resulting variational principle proves to be relatively straightforward. While the latter still requires some work, the random matrix aspect of the analysis is completely subsumed by Theorem 3.2.

**5 The polynomial method.** A new approach to strong convergence, the polynomial method, was recently introduced in the work of Chen, Garza-Vargas, Tropp, and the author [26] and further developed in [27, 67, 69]. In contrast to previous approaches, this method is largely based on soft arguments that require limited problem-specific input. This has led to a series of new developments and applications that appear to be difficult to approach by other methods. Surprisingly, the polynomial method also yields the strongest known quantitative results in several problems that were previously approached by other means.

The polynomial method is not specific to strong convergence problems, but is rather a general method for capturing cancellations in spectral problems that possess some regular structure. In this section, we aim to sketch in a general setting how this method works and in what situations it is applicable.

In the following, we fix a sequence of self-adjoint random matrices  $Z^N$  of dimension  $N$  and a limiting operator  $Z^\infty$  in a  $C^*$ -algebra  $\mathcal{A}$ . For example, to apply the method to strong convergence as in Definition 1.1 we will choose  $Z^N = P(\mathbf{X}^N)$  and  $Z^\infty = P(\mathbf{x})$ . In the present setting, we consider the following two properties.

- Strong convergence in the sense that  $\|Z^N\| \xrightarrow{N \rightarrow \infty} \|Z^\infty\|$  in probability.
- Weak convergence in the sense that  $\mathbf{E}[\text{tr } h(Z^N)] \xrightarrow{N \rightarrow \infty} \tau(h(Z^\infty))$  for every polynomial  $h \in \mathbb{R}[z]$ .

Here  $\text{tr } M = \frac{1}{n} \text{Tr } M$  is the normalized trace of an  $N \times N$  matrix  $M$ , and  $\tau$  is a faithful normalized trace on  $\mathcal{A}$  (equivalently, we could write  $\tau(h(Z^\infty)) = \int h d\nu_0$  where  $\nu_0$  is spectral distribution of  $Z^\infty$ ).

The basic obstacle that is shared by essentially all methods for proving strong convergence is that the norm  $\|Z^N\|$ , a complicated function of the entries of  $Z^N$ , is typically not directly amenable to computations. In contrast, when  $h \in \mathbb{R}[z]$  is a polynomial, the spectral statistics  $\mathbf{E}[\text{tr } h(Z^N)]$  are expected polynomials of the entries of  $Z^N$  which often admit (albeit complicated) explicit formulas that provide a basis for their analysis. For this reason, establishing weak convergence is generally much easier than establishing strong convergence.

**5.1 Weak and strong asymptotic expansions.** A phenomenon that is observed in many random matrix models is that the spectral statistics  $\mathbf{E}[\text{tr } h(Z^N)]$  for polynomial  $h$  behave in a regular way as a function of  $N$ :

$$(5.1) \quad \mathbf{E}[\text{tr } h(Z^N)] = \nu_0(h) + \frac{\nu_1(h)}{N} + \frac{\nu_2(h)}{N^2} + \cdots + \frac{\nu_q(h)}{N^q} + O\left(\frac{1}{N^{q+1}}\right),$$

where  $\nu_k : \mathbb{R}[z] \rightarrow \mathbb{R}$  are linear functionals on the space of polynomials. This *weak asymptotic expansion* may be viewed as an extension of the notion of weak convergence to higher order, since by construction  $\nu_0(h) = \tau(h(Z^\infty))$ .

In several situations, the polynomial spectral statistics are even rational functions of  $\frac{1}{N}$  [34, 36], while in more complicated models the existence of such an expansion is a nontrivial fact [68, 2, 52].

We now formulate a basic question that has been considered, e.g., in [47, 89, 46, 81, 80]:

*Does (5.1) remain valid for smooth test functions  $h \in C^\infty(\mathbb{R})$ ?*

In this case, the linear functionals  $\nu_k$  must extend to Schwartz distributions  $\nu_k : C^\infty(\mathbb{R}) \rightarrow \mathbb{R}$ . Random matrix models for which this holds will be said to admit a *strong asymptotic expansion*. The significance of this question is that a strong asymptotic expansion provides a simple criterion for strong convergence, which forms the basis for the work of Haagerup–Thorbjørnsen [47] and Schultz [89]. We only state the upper bound, since the lower bound is typically an easy consequence of weak convergence (see, e.g., [26, Appendix A]).

