A RANDOM MATRIX APPROACH TO ABSORPTION IN FREE PRODUCTS

BEN HAYES, DAVID JEKEL, BRENT NELSON, AND THOMAS SINCLAIR

ABSTRACT. This paper gives a free entropy theoretic perspective on amenable absorption results for free products of tracial von Neumann algebras. In particular, we give the first free entropy proof of Popa’s famous result that the generator MASA in a free group factor is maximal amenable, and we partially recover Houdayer’s results on amenable absorption and Gamma stability. Moreover, we give a unified approach to all these results using 1-bounded entropy. We show that if $\mathcal{M} = \mathcal{P} \ast \mathcal{Q}$, then $\mathcal{P}$ absorbs any subalgebra of $\mathcal{M}$ that intersects it diffusely and that has 1-bounded entropy zero (which includes amenable and property Gamma algebras as well as many others). In fact, for a subalgebra $\mathcal{P} \subseteq \mathcal{M}$ to have this absorption property, it suffices for $\mathcal{M}$ to admit random matrix models that have exponential concentration of measure and that “simulate” the conditional expectation onto $\mathcal{P}$.

INTRODUCTION

A foundational discovery of Popa [Pop83a] showed that the generator MASA $\mathcal{A} = L(\mathbb{Z})$ in a free group factor $\mathcal{M} = L(Z * F_{d-1})$ is a maximal amenable subalgebra (or equivalently a maximal hyperfinite subalgebra by [Con76]). This provided the first example of an abelian subalgebra of a II$_1$ factor that is maximal amenable (it is even maximal property Gamma), and it answered in the negative the question, stated by Kadison at the 1967 Baton Rouge conference, of whether any self-adjoint operator in a II$_1$ factor is contained in a hyperfinite subfactor. The fundamental insight of Popa was a detailed analysis of the $\mathcal{A}$-central sequences in $\mathcal{M}$ through his asymptotic orthogonality property. The approach of Popa is very fruitful and many authors have built on his work to establish maximal amenability in several cases, see [Ge96, She06, Pan07, CFRW10, Gao10, Bro14, Hou14a, Hou14b, Lea19]. Several other celebrated structural results for free group factors can be shown using Popa’s breakthrough deformation/rigidity theory initiated in [Pop06c, Pop06a, Pop06d, Pop06e]. Moreover, deformation/rigidity theory allows one to prove structural results not just for free group factors, but also von Neumann algebras of more general groups. We refer the reader to [Pop07, Vae07, Vae10, Ioa14, Ioa] for a survey of these results, many of which include solutions to long-standing open problems. This theory has been particularly fruitful in proving absence of Cartan subalgebras, as well as uniqueness of Cartan subalgebras for crossed product algebras. See, e.g., [OP10a, OP10b, CP13, CS13, PV14a, PV14b, Ioa15].

In this paper, we present an approach to structural results for free products through Voiculescu’s free entropy dimension theory. Free entropy dimension theory was initiated by Voiculescu in a series of papers [Vo94, Vo96], and gives powerful tools to prove various indecomposability and structural results on free group factors. In one celebrated achievement, Voiculescu gave the first proof of absence of Cartan subalgebras for free group factors using free entropy methods [Vo96, Theorem 5.2]. Shortly thereafter Ge [Ge98] and Ge and Popa [GP98] used this machinery to prove that free group factors are prime and thin (cannot be decomposed as a product of hyperfinite factors), respectively. See [Dyk97, Jun07b, HS11] for other applications of free entropy dimension to the structure of free group factors. While Popa’s deformation/rigidity theory applies to a wider class of algebras than free products, and does not require that the algebras have the Connes approximate embedding property, the tools of free entropy dimension are well-suited for studying free group factors and free products of Connes embeddable algebras due to the presence of natural random matrix models that come from Voiculescu’s asymptotic freeness theorem [Vo91, Vo98]. Additionally, there are certain indecomposability results for free group factors that can be proved using free entropy techniques which are currently out of reach by other methods. (See, e.g., [Hay18] Theorem 1.3, Corollary 1.4, Corollary 1.7, as well as [Dyk97, GP98].)

B. Hayes was partially supported by NSF grants DMS-1600802 and DMS-1827376. D. Jekel was partially supported by NSF grant DMS-1762360. B. Nelson was partially supported by NSF grant DMS-1856683. T. Sinclair was partially supported by NSF grant DMS-1600867.
Given the success of free entropy theory and random matrix theory in proving other structural results for free products, it is reasonable to suppose that the maximal amenability of the generator MASA in a free group factor can be proved using random matrices and free entropy dimension. We provide such a proof using the notion of 1-bounded entropy, which is implicit in [Jun07b] and explicitly defined in [Hay18]. The 1-bounded entropy is defined similarly to free entropy dimension, but it is useful specifically for studying von Neumann algebras with free entropy dimension 1; the 1-bounded entropy is less than \(+\infty\) if and only if the algebra is strongly 1-bounded in the sense of Jung [Jun07b]. The 1-bounded entropy has the advantage that it is known to be an invariant for tracial von Neumann algebras and can be computed on any set of generators. This invariant provides a unified and efficient way to recast the proofs of Voiculescu and Ge on structural properties of free group factors, see Section 1.2.

We now state the main result of the paper. As a convention, when the context is clear we will write “\(N \leq M\)” to indicate that \(N\) is a von Neumann subalgebra of \(M\).

**Theorem A.** Let \(M_1\) and \(M_2\) be tracial W*-algebras such that every separable subalgebra is embeddable into \(\mathcal{R}^\omega\). Let \((D,\tau_D)\) be a common subalgebra of \(M_1\) and \(M_2\) which is totally atomic (that is, a direct sum of at most countably many matrix algebras), and consider the amalgamated free product \(M = M_1 *_D M_2\). Then for any \(N \leq M\) with \(h(N : M) = 0\) and \(N \cap M_1\) diffuse, one has \(N \subseteq M_1\).

We remark that if \(D\) is diffuse, then the conclusions of Theorem A can fail for rather trivial reasons (in contrast to the results of Brown–Dykema–Jung [BDJ08] on free entropy dimension for free products with amalgamation over any amenable \(D\)). For instance, suppose that \(D\) is diffuse abelian, let \(K_1\) and \(K_2\) be finite-dimensional algebras, and let \(M_1 = D \otimes K_1\) and \(M_2 = D \otimes K_2\). Then, in the algebra \(M\), the intersection \(M_1 \cap M_2\) is diffuse and \(h(M_2 : M) = 0\), yet \(M_2\) is not contained in \(M_1\).

The maximal amenability of the generator MASA in a free group factor is obtained since the above result implies that it is a maximal subalgebra with 1-bounded entropy zero, and ameniable von Neumann algebras always have 1-bounded entropy zero. This proof will have other consequences besides maximal amenability, since there are several other classes of von Neumann algebras that have 1-bounded entropy \(\leq 0\) (zero if Connes-embeddable and \(-\infty\) otherwise). This includes any von Neumann algebra which is non-prime, any von Neumann algebra so that every nonzero direct summand has property Gamma, as well as any von Neumann algebra that has a Cartan subalgebra (see [J12]). Thus, for instance, our proof implies that the generator MASA is a maximal subalgebra with property Gamma. In this way, our approach gives a unified proof for amenability, property Gamma, having Cartan subalgebras, and so forth. However, a significant limitation is that we can only handle free absorption results when each side is finite, and so for instance, our methods cannot reach type III algebras. Moreover, unlike the asymptotic orthogonality or deformation/rigidity approaches, we have to assume \(M_1, M_2\) both embed into an ultrapower of \(\mathcal{R}\), and our results do not have concrete applications beyond free products.