LEMMA 5.1. *If  $Z^N$  admits a strong asymptotic expansion and  $\text{supp } \nu_1 \subseteq [-\|Z^\infty\|, \|Z^\infty\|]$ , then*

$$\|Z^N\| \leq (1 + o(1))\|Z^\infty\| \quad \text{in probability.}$$

*Proof.* Choose  $h \in C^\infty(\mathbb{R})$ ,  $h \geq 0$  with  $h(z) = 0$  for  $|z| \leq \|Z^\infty\|$  and  $h(z) = 1$  for  $|z| \geq \|Z^\infty\| + \varepsilon$ . Then

$$\mathbf{P}[\|Z^N\| \geq \|Z^\infty\| + \varepsilon] \leq \mathbf{E}[\#\{\text{eigenvalues } \lambda \text{ of } Z^N \text{ with } |\lambda| \geq \|Z^\infty\| + \varepsilon\}] \leq \mathbf{E}[\text{Tr } h(Z^N)],$$

where we note the unnormalized trace  $\text{Tr} = N \text{tr}$  on the right-hand side. But as  $h$  vanishes on  $[-\|Z^\infty\|, \|Z^\infty\|]$ , we have  $\nu_0(h) = \tau(h(Z^\infty)) = 0$  and  $\nu_1(h) = 0$ . Thus (5.1) yields  $\mathbf{E}[\text{Tr } h(Z^N)] = O(\frac{1}{N})$ .  $\square$

The key difficulty in applying this criterion is that it is far from clear why random matrix models that admit a weak asymptotic expansion should also admit a strong asymptotic expansion. A strong asymptotic expansion implies, for example, that weak convergence takes place at the same rate  $\frac{1}{N}$  for polynomial and smooth test functions; it is not at all obvious why this should always be the case. Consequently, applications of this approach were restricted to situations where smooth spectral statistics could be analyzed directly using analytic techniques (such as integration by parts), leaving more complicated models out of reach.

In essence, the punchline of the polynomial method is that under mild assumptions, the existence of a *weak* asymptotic expansion automatically implies the existence of a *strong* asymptotic expansion. This opens the door to establishing strong convergence in a many situations where weak asymptotic expansions are accessible but strong asymptotic expansions had previously remained out of reach.

**5.2 From weak to strong.** We now aim to sketch how the polynomial method works. To simplify the presentation, let us assume a uniform a priori bound  $\|Z^N\| \leq K$  on the random matrices, and that

$$\mathbf{E}[\text{tr } h(Z^N)] = \Phi_h(\frac{1}{N})$$

is given by a polynomial  $\Phi_h$  of degree  $q$  for every polynomial test function  $h \in \mathbb{R}[z]$  of degree  $q$  (so that there is no error term in (5.1)). These two assumptions do not actually hold simultaneously for any random matrix model, so that additional arguments are needed to truncate the expansion or the support of the random matrices. However, we will ignore these issues in order to focus on the core ideas behind the method.

For sake of illustration, we explain how to show that  $\nu_1 : \mathbb{R}[z] \rightarrow \mathbb{R}$  extends continuously to smooth functions. Completely analogous arguments apply to the higher-order terms  $\nu_k$  and to the error term in the expansion.

**Step 1.** The basic observation is that we can view  $\nu_1(h) = \Phi'_h(0)$  as the derivative of the expansion at zero. The secret weapon of the method is the following classical theorem of A. Markov [28, p. 91].

THEOREM 5.2 (Markov). *For any  $f \in \mathbb{R}[z]$  of degree  $q$ , we have  $\|f'\|_{L^\infty[0,a]} \leq \frac{2q^2}{a} \|f\|_{L^\infty[0,a]}$ .*

In the present setting, as  $\|Z^N\| \leq K$ , we have a trivial a priori bound  $|\Phi_h(\frac{1}{N})| = |\mathbf{E}[\text{tr } h(Z^N)]| \leq \|h\|_{L^\infty[-K,K]}$ . This can be exploited by applying Theorem 5.2 to  $\Phi_h$  twice. First, the Markov inequality ensures that  $\Phi_h$  cannot change rapidly between the discrete points  $\frac{1}{N}$ , so that it remains bounded on a continuous interval. Second, applying the Markov inequality again yields a bound on  $\nu_1(h) = \Phi'_h(0)$ . Combining these arguments yields

$$(5.2) \quad |\nu_1(h)| \leq Cq^4 \|h\|_{L^\infty[-K,K]}$$

for any polynomial test function  $h \in \mathbb{R}[z]$  of degree  $q$ , where  $C$  is a universal constant.

**Step 2.** To exploit this estimate, let  $T_k$  be the Chebyshev polynomial defined by  $T_k(\cos \theta) = \cos(k\theta)$ , and express  $h \in \mathbb{R}[z]$  as  $h(x) = \sum_{k=0}^q a_k T_k(x/K)$ . By applying (5.2) to each term separately, we can estimate

$$(5.3) \quad |\nu_1(h)| \leq \sum_{k=0}^q C k^4 |a_k| \lesssim \|h\|_{C^5[-K, K]},$$

where we used that  $\|T_k\|_{L^\infty[-1, 1]} = 1$ . Here the last inequality is a simple fact of Fourier analysis, since  $a_k$  are the Fourier coefficients of the function  $f(\theta) = h(K \cos \theta)$ . We have now accomplished precisely what we wish to show, since (5.3) ensures that  $\nu_1$  extends continuously to every  $h \in C^5(\mathbb{R})$ .

We emphasize that the miracle of (5.2) is that it is the uniform norm  $\|h\|_{L^\infty[-K, K]}$  of  $h(z) = \sum_{k=0}^q b_k z^k$  that appears on the right-hand side, as opposed to the norm of the coefficients  $\sum_{k=0}^q |b_k|$  which is elementary. The latter is typically exponentially larger in  $q$  than the former, which would make it impossible to extend  $\nu_1$  beyond analytic functions. The remarkable feature of the Markov inequality is that it is able to capture cancellations between the coefficients of  $h$ , which is the key to the success of the method.

**Further ingredients.** For expository purposes, we assumed above that  $\Phi_h$  is itself a polynomial. This is not the case in most applications, which requires some adaptations. The polynomial method was initially developed [26] in a setting where  $\Phi_h$  is a rational function, to which the above arguments are readily adapted; see [92, §3] for a self-contained exposition. It was later realized in [69] that the method can be adapted to work assuming only that the weak asymptotic expansion (5.1) holds with a modest estimate on the error term (viz., of Gevrey type  $(q!)^C N^{-q}$  for any  $C > 0$ ), which greatly expands its range of applications. The method can also be adapted to situations where  $\|Z^N\|$  is not uniformly bounded, such as gaussian models [27].

We have not mentioned so far the second ingredient needed by Lemma 5.1, which is a bound on the support of  $\nu_1$ . Such a bound can be achieved using the fact [26, Lemma 4.9], which holds for any compactly supported Schwartz distribution  $\mu$ , that  $\text{supp } \mu \subseteq [-\varrho, \varrho]$  with  $\varrho = \limsup_{p \rightarrow \infty} |\mu(z^p)|^{1/p}$ . Thus the problem reduces to understanding the exponential growth rate of the moments of  $\nu_1$ , which is accessible since these moments can be computed explicitly in practice. Several other tools, such as positivization [67], bootstrapping [27], and supersymmetry arguments [27], have been developed to facilitate this part of the analysis.

**5.3 Open questions.** While the polynomial method has already led to a series of applications [26, 27, 67, 25, 69, 51, 52], its full potential remains unclear. We highlight two general questions in this direction.

First, the polynomial method relies on the weak asymptotic expansion (5.1). While the existence of such an expansion is an easy fact for the most basic random matrix models, it remains an open question whether or not such an expansion holds in many interesting cases. For example, it is unclear which discrete groups admit random permutation representations that have a weak asymptotic expansion.

Second, in various situations where strong convergence is of considerable interest (e.g., random Schreier graphs of finite simple groups of Lie type), a weak asymptotic expansion in the sense of (5.1) does not appear to hold. Could an approach in the spirit of the polynomial method nonetheless address such questions by exploiting other forms of regular behavior of the weak spectral statistics? At present, this question is entirely speculative.