The conclusion of Theorem A relates to previous results about amenable absorption. For example, in [Hou15] Houdayer defined a generalization of Popa’s asymptotic orthogonal property now called the strong asymptotic orthogonality property from which one can deduce an amenable absorption property: if \(Q \leq M\) is amenable and \(Q \cap A\) is diffuse, then \(Q \subseteq A\). This strengthening of the asymptotic orthogonality property was then used in [Wen16, BW16, PSW18] to prove that one has amenable absorption for maximal amenable subalgebras of free group factors. We remark in passing that it is a conjecture of Peterson and Thom (see the discussion following [PT11, Proposition 7.7]) that every maximal amenable subalgebra of a free group factor has the amenable absorption property. There is also another recent framework for analyzing maximal amenability in terms of singular states due to Boutonnet–Carderi [BC15], which was modified by Ozawa to give a short proof of maximal amenability of the generator MASA in [Oza15]. This was carried out to great effect by Boutonnet–Houdayer in [BHS18] to give complete results on amenable absorption in amalgamated free products.

Since amenable algebras always have 1-bounded entropy zero, Theorem A recovers these amenable absorption results for certain amalgamated free products. Moreover, since factors with property Gamma also have 1-bounded entropy zero, we recover some of the Gamma stability results shown in [Hou15] and more. For example: if \(Q \leq M_1 * M_2\) and \(Q \cap M_1\) is diffuse, and if \(Q\) has 1-bounded entropy zero, then \(Q \subseteq M_1\). So if \(M_1\) is non-prime then \(M_1\) absorbs any \(Q\) that is non-prime, has the property that every nonzero direct summand has property Gamma, or has a Cartan, provided that \(Q\) intersects \(M_1\) diffusely. In the spirit of Gamma stability, we can replace the assumption that \(Q\) is non-prime, or that every nonzero direct
summand has property Gamma, or that it has a Cartan, with the requirement that $Q' \cap M^\omega$ is diffuse and still conclude that $Q \subseteq M_1$ (provided $Q \cap M_1$ is diffuse). Additionally, by [Hou14a, HS11, Hay18] one can find von Neumann algebras which are strongly solid, have the Haagerup property, and the complete metric approximation property and yet still have 1-bounded entropy zero, and so similar remarks apply to these examples.

The proof of Theorem [A] proceeds (modulo a quick reduction to the separable case) by constructing random matrix models that have certain properties relative to the inclusion $M_1 \subseteq M_1 \ast M_2$. In fact, all our conclusions hold for an inclusion $\mathcal{P} \subseteq \mathcal{M}$ if we only assume the existence of random matrix models satisfying these properties. We state this result as a theorem in its own right.

**Theorem B.** Suppose that $\mathcal{P} \subseteq \mathcal{M}$ is an inclusion of tracial $W^*$-algebras. Let $\mathbf{x} = (x_i)_{i \in I}$ be a tuple of bounded self-adjoint generators for $\mathcal{M}$. Suppose $n(k) \to \infty$ and that $\mathbf{X}(k) = (X_i^{(k)})_{i \in I}$ is a tuple of random $n(k) \times n(k)$ self-adjoint matrices satisfying the following properties:

1. We have $\|X_i^{(k)}\| \leq R_i$ for some $R_i$ independent of $k$.
2. For every non-commutative polynomial $p \in \mathbb{C}(t_i : i \in I)$, we have $\tau_{n(k)}(p(\mathbf{X}(k))) \to \tau(p(\mathbf{x}))$ in probability.
3. The probability distributions of $(X_i^{(k)})_{i \in F}$ exhibit exponential concentration of measure at the scale $n(k)^2$ as $k \to \infty$ for each finite $F \subseteq I$. (See [L3, §2.3] for the precise definition.)
4. For each non-commutative polynomial $p \in \mathbb{C}(t_i : i \in I)$, we have
   $$\lim_{k \to \infty} \|E_p[\mathbf{X}(k)]\|_2 = \|E_p[\mathbf{x}]\|_2.$$ 

Here $E_\mathcal{P}$ is the trace-preserving conditional expectation of $\mathcal{M}$ onto $\mathcal{P}$.

Then for any $W^*$-algebra $\mathcal{N} \subseteq \mathcal{M}$ with $h(\mathcal{N} : \mathcal{M}) = 0$ and $\mathcal{N} \cap \mathcal{P}$ diffuse, one has $\mathcal{N} \subseteq \mathcal{P}$.

The hypotheses (1) and (2) are standard descriptions of how random matrix models simulate the given operators $\mathbf{x}$ in the large $k$ limit. Exponential concentration (3) is also familiar in random matrix theory (at least in the case when $I$ is finite); see for instance [AGZ09, §2.3 and §4.4]. The hypothesis (4) is new to our work, and it says intuitively that the probabilistic expectation for the random matrix models simulates the (trace-preserving) conditional expectation of $\mathcal{M}$ onto $\mathcal{P}$ rather than the trace.

In the case of Theorem [A] where $\mathcal{M} = M_1 \ast M_2$ and $\mathcal{P} = M_1$, such random matrix models will be constructed by adapting the methods of Brown–Dykema–Jung [BDJ08], who gave a version of Voiculescu’s asymptotic freeness results for amalgamation over a finite-dimensional algebra.

The application of Theorem [B] in the case where $h(\mathcal{P} : \mathcal{M}) = 0$ naturally motivates the following definition.

**Definition.** Let $(\mathcal{M}, \tau)$ be a tracial von Neumann algebra. We say that a diffuse $\mathcal{P} \subseteq \mathcal{M}$ is a Pinsker algebra if $h(\mathcal{P} : \mathcal{M}) = 0$, and given any $\mathcal{P} \subseteq \mathcal{Q} \subseteq \mathcal{M}$ with $h(\mathcal{Q} : \mathcal{M}) = 0$ we have $\mathcal{Q} = \mathcal{P}$.

It follows from general properties of 1-bounded entropy that if $\mathcal{Q} \subseteq \mathcal{M}$ is diffuse and $h(\mathcal{Q} : \mathcal{M}) \leq 0$, then there is a unique Pinsker algebra $\mathcal{P} \subseteq \mathcal{M}$ with $\mathcal{P} \supseteq \mathcal{Q}$. In particular, two Pinsker algebras which intersect diffusely are equal. We refer the reader to Section [1,2] for a more detailed discussion. We also discuss in Section [1,2] the motivation for the terminology “Pinsker algebra,” which comes from entropy in ergodic theory. Hypotheses (2)-[1] of Theorem [B] have natural interpretations in ergodic theory. For example, exponential concentration in the ergodic theory context naturally leads to the “almost blowing up property,” a condition on a probability measure-preserving action which is equivalent to the action being Bernoulli. We refer the reader to [MS94] as well as [Sh98, Section III.4] for more information. We remark that [Hay18, Theorem 3.8] (which is the main result of that paper) can be rephrased as saying that if $\mathcal{P} \subseteq \mathcal{M}$ is Pinsker, then $\mathcal{P} \subseteq \mathcal{M}$ is coarse in the sense of Popa [Pop]. So Pinsker algebras connect to coarse embeddings, and thus also to the coarseness conjecture independently formulated by the first named author [Hay18, Conjecture 1.12] and Popa [Pop, Conjecture 5.2]. As a corollary of Theorem [B] we obtain examples of Pinsker algebras in amalgamated free products.

**Corollary.** Let $\mathcal{P} \subseteq \mathcal{M}$ be as in Theorem [2] and suppose that $h(\mathcal{P} : \mathcal{M}) = 0$. Then $\mathcal{P}$ is a Pinsker algebra. In particular, if $\mathcal{M} = M_1 \ast \mathcal{M}_2$ is as in Theorem [2] and $h(M_1 : \mathcal{M}) = 0$, then $M_1$ is a Pinsker algebra in $\mathcal{M}$.

We remark that the proofs of Theorems [A] and [B] are designed to work even with infinite generating sets (that is, the index set $I$ in Theorem [B] may be infinite). Of course, the statement and proof would be...
marginally simpler to write down in the finitely generated case, and random matrix results often use finite generating sets. However, some of the most natural applications of 1-bounded entropy and Pinsker algebras deal with cases where we do not know a priori that the W*-algebras are finitely generated (and it is unknown whether every tracial W*-algebra with separable predual is finitely generated). That is why we have taken care to set up our machinery to handle infinite generating sets. We also show that the existence of random matrix models satisfying the hypotheses of Theorem B is independent of the choice of generators for M; see Proposition 6.2 and the discussion preceding it.

The rest of the paper is organized as follows. In §1 we review background, definitions, and notation. In particular, we summarize the properties of 1-bounded entropy that are needed for the applications of our main theorems discussed in the introduction. In §2 we present a multivariable functional calculus, based on the work of the second author in [Jek18, Jek19], that will be used as a technical tool to transform between microstate spaces for different choices of generators. In §3 we prove Theorem B. In §4 we apply Theorem B to the generator MASA in free group factors and we describe the relationship with previous work on random matrix models with convex potentials. In §5, we prove Theorem A which handles the case of free products with amalgamation over totally atomic subalgebras. In §6 we further discuss the random matrix models used in Theorem B relating them with other ideas from von Neumann algebras and ergodic theory.

Acknowledgements. The initial stages of this work were carried out at the Hausdorff Research Institute for Mathematics during the 2016 trimester program “Von Neumann Algebras.” The first, third, and last named authors thank the Hausdorff Institute for their hospitality. The first named author thanks Tim Austin and Lewis Bowen for discussions related to exponential concentration of measure in the sofic entropy context which were inspirational for this work. In particular, he thanks Bowen for sharing notes on exponential concentration of measure in the sofic entropy context which led to the initial stages of this paper. All the authors thank the Banff International Research Station for their hospitality during the workshop “Classification Problems in von Neumann Algebras” in Fall 2019.

1. Background

1.1. Algebras, Traces, and Laws. We begin by recalling some basic definitions, concepts, and notations from operator algebras.

A tracial W*-algebra is pair (M, τ), where M is a W*-algebra and τ is a faithful, normal, tracial state. This is equivalent (in the separable case) to a finite von Neumann algebra with a designated choice of faithful normal trace. We denote by Msa the set of self-adjoint elements of M. We denote ∥x∥2 = τ(x∗x)1/2.

We will view M as an algebra of “bounded random variables” and τ as the “expectation.” Given (M, τ) and a W*-subalgebra N, there is a unique trace-preserving conditional expectation E_N : M → N. In this paper, the default meaning of “conditional expectation M → N” will be this trace-preserving conditional expectation.

Given a (possibly infinite) index set I, we want to discuss the non-commutative law of a tuple (x_i)_{i ∈ I} ∈ Msa^I, as well as a topology on the space of non-commutative laws. These laws are defined in terms of the trace of non-commutative polynomials in the variables (x_i)_{i ∈ I}.

Let C⟨t_i : i ∈ I⟩ be the algebra of non-commutative complex polynomials in (t_i)_{i ∈ I} (i.e. the free C-algebra on the set I). We give C⟨t_i : i ∈ I⟩ the unique *-algebra structure which makes the t_i self-adjoint. If M is a W*-algebra and x = (x_i)_{i ∈ I} ∈ Msa^I, then there is a unique *-homomorphism ev_x : C⟨t_i : i ∈ I⟩ → M such that ev_x(t_i) = x_i. For a non-commutative polynomial p ∈ C⟨t_i : i ∈ I⟩, we define p(x) = p((x_i)_{i ∈ I}) to be ev_x(p).

A tracial non-commutative law of a self-adjoint I-tuple is a linear functional λ : C⟨t_i : i ∈ I⟩ → C that is

1. unital; that is, λ(1) = 1,
2. positive; that is, λ(p∗p) ≥ 0,
3. tracial; that is, λ(pq) = λ(qp),
4. exponentially bounded; that is, for some (R_i)_{i ∈ I} ∈ (0, +∞)^I, we have

|λ(t_{i(1)}...t_{i(ℓ)})| ≤ R_{i(1)}...R_{i(ℓ)}

for all ℓ and all i(1), ..., i(ℓ) ∈ I.
Given $R = (R_i)_{i \in I} \in (0, +\infty)^I$, we define $\Sigma_R = \Sigma(R_i)_{i \in I}$ to be the set of non-commutative laws satisfying (4) for our given choice of $(R_i)_{i \in I}$. We equip $\Sigma_R$ with the topology of pointwise convergence on $\mathbb{C}(t_i : i \in I)$.

In this topology, $\Sigma_R$ is a compact Hausdorff space, and if $I$ is countable, then $\Sigma_R$ is metrizable.

Given a tracial $W^*$-algebra $(M, \tau)$, a tuple $x = (x_i)_{i \in I} \in M_{sa}^I$, and $R = (R_i)_{i \in I} \in (0, +\infty)^I$ satisfying $\|x_i\| \leq R_i$, we define the non-commutative law of $x$ as the map

$$\lambda_x : \mathbb{C}(t_i : i \in I) \to \mathbb{C} : p \mapsto \tau(p(x)).$$

It is straightforward to verify that $\lambda_x$ is in $\Sigma_R$. Conversely, given any $\lambda \in \Sigma_R$, there exists some $(M, \tau)$ and $x \in M_{sa}^I$ such that $\lambda_x = \lambda$ and $\|x_i\| \leq R_i$ for all $i \in I$. This follows from a version the GNS construction; the proof is the same as in [AGZ09, Proposition 5.2.14(d)]. We also remark that $(M, \tau)$ could be $M_n(\mathbb{C})$ with the normalized trace $\tau_n = (1/n)\text{Tr}$. Thus, if $x \in M_n(\mathbb{C})_{sa}$, then $\lambda_x$ is a well-defined non-commutative law.

At several points in the paper we will use the following folklore result, whose proof we leave as an exercise to the reader. See, e.g., [Jun07a, Lemma 2.9] or [AK19, Lemma 1.10 and Proposition 1.7].