## References

- [1] J. ACHARYA, A. DHARMAVARAPU, Y. LIU, AND N. YU, *Pauli measurements are near-optimal for single-qubit tomography*, 2025. arxiv:2507.22001.
- [2] N. ANANTHARAMAN AND L. MONK, *Friedman-Ramanujan functions in random hyperbolic geometry and application to spectral gaps. I, II*, 2025. arxiv:2304.02678 and arxiv:2502.12268.
- [3] G. W. ANDERSON, *Convergence of the largest singular value of a polynomial in independent Wigner matrices*, Ann. Probab., 41 (2013), pp. 2103–2181.
- [4] J. ANDERSON, *A  $C^*$ -algebra  $\mathcal{A}$  for which  $\text{Ext}(\mathcal{A})$  is not a group*, Ann. of Math. (2), 107 (1978), pp. 455–458.
- [5] T. AUSTIN, *Annealed almost periodic entropy*, 2025. arxiv:2507.08909.
- [6] W. BALLMANN, S. MONDAL, AND P. POLYMERAKIS, *On the spectral stability of finite coverings*, 2025. arxiv:2507.17466.



- [7] A. S. BANDEIRA, M. T. BOEDIHARDJO, AND R. VAN HANDEL, *Matrix concentration inequalities and free probability*, Invent. Math., 234 (2023), pp. 419–487.
- [8] A. S. BANDEIRA, G. CIPOLLONI, D. SCHRÖDER, AND R. VAN HANDEL, *Matrix concentration inequalities and free probability II. Two-sided bounds and applications*, 2024. arxiv:2406.11453.
- [9] A. S. BANDEIRA, K. LUCCA, P. NIZIĆ-NIKOLAC, AND R. VAN HANDEL, *Matrix chaos inequalities and chaos of combinatorial type*, 2025. arxiv:2412.18468.
- [10] M. BANNA, M. CAPITAINE, AND G. CÉBRON, *Strong convergence of multiplicative brownian motions on the general linear group*, 2025. arxiv:2507.13922.
- [11] B. BEKKA, P. DE LA HARPE, AND A. VALETTE, *Kazhdan’s property (T)*, Cambridge, 2008.
- [12] S. BELINSCHI AND M. CAPITAINE, *Strong convergence of tensor products of independent GUE matrices*, 2022. arXiv:2205.07695.
- [13] S. BELINSCHI, B. COLLINS, AND I. NECHITA, *Eigenvectors and eigenvalues in a random subspace of a tensor product*, Invent. Math., 190 (2012), pp. 647–697.
- [14] S. T. BELINSCHI AND M. CAPITAINE, *Spectral properties of polynomials in independent Wigner and deterministic matrices*, J. Funct. Anal., 273 (2017), pp. 3901–3963.
- [15] S. T. BELINSCHI, B. COLLINS, AND I. NECHITA, *Almost one bit violation for the additivity of the minimum output entropy*, Comm. Math. Phys., 341 (2016), pp. 885–909.
- [16] B. BLACKADAR AND E. KIRCHBERG, *Generalized inductive limits of finite-dimensional  $C^*$ -algebras*, Math. Ann., 307 (1997), pp. 343–380.
- [17] C. BORDENAVE AND B. COLLINS, *Eigenvalues of random lifts and polynomials of random permutation matrices*, Ann. of Math. (2), 190 (2019), pp. 811–875.
- [18] C. BORDENAVE AND B. COLLINS, *Norm of matrix-valued polynomials in random unitaries and permutations*, 2024. arxiv:2304.05714.
- [19] C. BORDENAVE AND B. COLLINS, *Strong asymptotic freeness for independent uniform variables on compact groups associated to nontrivial representations*, Invent. Math., 237 (2024), pp. 221–273.
- [20] C. BORDENAVE AND H. LACOIN, *Cutoff at the entropic time for random walks on covered expander graphs*, J. Inst. Math. Jussieu, 21 (2022), pp. 1571–1616.
- [21] T. BRAILOVSKAYA AND R. VAN HANDEL, *Universality and sharp matrix concentration inequalities*, Geom. Funct. Anal., 34 (2024), pp. 1734–1838.
- [22] L. G. BROWN, R. G. DOUGLAS, AND P. A. FILLMORE, *Unitary equivalence modulo the compact operators and extensions of  $C^*$ -algebras*, vol. 345 of Lecture Notes in Math., Springer, 1973, pp. 58–128.
- [23] R. L. BRYANT, *Minimal surfaces of constant curvature in  $S^n$* , Trans. Amer. Math. Soc., 290 (1985), pp. 259–271.
- [24] M. W. BUCK, *Expanders and diffusers*, SIAM J. Algebraic Discrete Methods, 7 (1986), pp. 282–304.
- [25] E. CASSIDY, *Random permutations acting on  $k$ -tuples have near-optimal spectral gap for  $k = \text{poly}(n)$* , 2024. arxiv:2412.13941.
- [26] C.-F. CHEN, J. GARZA-VARGAS, J. A. TROPP, AND R. VAN HANDEL, *A new approach to strong convergence*, Ann. of Math., (2025). To appear.
- [27] C.-F. CHEN, J. GARZA-VARGAS, AND R. VAN HANDEL, *A new approach to strong convergence II. The classical ensembles*, 2025. arxiv:2412.00593.