**Lemma 1.1.** Let $(M, \tau)$ be a hyperfinite tracial von Neumann algebra with separable predual. Let $x \in M_{sa}^I$. Suppose that $(n(k))_k$ is a sequence of positive integers with $n(k) \to \infty$. If $A^{(k)} \in M_{n(k)}(\mathbb{C})_{sa}^I$ satisfy

$$\sup_k \|A^{(k)}_i\|^2, \sup_k \|B^{(k)}_i\|^2 < \infty \quad \text{for all } i \in I, \quad \text{and } \lambda_{A^{(k)}} \to \lambda_x, \lambda_{B^{(k)}} \to \lambda_x$$

then there is a sequence $U^{(k)} \in M_{n(k)}(\mathbb{C})$ with

$$\|U^{(k)}_i A^{(k)}_i (U^{(k)}_i)^* - B^{(k)}_i\|_2 \to_k \to \infty 0,$$

for all $i \in I$.

1.2. **Microstate Spaces and 1-Bounded Entropy.** Microstates free entropy as well as 1-bounded entropy are defined by looking at the exponential growth rate of volumes/covering numbers of so-called matricial microstate spaces, that is, spaces of matrices that have approximately the same non-commutative law as a given tuple $y$.

Let $R \in (0, +\infty)^I$, let $y \in M_{sa}^I$ be a self-adjoint tuple with $\|y_j\| \leq R_j$, and let $U \subseteq \Sigma_R$ be a neighborhood of $\lambda_y$. Then we define the microstate space

$$\Gamma_{R,n}(y : U) := \{B \in M_n(\mathbb{C})_{sa}^I : \|B_j\| \leq R_j \text{ for all } j \} \in U\}.$$

These microstate spaces are invariant under conjugating a tuple $B$ by a fixed $n \times n$ unitary matrix. In order to remove the inherent ambiguity of unitary equivalence, we will often fix a self-adjoint $z \in M$ and fix a sequence $(C^{(n)})_{n \in \mathbb{N}}$ such that $C^{(n)} \in M_n(\mathbb{C})$, $\|C^{(n)}\|$ is bounded, and $\lambda_{C^{(n)}} \to \lambda_z$. For shorthand, we will often denote this as $C^{(n)} \to z$. Thus, we define microstate spaces for $y$ relative to $C^{(n)} \to z$ as follows: For each $R \in (0, +\infty)^I_{<0}$ with $\|y_j\| < R_j$ and $\|z\| < R_0$, for each sequence of microstates $C^{(n)}$ for $z$ as above with $\|C^{(n)}\| < R_0$, and for each neighborhood $U$ of $\lambda_{y,z} \in \Sigma_R$, define

$$\Gamma_{R,n}(y : C^{(n)} \to z : U) := \{B \in M_n(\mathbb{C})_{sa}^I : \|B_j\| \leq R_j, \lambda_{y,B,C^{(n)}} \in U\}.$$
However, to minimize the technical background, we will state all our definitions directly for this particular case rather than importing the entire formalism of uniform structures.

Given $A \in M_n(\mathbb{C})_{sa}^I$, we define for each finite $F \subseteq I$ and $\varepsilon > 0$,

$$N_{F,\varepsilon}(A) = \{ B \in M_n(\mathbb{C})_{sa}^I : \| A_i - B_i \|_2 < \varepsilon \text{ for all } i \in F \},$$

and we refer to this set as the $(F, \varepsilon)$-neighborhood centered at $A$. Similarly, we define the $(F, \varepsilon)$-neighborhood of a set $\Omega \subseteq M_n(\mathbb{C})_{sa}^I$

$$N_{F,\varepsilon}(\Omega) = \bigcup_{A \in \Omega} N_{F,\varepsilon}(A),$$

that is, the set of $B$ such that there exists $A \in \Omega$ with $\| A_i - B_i \|_2 < \varepsilon$ for all $i \in F$.

For a set $\Omega \subseteq M_n(\mathbb{C})_{sa}^I$, we define the covering number

$$K_{F,\varepsilon}(\Omega) = \inf \{|\Omega_0| : \Omega_0 \subseteq \Omega \text{ and } \Omega \subseteq N_{F,\varepsilon}(\Omega_0)\}.$$ 

In other words, this is the minimum number of points whose $(F, \varepsilon)$-neighborhoods cover $\Omega$. (This will be finite if $\Omega$ is contained in a product of $\| \cdot \|_2$-balls $\{ B_i(x_i, R_i) \}_{i \in I}$.)

**Definition 1.2.** Let $(M, \tau)$ be a diffuse tracial $W^*$-algebra and $\mathcal{N} \subseteq \mathcal{M}$. Fix some $z \in \mathcal{N}_{sa}$ with diffuse spectrum, some $R_0 > \| z \|$, and a sequence $C^{(n)} \in M_n(\mathbb{C})$ such that $\| C^{(n)} \| \leq R_0$ and $C^{(n)} \tau \rightarrow z$. Fix a tuple $y \in \mathcal{N}_{sa}^J$ of generators for $\mathcal{N}$ and a tuple $x \in \mathcal{M}_{sa}^I$ of generators for $\mathcal{M}$. Fix $R \in (0, +\infty)^{I \cup J}$ and $R_i > \| x_i \|$ for $i \in I$ and $R_j > \| y_j \|$ for $j \in J$. Then we define

$$h(\mathcal{N} : \mathcal{M}) := \sup_{\mathcal{U}} \sup_{n \rightarrow \infty} \frac{1}{n^2} \log K_{F,\varepsilon}(\Gamma_{R,n}(y : x | C^{(n)} \tau \rightarrow z ; \mathcal{U})),$$

where the infimum is taken over all neighborhoods $\mathcal{U}$ of $\lambda_{(x,y,z)}$, and the supremum is taken over all $F \subseteq I$ finite and all $\varepsilon > 0$.

It is implicit from [Jun07] Theorem 3.2 and explicitly shown in [Hay18] Theorem A.9 that $h(\mathcal{N} : \mathcal{M})$ is well-defined in the sense that it is independent of the choice of generators $x$ and $y$, the choice of $z \in \mathcal{N}_{sa}$, and the choice of microstates $C^{(n)}$. The fact that it is independent of the choice of $R$ follows from completely standard techniques (compare [Voi94] proof of Proposition 2.4 and [BB03]). We call $h(\mathcal{N} : \mathcal{M})$ the 1-bounded entropy in the presence of $\mathcal{M}$. We write $h(\mathcal{M}) := h(\mathcal{M} : \mathcal{M})$ call this the 1-bounded entropy of $\mathcal{M}$.

We list here some important properties of 1-bounded entropy, along with pointers in the literature to where they are proved. All von Neumann algebras listed in the properties below are assumed to be tracial and embeddings are assumed to be trace-preserving.

**P1:** $h(\mathcal{N} : \mathcal{M}) \geq 0$ if $\mathcal{N} \subseteq \mathcal{M}$ and every von Neumann subalgebra of $\mathcal{M}$ with separable predual embeds into an ultrapower of $\mathcal{R}$, and $h(\mathcal{N} : \mathcal{M}) = -\infty$ if there exists a von Neumann subalgebra of $\mathcal{M}$ with separable predual which does not embed into an ultrapower of $\mathcal{R}$. (Exercise from the definitions.)