- [28] E. W. CHENEY, *Introduction to approximation theory*, AMS, Providence, RI, 1998.
- [29] S. Y. CHENG, *Eigenvalue comparison theorems and its geometric applications*, Math. Z., 143 (1975), pp. 289–297.
- [30] B. COLLINS, *Haagerup’s inequality and additivity violation of the minimum output entropy*, Houston J. Math., 44 (2018), pp. 253–261.
- [31] B. COLLINS, A. DAHLQVIST, AND T. KEMP, *The spectral edge of unitary Brownian motion*, Probab. Theory Related Fields, 170 (2018), pp. 49–93.
- [32] B. COLLINS, A. GUIONNET, AND F. PARRAUD, *On the operator norm of non-commutative polynomials in deterministic matrices and iid GUE matrices*, Camb. J. Math., 10 (2022), pp. 195–260.
- [33] B. COLLINS AND C. MALE, *The strong asymptotic freeness of Haar and deterministic matrices*, Ann. Sci. Éc. Norm. Supér. (4), 47 (2014), pp. 147–163.
- [34] B. COLLINS, S. MATSUMOTO, AND J. NOVAK, *The Weingarten calculus*, Notices Amer. Math. Soc., 69 (2022), pp. 734–745.
- [35] B. COLLINS AND W. YUAN, *Strong convergence for tensor GUE random matrices*, 2024. arXiv:2407.09065.
- [36] F. D. CUNDEN, F. MEZZADRI, N. O’CONNELL, AND N. SIMM, *Moments of random matrices and hypergeometric orthogonal polynomials*, Comm. Math. Phys., 369 (2019), pp. 1091–1145.
- [37] J. FRIEDMAN, *Relative expanders or weakly relatively Ramanujan graphs*, Duke Math. J., 118 (2003), pp. 19–35.
- [38] J. FRIEDMAN, *A proof of Alon’s second eigenvalue conjecture and related problems*, Mem. Amer. Math. Soc., 195 (2008), pp. viii+100.
- [39] J. FRIEDMAN, A. JOUX, Y. ROICHMAN, J. STERN, AND J.-P. TILICH, *The action of a few random permutations on  $r$ -tuples and an application to cryptography*, in STACS 96 (Grenoble, 1996), vol. 1046 of Lecture Notes in Comput. Sci., Springer, Berlin, 1996, pp. 375–386.
- [40] M. FUKUDA, T. HASEBE, AND S. SATO, *Additivity violation of quantum channels via strong convergence to semi-circular and circular elements*, Random Matrices Theory Appl., 11 (2022), pp. Paper No. 2250012, 36.
- [41] A. GAMBURD, D. JAKOBSON, AND P. SARNAK, *Spectra of elements in the group ring of  $SU(2)$* , J. Eur. Math. Soc. (JEMS), 1 (1999), pp. 51–85.
- [42] A. GUIONNET, *Asymptotics of random matrices and related models: the uses of Dyson-Schwinger equations*, vol. 130 of CBMS Regional Conference Series in Mathematics, AMS, 2019.
- [43] A. GUIONNET AND D. SHLYAKHTENKO, *Free diffusions and matrix models with strictly convex interaction*, Geom. Funct. Anal., 18 (2009), pp. 1875–1916.
- [44] U. HAAGERUP, *Quasitraces on exact  $C^*$ -algebras are traces*, C. R. Math. Acad. Sci. Soc. R. Can., 36 (2014), pp. 67–92.
- [45] U. HAAGERUP, H. SCHULTZ, AND S. THORBJØRNSSEN, *A random matrix approach to the lack of projections in  $C_{\text{red}}^*(\mathbb{F}_2)$* , Adv. Math., 204 (2006), pp. 1–83.
- [46] U. HAAGERUP AND S. THORBJØRNSSEN, *Asymptotic expansions for the Gaussian unitary ensemble*, Infin. Dimens. Anal. Quantum Probab. Relat. Top., 15 (2012), pp. 1250003, 41.
- [47] U. HAAGERUP AND S. THORBJØRNSSEN, *A new application of random matrices:  $\text{Ext}(C_{\text{red}}^*(F_2))$  is not a group*, Ann. of Math. (2), 162 (2005), pp. 711–775.
- [48] A. HATCHER, *Algebraic topology*, Cambridge University Press, Cambridge, 2002.