**P2:** $h(\mathcal{N}_1 : \mathcal{M}_1) \leq h(\mathcal{N}_2 : \mathcal{M}_2)$ if $\mathcal{N}_1 \subseteq \mathcal{N}_2 \subseteq \mathcal{M}_2 \subseteq \mathcal{M}_1$, if $\mathcal{N}_1$ is diffuse. (Exercise from the definitions.)

**P3:** $h(\mathcal{N} : \mathcal{M}) = 0$ if $\mathcal{N} \subseteq \mathcal{M}$ and $\mathcal{N}$ is diffuse and hyperfinite. (Exercise from Lemma 1.1 using the “orbital” definition of 1-bounded entropy given in [Hay18] Corollary A.6.)

**P4:** For $\mathcal{M}$ diffuse, $h(\mathcal{M}) < \infty$ if and only if $\mathcal{M}$ is strongly 1-bounded in the sense of Jung. (See [Hay18] Proposition A.16.)

**P5:** $h(\mathcal{M}) = \infty$ if $\mathcal{M} = W^*(x_1, \ldots, x_n)$ where $x_j \in \mathcal{M}_{sa}$ for all $1 \leq j \leq n$ and $\delta(x_1, \ldots, x_n) > 1$. For example, this applies if $\mathcal{M} = L(\mathcal{F}_n)$, for $n > 1$. (This follows from Property 4 and Jun07 Corollary 3.5.))

**P6:** $h(\mathcal{N}_1 \vee \mathcal{N}_2 : \mathcal{M}) \leq h(\mathcal{N}_1 : \mathcal{M}) + h(\mathcal{N}_2 : \mathcal{M})$ if $\mathcal{N}_1, \mathcal{N}_2 \subseteq \mathcal{M}$ and $\mathcal{N}_1 \cap \mathcal{N}_2$ is diffuse. (See [Hay18] Lemma A.12.)

**P7:** Suppose that $(\mathcal{N}_\alpha)_\alpha$ is an increasing chain of diffuse von Neumann subalgebras of a von Neumann algebra $\mathcal{M}$. Then

$$h \left( \bigvee_\alpha \mathcal{N}_\alpha : \mathcal{M} \right) = \sup_\alpha h(\mathcal{N}_\alpha : \mathcal{M}).$$

(See [Hay18] Lemma A.10.)

**P8:** $h(\mathcal{N} : \mathcal{M}) = h(\mathcal{N} : \mathcal{M}^\omega)$ if $\mathcal{N} \subseteq \mathcal{M}$ is diffuse, and $\omega$ is a free ultrafilter on an infinite set. (See [Hay18] Proposition 4.5.)
P9: $h(W^*(N_M(N)) : M) = h(N : M)$ if $N \leq M$ is diffuse. Here $N_M(N) = \{u \in U(M) : uN^*u = N\}$. (This is a special case of [Hay18 Theorem 3.8].)

By [Hay18 Theorem 3.8 and Proposition 3.2], we can replace the normalizer in Property 9 with various other weakenings of the normalizer. For example, this works for the quasi-normalizer as defined in [LP98, PP86, Pop99] (ideas behind the quasi-normalizer trace back to [Pop83]), qG-normalizer as defined in [Pop06a, Pop06b, LP08, GP17]. In fact there is a single subspace of $L^2(M)$ called the “singular subspace” so that the conclusion of 9 holds with the normalizer replaced by the singular subspace, and so that the singular subspace contains all previously defined “weakenings” of the normalizer. We refer the reader to [Hay18] (in particular Theorem 3.8 and Proposition 3.2 of that paper) for a more detailed discussion.

We briefly give some examples of tracial von Neumann algebras with nonpositive 1-bounded entropy (in particular 1-bounded entropy zero if they satisfy the Connes embedding conjecture). The main starting point is Property 9.

Example 1. If $M$ has a Cartan subalgebra, then $h(M) \leq 0$. This follows from Properties 3 and 9. If we analyze the proofs of these properties, then the proof of the fact that $M$ having a Cartan implies $h(M) \leq 0$ is not substantially different than the proof of absence of Cartan for $L(F_n)$, $n > 1$ in [Voi96].

Example 2. If $M$ is non-prime, then $h(M) \leq 0$. To see this, suppose that $M = M_1 \overline{\otimes} M_2$, where $M_1, M_2$ are diffuse. Fix $A_j \leq M_j$ diffuse and abelian. Then $h(M_1 \overline{\otimes} A_2 : M), h(A_1 \overline{\otimes} M_2 : M) \leq 0$ by Properties 9, 2 and 2. Thus by Property 9

$$h(M) = h(M : M) \leq h(M_1 \overline{\otimes} A_2 : M) + h(A_1 \overline{\otimes} M_2 : M) \leq 0.$$ 

This is similar to Ge’s proof of primeness of free group factors in [Ge98].

Example 3. If every non-zero direct summand of $M$ has property Gamma, then $h(M) \leq 0$. To see this, note that by [Dix69 Proposition 1.10] having property Gamma for every non-zero direct summand implies that we can find a sufficiently large infinite set $I$, and a free ultrafilter $\omega$ on $I$, so that there is a diffuse, abelian $A \leq M' \cap M^{\omega}$. We then have by Properties 8, 2, 9, and 3 that

$$h(M) = h(M : M) = h(M : M^{\omega}) \leq h(W^*(N_{M^{\omega}}(A)) : M^{\omega}) = h(A : M^{\omega}) \leq h(A : A) = h(A) = 0.$$ 

This is similar to Voiculescu’s proof that free group factors do not have Property Gamma in [Voi96]. Note that the same proof shows that if $Q \leq M$ is diffuse and $Q' \cap M^{\omega}$ is diffuse, then $h(Q : M) \leq 0$.

More generally, any von Neumann algebra with Property (C') in the sense of Galatan-Popa [GP17 Definition 3.6] has 1-bounded entropy zero (see [Hay18 Corollary 4.8]). There are even examples of von Neumann algebras which have the Haagerup property, are strongly solid, and have the complete metric approximation property which have nonpositive 1-bounded entropy (see [Hou13a, HST11], as well as [Hay18 Corollary 4.13]). We remark that it is not known if $h(M) < \infty$ implies $h(M) \leq 0$. Thus there is a priori a gap between strongly 1-bounded algebras and algebras with nonpositive 1-bounded entropy. Another important consequence of the general properties of 1-bounded entropy is the existence of Pinsker algebras. Indeed, by Properties 9 and 7 if $N \leq M$ is diffuse and has $h(N : M) = 0$, then there is a unique Pinsker algebra $P \leq M$ with $N \subseteq P$. Moreover, as consequence of Property 9 if $P_1, P_2$ are two Pinsker algebras and $P_1 \cap P_2$ is diffuse, then $P = P_2$.