- [49] B. HAYES, *A random matrix approach to the Peterson-Thom conjecture*, Indiana Univ. Math. J., 71 (2022), pp. 1243–1297.
- [50] B. HAYES, D. JEKEL, AND S. KUNNAWALKAM ELAYAVALLI, *Consequences of the random matrix solution to the Peterson-Thom conjecture*, Anal. PDE, 18 (2025), pp. 1805–1834.
- [51] W. HIDE, D. MACERA, AND J. THOMAS, *Spectral gap with polynomial rate for random covering surfaces*, 2025. arxiv:2505.08479.
- [52] W. HIDE, D. MACERA, AND J. THOMAS, *Spectral gap with polynomial rate for Weil-Petersson random surfaces*, 2025. arxiv:2508.14874.
- [53] W. HIDE AND M. MAGEE, *Near optimal spectral gaps for hyperbolic surfaces*, Ann. of Math. (2), 198 (2023), pp. 791–824.
- [54] W. HIDE, J. MOY, AND F. NAUD, *On the spectral gap of negatively curved surface covers*, 2025. arxiv:2502.10733.
- [55] S. HOORY, N. LINIAL, AND A. WIGDERSON, *Expander graphs and their applications*, Bull. Amer. Math. Soc. (N.S.), 43 (2006), pp. 439–561.
- [56] H. HUBER, *Über den ersten Eigenwert des Laplace-Operators auf kompakten Riemannschen Flächen*, Comment. Math. Helv., 49 (1974), pp. 251–259.
- [57] D. JEKEL, Y. LEE, B. NELSON, AND J. PI, *Strong convergence to operator-valued semicirculars*, 2025. arxiv:2506.19940.
- [58] H. KESTEN, *Symmetric random walks on groups*, Trans. Amer. Math. Soc., 92 (1959), pp. 336–354.
- [59] C. LANCIEN, *Optimal quantum (tensor product) expanders from unitary designs*, 2024. arxiv:2409.17971.
- [60] C. LANCIEN AND P. YOUSSEF, *A note on quantum expanders*, 2023. arxiv:2302.07772.
- [61] F. LEHNER, *Computing norms of free operators with matrix coefficients*, Amer. J. Math., 121 (1999), pp. 453–486.
- [62] L. LOUDER AND M. MAGEE, *Strongly convergent unitary representations of limit groups*, J. Funct. Anal., 288 (2025), p. Paper No. 110803.
- [63] E. LUBETZKY AND Y. PERES, *Cutoff on all Ramanujan graphs*, Geom. Funct. Anal., 26 (2016), pp. 1190–1216.
- [64] A. LUBOTZKY, R. PHILLIPS, AND P. SARNAK, *Ramanujan graphs*, Combinatorica, 8 (1988), pp. 261–277.
- [65] M. MAGEE, *Strong convergence of unitary and permutation representations of discrete groups*, 2024. Proceedings of the ECM, to appear.
- [66] M. MAGEE AND M. DE LA SALLE,  *$SL_4(\mathbf{Z})$  is not purely matricial field*, C. R. Math. Acad. Sci. Paris, 362 (2024), pp. 903–910.
- [67] M. MAGEE AND M. DE LA SALLE, *Strong asymptotic freeness of Haar unitaries in quasi-exponential dimensional representations*, 2024. arXiv:2409.03626.
- [68] M. MAGEE AND D. PUDER, *The asymptotic statistics of random covering surfaces*, Forum Math. Pi, 11 (2023), pp. Paper No. e15, 51.
- [69] M. MAGEE, D. PUDER, AND R. VAN HANDEL, *Strong convergence of uniformly random permutation representations of surface groups*, 2025. arxiv:2504.08988.
- [70] M. MAGEE AND J. THOMAS, *Strongly convergent unitary representations of right-angled Artin groups*, 2023. arxiv:2308.00863.

- [71] C. MALE, *The norm of polynomials in large random and deterministic matrices*, Probab. Theory Related Fields, 154 (2012), pp. 477–532. With an appendix by Dimitri Shlyakhtenko.
- [72] M. MIRZAKHANI, *Growth of Weil-Petersson volumes and random hyperbolic surfaces of large genus*, J. Differential Geom., 94 (2013), pp. 267–300.
- [73] A. MONTANARI AND E. RICHARD, *A statistical model for tensor PCA*, in Proceedings of the 27th Conference on Neural Information Processing Systems, NIPS’14, MIT Press, 2014, pp. 2897–2905.
- [74] J. D. MOORE, *Introduction to global analysis: minimal surfaces in Riemannian manifolds*, AMS, 2017.
- [75] S. C. MORAMPUDI AND C. R. LAUMANN, *Many-body systems with random spatially local interactions*, Phys. Rev. B, 100 (2019), p. 245152.
- [76] J. MOY, *Spectral gap of random covers of negatively curved noncompact surfaces*, 2025. arxiv:2505.07056.
- [77] A. NICA AND R. SPEICHER, *Lectures on the combinatorics of free probability*, Cambridge, 2006.
- [78] N. OZAWA, *Amenability for unitary groups of simple monotracial  $C^*$ -algebras*, 2023. arxiv:2307.08267.
- [79] F. PARRAUD, *On the operator norm of non-commutative polynomials in deterministic matrices and iid Haar unitary matrices*, 2021. arxiv:2005.13834.
- [80] F. PARRAUD, *Asymptotic expansion of smooth functions in deterministic and iid Haar unitary matrices, and application to tensor products of matrices*, 2023. arxiv:2302.02943.
- [81] F. PARRAUD, *Asymptotic expansion of smooth functions in polynomials in deterministic matrices and iid GUE matrices*, Comm. Math. Phys., 399 (2023), pp. 249–294.
- [82] F. PARRAUD, *The spectrum of a tensor of random and deterministic matrices*, 2024. arXiv:2410.04481.
- [83] J. PETERSON AND A. THOM, *Group cocycles and the ring of affiliated operators*, Invent. Math., 185 (2011), pp. 561–592.
- [84] G. PISIER, *A simple proof of a theorem of Kirchberg and related results on  $C^*$ -norms*, J. Operator Theory, 35 (1996), pp. 317–335.
- [85] G. PISIER, *Random matrices and subexponential operator spaces*, Israel J. Math., 203 (2014), pp. 223–273.
- [86] G. PISIER, *On a linearization trick*, Enseign. Math., 64 (2018), pp. 315–326.
- [87] I. RIVIN AND N. T. SARDARI, *Quantum chaos on random Cayley graphs of  $SL_2[\mathbb{Z}/p\mathbb{Z}]$* , Exp. Math., 28 (2019), pp. 328–341.
- [88] J. SALEZ, *Modern aspects of Markov chains: entropy, curvature and the cutoff phenomenon*, 2025. arxiv:2508.21055.
- [89] H. SCHULTZ, *Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases*, Probab. Theory Related Fields, 131 (2005), pp. 261–309.
- [90] A. SONG, *Random harmonic maps into spheres*, 2025. arxiv:2402.10287.
- [91] J. A. TROPP, *Second-order matrix concentration inequalities*, Appl. Comput. Harmon. Anal., 44 (2018), pp. 700–736.
- [92] R. VAN HANDEL, *The strong convergence phenomenon*, in Current Developments in Mathematics 2026.
- [93] D. VOICULESCU, *Around quasidiagonal operators*, Integral Eq. Operator Theory, 17 (1993), pp. 137–149.
- [94] D. VOICULESCU, *The topological version of free entropy*, Lett. Math. Phys., 62 (2002), pp. 71–82.
- [95] A. S. WEIN, A. EL ALAOU, AND C. MOORE, *The Kikuchi hierarchy and tensor PCA*, in 2019 IEEE 60th Annual Symposium on Foundations of Computer Science, IEEE Press, 2019, pp. 1446–1468.