The motivation and terminology for Pinsker algebras comes from ergodic theory. In the context of dynamical entropy of probability measure-preserving actions of groups it is well known that there is a maximal factor action which has nonpositive entropy, called the Pinsker factor, and that the original action is a complete positive entropy extension of the Pinsker factor action. For the case of $\mathbb{Z}$, this was first established by Pinsker in [Pin60]. For sofic entropy this is a folklore result, and for Rokhlin entropy this is due to Seward [Sewb]. We should note here that the Pinsker algebra as we have defined it is really more analogous to the outer Pinsker factor that appears in sofic (implicit in [Ker13] and explicitly defined in [Hay17]) and Rokhlin entropy (see [Sewb]). It appears that outer Pinsker factors have better permanence properties than Pinsker factors (see e.g., [Hay17, Hay, Sewb, Sewa]), and we expect the same to hold in the 1-bounded entropy setting. For example, a probability measure-preserving action always has Lebesgue spectrum over its outer Pinsker factor (see [Hay17, Sewa]). In the context of 1-bounded entropy, the analogous result is if $P \leq M$ is Pinsker, then $P \leq M$ is coarse in the sense of Popa [Pop]. This is merely a rephrasing of [Hay18 Theorem 3.8].
1.3. Concentration of Measure. As in Theorem [3] we will consider a random matrix tuple $X^{(k)} \in M_{n(k)}(\mathbb{C})^I$ for a possibly infinite index set $I$. The distribution of such a random variable is a Borel probability measure $\mu^{(k)}$ on the product space $M_{n(k)}(\mathbb{C})^I$, endowed with the product topology coming from the usual topology on $M_{n(k)}(\mathbb{C})$.

**Definition 1.3.** Given a probability measure $\mu$ on $M_n(\mathbb{C})^I$, a finite $F \subseteq I$, and $\varepsilon > 0$, we define the concentration function of $\mu$ by

$$\alpha_\mu(F, \varepsilon) = \sup\{\mu(N_{F,\varepsilon}(\Omega)) : \Omega \subseteq M_n(\mathbb{C})^I \text{ Borel}, \mu(\Omega) \geq 1/2\},$$

where $N_{F,\varepsilon}$ is an $(F, \varepsilon)$-neighborhood.

We have slightly modified the usual definition of the concentration function for a probability measure on a metric space. The concentration $\alpha_\mu(F, \varepsilon)$ used here is simply the metric concentration function for the marginal on the coordinates indexed by $F$ (where the metric is given by the maximum of the 2-norms of the coordinates). Alternatively, $\alpha_\mu(F, \varepsilon)$ can be viewed as the concentration function of a uniform structure (rather than a metric) on $M_n(\mathbb{C})^I$, namely the uniform structure given by the neighborhoods $N_{F,\varepsilon}(A)$, which is the product of the uniform structures on $M_n(\mathbb{C})$ given by the $\|\cdot\|_2$ metric.

**Definition 1.4.** Given a sequence $n(k) \to \infty$ and a sequence of probability measures $\mu^{(k)}$ on $M_{n(k)}(\mathbb{C})^I$, we say that $\mu^{(k)}$ has exponential concentration if for every finite $F \subseteq I$ and every $\varepsilon > 0$,

$$\limsup_{k \to \infty} \frac{1}{n(k)^2} \log \alpha_{\mu^{(k)}}(F, \varepsilon) < 0.$$

Intuitively, exponential concentration of measure says that if a sequence of Borel sets $\Omega_k \subseteq M_{n(k)}(\mathbb{C})^I$ have measure at least $1/2$, then their $(F, \varepsilon)$-neighborhoods will include everything except for a set of exponentially small measure as $k \to \infty$. As we will explain later, this concentration phenomenon occurs in many natural examples from random matrix theory. Another consequence of concentration is that if $\Omega_k$ is not exponentially small, then the complement of the $(F, 2\varepsilon)$ neighborhood of $\Omega_k$ must be exponentially small, which follows from the next lemma.

**Lemma 1.5.** Let $\mu$ be a probability measure on $M_n(\mathbb{C})^I$. Then for every Borel set $\Omega$, every finite $F \subseteq I$, and every $\varepsilon > 0$, we have

$$\mu(\Omega) > \alpha_\mu(F, \varepsilon) \implies \mu(N_{F,2\varepsilon}(\Omega)) \geq 1 - \alpha_\mu(F, \varepsilon).$$

**Proof.** Suppose $\mu(\Omega) > \alpha_\mu(\varepsilon)$. Let $\Upsilon = N_{F,\varepsilon}(\Omega)^c$. Note that $\Omega \subseteq N_{F,\varepsilon}(\Upsilon)^c$. Thus, $\mu(N_{F,\varepsilon}(\Upsilon)^c) > \alpha_\mu(F, \varepsilon)$, which implies that $\mu(\Upsilon) < 1/2$, by applying the definition of $\alpha_\mu(F, \varepsilon)$ to the set $\Omega$ in contrapositive. Thus, $\mu(N_{F,\varepsilon}(\Omega)) = \mu(\Upsilon^c) \geq 1/2$. So by the definition of the concentration function, we have $\mu(N_{F,2\varepsilon}(\Omega)^c) \leq \alpha_\mu(F, \varepsilon)$. Hence, $\mu(N_{F,2\varepsilon}(\Omega)) \geq 1 - \alpha_\mu(F, \varepsilon)$. \qed

2. An $L^2$-Continuous Functional Calculus

Following the ideas of the second author from [Jek15] and [Jek19], we define a notion of “functions of several non-commuting real variables” that is more flexible than the notions of non-commutative polynomials. The space $\mathcal{F}_{R,\infty}$ which we define has the following useful properties:

- Given $(M, \tau)$ and $x \in M_I^I$ with $\|x_i\| \leq R_i$, every element of $W^*(x)$ can be realized as $f(x)$ for some $f \in \mathcal{F}_{R,\infty}$ (see Proposition 2.4).
- Given a tuple $f \in \mathcal{F}_{R,\infty}$, the law of $f(x)$ depends continuously on the law of $x$ (see Proposition 2.6), which will mean that $f$ will map microstate spaces for $x$ into microstate spaces of $f(x)$ (see Corollary 2.7).
- Given $f \in \mathcal{F}_{R,\infty}$, the evaluation $x \mapsto f(x)$ is $L^2$-uniformly continuous in a certain sense (see Proposition 2.8), which will be helpful later for “pushing forward” concentration of measure (see Corollary 2.11).

Here we approach the definitions from a slightly different point of view than [Jek19]. Moreover, we generalize to the setting of infinite tuples without assuming embeddability into $R^\omega$. However, we make the simplifying restriction of only handling functions that are defined in a product of operator norm balls rather than on arbitrary self-adjoint tuples. As explained in [Jek19] §3.1, the space $\mathcal{F}_{R,\infty}$ is closely related to the trace polynomials used in previous work such as [Rai97] [Ceb13] [DHK13] [Kem16] [Kem17] [DGS16].
Moreover, from the model-theoretic viewpoint of [FHS13, FHS14a, FHS14b], the functions in $F_{R,\infty}$ are certain quantifier-free definable functions in the language of tracial von Neumann algebras; see [Jek19, Remark 3.3].

2.1. Construction of the Space.

**Definition 2.1.** Fix an index set $I$ and $R \in (0, +\infty)^I$. Consider the space

$$A_R = C(\Sigma_R) \otimes C(t_i : i \in I).$$

Given $(M, \tau)$ and $x \in M^I_{sa}$ with $\|x_i\| \leq R_i$, we define the evaluation map

$$\text{ev}_x : A_R \to M$$

$$\phi \otimes p \mapsto \phi(\lambda_x)p(x).$$

Then we define a semi-norm on $C(\Sigma_R) \otimes C(t_i : i \in I)$ by

$$\|f\|_{R,2} = \sup_{(M, \tau), x} \|\text{ev}_x(f)\|_{L^2(M)},$$

where the supremum is over all tracial $W^*$-algebras $(M, \tau)$ and all $x \in M^I_{sa}$ with $\|x_i\| \leq R_i$. Denote by $F_{R,2}$ the completion of $A_R/\{f \in A_R : \|f\|_{R,2} = 0\}$.

It is immediate that for every $(M, \tau)$, for every self-adjoint tuple $x \in M^I_{sa}$ with $\|x_i\| \leq R_i$, the evaluation map $\text{ev}_x : A_R \to M$ passes to a well-defined map $F_{R,2} \to L^2(M)$, which we continue to denote by $\text{ev}_x$, and we will also write $f(x) = \text{ev}_x(f)$. Moreover, it is clear that $f(x) := \text{ev}_x(f)$ always lies in $L^2(W^*(x))$ because this holds when $f$ is a simple tensor.

It will be convenient often to restrict our attention to elements of $F_{R,2}$ that are bounded in operator norm, and we will show that these in fact form a $C^*$-algebra.

**Definition 2.2.** For $f \in F_{R,2}$, let us define

$$\|f\|_{R,\infty} = \sup_{(M, \tau), x} \|\text{ev}_x(f)\|,$$

and then set

$$F_{R,\infty} = \{ f \in F_{R,2} : \|f\|_{R,\infty} < +\infty \}.$$

**Lemma 2.3.** $F_{R,\infty}$ is a $C^*$-algebra with respect the norm $\|\cdot\|_{R,\infty}$ and the multiplication and $*$-operation arising from the natural ones on simple tensors.

**Proof.** We can define a multiplication map $A_R \times A_R \to A_R$ by

$$(\phi \otimes p, \psi \otimes q) \mapsto \phi \psi \otimes pq,$$

and a $*$-operation $(\phi \otimes p)^* = \overline{\phi} \otimes p^*$. Then the evaluation maps $\text{ev}_x$ are $*$-homomorphisms.

Note that for $f \in A_R$, we have $\|f\|_{R,\infty} < +\infty$. Moreover, $f, g \in A_R$,

$$\|fg\|_{R,2} \leq \|f\|_{R,\infty}\|g\|_{R,2},$$

because $\|f(x)g(x)\|_2 \leq \|f(x)\|_\infty\|g(x)\|_2$ for every $x$ coming from a tracial von Neumann algebra. Therefore, the multiplication passes to a well-defined map

$$A_R \times F_{R,2} \to F_{R,2}.$$

On the other hand, given $f \in A_R$ and $g \in F_{R,\infty}$, we can check that

$$\|fg\|_{R,2} \leq \|f\|_{R,\infty}\|g\|_{R,\infty},$$

so that multiplication is well-defined $F_{R,2} \times F_{R,\infty} \to F_{R,2}$. Then by checking that $\|fg\|_{R,\infty} \leq \|f\|_{R,\infty}\|g\|_{R,\infty}$, we see that $F_{R,\infty}$ has a well-defined multiplication operation.

This multiplication operation is characterized by the fact that for self-adjoint tuples $x$ from $(M, \tau)$ with $\|x_i\| \leq R_i$, we have $(fg)(x) = f(x)g(x)$, since $fg$ is uniquely determined by its evaluation on self-adjoint tuples. This easily implies associativity of multiplication, compatibility with the $*$-operation, and the $C^*$-identity for the norm. Completeness of $F_{R,\infty}$ is also a standard exercise (every Cauchy sequence in $\|\cdot\|_{R,\infty}$ would also be Cauchy in $\|\cdot\|_{R,2}$, and so forth).
Proposition 2.6. Given $(\mathcal{M}, \tau)$ and $x \in \mathcal{M}_{sa}^I$ with $\|x_i\| \leq R_i$, the evaluation map $ev_x : \mathcal{F}_{R,2} \to L^2(W^*(x))$ is surjective, and it restricts to a surjective $*$-homomorphism $\mathcal{F}_{R,\infty} \to W^*(x)$. 

Proof. Let $x$ be a self-adjoint tuple from $(\mathcal{M}, \tau)$. Let $z \in L^2(W^*(x))$. Then there is a sequence of non-commutative polynomials $\{p_k\}_{k \in \mathbb{N}}$ such that $p_k(x) \to z$ in $L^2(W^*(x))$. By passing to a subsequence, we can assume without loss of generality that $\|p_{k+1}(x) - p_k(x)\|_2 < 1/2^k$ for $k \geq 1$. Now define $\mathcal{U}_k \subseteq \Sigma_R$ by

$$\mathcal{U}_k = \{ \lambda : \lambda((p_{k+1} - p_k)^2) < 1/2^k \}.$$  

Then $\mathcal{U}_k$ is an open subset of $\Sigma_R$ containing $\lambda_x$. By Urysohn’s lemma, there exists $\phi_k \in C(\Sigma_R)$ such that

$$0 \leq \phi_k \leq 1, \quad \phi_k(x) = 1, \quad \phi_k|_{\mathcal{U}_k} = 0.$$  

This implies that

$$\|\phi_k \otimes (p_{k+1} - p_k)\|_{\mathcal{F}_{R,2}} \leq \frac{1}{2^k}.$$  

Therefore,

$$1 \otimes p_1 + \sum_{k=1}^{\infty} \phi_k \otimes (p_{k+1} - p_k)$$

converges absolutely in $\mathcal{F}_{R,2}$ to some function $f$. By construction,

$$f(x) = p_1(x) + \sum_{k=1}^{\infty} (p_{k+1}(x) - p_k(x)) = z.$$  

Now we turn to the case of $ev_x : \mathcal{F}_{R,\infty} \to W^*(x)$, which we already showed is a $*$-homomorphism. Fix $z \in W^*(x)$. We can assume without loss of generality that $z$ is self-adjoint. Choose a sequence of non-commutative polynomials $\{p_k\}$ and continuous functions $\phi_k$ as above. Assume without loss of generality that $p_k = p_k^*$ and $\phi_k$ is real. Then let

$$f_n = 1 \otimes p_1 + \sum_{k=1}^{n} \phi_k \otimes (p_{k+1} - p_k),$$

which is a self-adjoint element of $\mathcal{F}_{R,\infty}$.

Choose $h \in C_0(\mathbb{R})$ satisfying $h(t) = t$ for $|t| \leq ||z||$ and $\|h\|_{C_0(\mathbb{R})} \leq ||z||$. Since $\mathcal{F}_{R,\infty}$ is a C*-algebra, $h(f_n)$ is well defined and $(h(f_n))(x) = h(f_n(x))$. We claim that $h(f_n)$ converges in $\|\cdot\|_{\mathcal{F}_{R,2}}$ to some $g \in \mathcal{F}_{R,\infty}$ and that $g(x) = z$.

We will use the observation that every $h \in C_0(\mathbb{R})$ is uniformly continuous with respect to $\|\cdot\|_2$ in the following sense: For every $\varepsilon > 0$ there exists $\delta > 0$ such that if $(\mathcal{M}, \tau)$ is a tracial von Neumann algebra and $a, b$ are self-adjoint operators in $\mathcal{M}$, then $\|a - b\|_2 < \delta$ implies $\|h(a) - h(b)\|_2 < \varepsilon$. Clearly, this holds if $h(t)$ is the resolvent $(t + i)^{-1}$, or if $h$ is a $*$-polynomial in $(t + i)^{-1}$. But $*$-polynomials in $(t + i)^{-1}$ are dense in $C_0(\mathbb{R})$ by the Stone-Weierstrass theorem, so the claim holds for all $h \in C_0(\mathbb{R})$ and hence for our particularly chosen $h$.

To show that $\{h(f_n)\}$ is Cauchy in $\|\cdot\|_{\mathcal{F}_{R,2}}$, fix $\varepsilon > 0$. By the $L^2$-uniform continuity of $h$, we may choose $\delta > 0$ such that $\|a - b\|_2 < \delta$ implies $\|h(a) - h(b)\|_2 < \varepsilon$ where $a, b$ as above. Since $\{f_n\}$ is Cauchy in $\|\cdot\|_{\mathcal{F}_{R,2}}$, we have $\|f_n - f_m\|_{\mathcal{F}_{R,2}} < \delta$. So for every self-adjoint tuple $y \in (\mathcal{N}, \tau)^{I_{sa}}$ with $\|y_i\| \leq R_i$, we have $\|f_n(y) - f_m(y)\|_2 < \delta$, so $\|h(f_n(y)) - h(f_m(y))\|_2 < \varepsilon$, thus making $\|h(f_n) - h(f_m)\|_{\mathcal{F}_{R,2}} \leq \varepsilon$.

So $\{h(f_n)\}$ converges in $\|\cdot\|_{\mathcal{F}_{R,2}}$ to some $g$. Since $\|h(f_n)\|_{\mathcal{F}_{R,\infty}} \leq ||z||$ for all $n$, we have $\|g\|_{\mathcal{F}_{R,\infty}} \leq ||z||$. Finally, since $f_n(x) \to z$ in $\|\cdot\|_2$, the $L^2$-uniform continuity of $h$ implies that $h(f_n(x)) \to h(z) = z$ in $\|\cdot\|_2$, and therefore $g(x) = z$ as desired. \hfill $\square$

2.2. Push-Forwards of Non-commutative Laws. Now we turn our attention to the way that tuples from $\mathcal{F}_{R,\infty}$ push forward non-commutative laws.

Definition 2.5. Let $I$ and $I'$ be index sets. Let $R \in (0, +\infty)^I$ and $R' \in (0, +\infty)^{I'}$. We define

$$\mathcal{F}_{R,R'} = \{ f = (f_i)_{i \in I'} \in \mathcal{F}_{R,\infty}^{I'} : \|f_i\|_{\mathcal{F}_{R,\infty}} \leq R'_i \text{ for all } i \in I' \}.$$  

Proposition 2.6. Let $R \in (0, +\infty)^I$ and $R' \in (0, +\infty)^{I'}$. Let $f = (f_i)_{i \in I'} \in \mathcal{F}_{R,R'}$. 

(1) Given \((\mathcal{M}, \tau)\) and \(x \in \mathcal{M}_a^i\) with \(\|x_i\| \leq R_i\), we set \(f(x) = (f_i(x))_{i \in I'}\). Then \(\lambda_{f(x)}\) is uniquely determined by \(\lambda_X\).
(2) Let \(f, \) be the “push-forward” mapping \(\Sigma_R \to \Sigma_{R'}\) defined by \(f, \lambda_x = \lambda_{f(x)}\) for all such tuples \(x\). Then \(f, \) is continuous.

Proof. It suffices to show that for every non-commutative polynomial \(p \in \mathbb{C}(t_i : i \in I')\), the quantity \(\tau(p(f(x)))\) is uniquely determined by \(\lambda_X\), and that it depends continuously on \(\lambda_X\). Now since \(\mathcal{F}_{R,\infty}\) is a C*-algebra, \(f := p(f)\) is an element of \(\mathcal{F}_{R,\infty} \subseteq \mathcal{F}_{R,2}\). Thus, it suffices to show that for \(f \in \mathcal{F}_{R,2}\), the quantity \(\tau(f(x))\) is uniquely determined by \(\lambda_X\) and depends continuously on it. Now \(|\tau(f(x))| \leq \|f(x)\|_2 \leq \|f\|_{R,2}\), so we can reduce to the case where \(f\) comes from a dense subset, say \(\mathcal{A}_R\) (or rather its image under the quotient map). Then by linearity, we reduce to the case where \(f\) is given by a simple tensor \(\phi \otimes p\). But in this case,

\[
\tau(f(x)) = \phi(\lambda_X)\tau(p(x)) = \phi(\lambda_X)\lambda_X(p),
\]

which only depends on \(\lambda_X\), and which is given by the continuous function \(\lambda \mapsto \phi(\lambda)\lambda(p)\) on \(\Sigma_R\).

The elements of \(\mathcal{F}_{R,R'}\) can be used to map between microstate spaces in the following way. Let \(f \in \mathcal{F}_{R,R'}\). If \(A \in M_n(\mathbb{C})^a_i\) satisfies \(\|A_i\| \leq R_i\), then \(f(A)\) is a well-defined element of \(M_n(\mathbb{C})^a_i\) (since \(M_n(\mathbb{C})\) is a tracial W*-algebra).

Corollary 2.7. Let \(R \in (0, +\infty)^I\), \(R' \in (0, +\infty)^{I'}\), and \(f \in \mathcal{F}_{R,R'}\). Let \(x\) be a self-adjoint tuple from \((\mathcal{M}, \tau)\) with \(\|x_i\| \leq R_i\).

1. For every neighborhood \(V\) of \(\lambda_{(x,f(x))}\) in \(\Sigma_{(R,R')}\), there exists a neighborhood \(U\) of \(\lambda_X\) such that \(f(\Gamma_{(R,R'),n}(x;U)) \subseteq \Gamma_{(R,R'),n}(f(x);V)\) for all \(n\).

2. Similarly, fix \(R_0\), a \(z \in \mathcal{M}_a\) with diffuse spectrum and microstates \((C(n))_n\) for \(z\) and \(f\). Then for every neighborhood \(V\) of \(\lambda_{(x,f(x),z)}\), there exists a neighborhood \(U\) of \(\lambda_{(x,z)}\) such that \(f(\Gamma_{(R,R'),n}(x;C(n) \to z;U)) \subseteq \Gamma_{(R,R'),n}(f(x);C(n) \to z;V)\) for all \(n\).

Proof. (1) It follows from the previous result that the non-commutative law of \((a, f(a))\) depends continuously on the non-commutative law of \(a\). This means that \(U = (id, f)^{-1}(V)\) is open in \(\Sigma_R\).

(2) The argument is similar, using the fact that the law of \((a, f(a), c)\) depends continuously on the law of \((a, c)\).

2.3. \(L^2\)-uniform Continuity. The elements of \(\mathcal{F}_{R,2}\) are all “\(L^2\)-uniformly continuous functions” in the following sense.

Proposition 2.8. Let \(I\) be an index set and \(R \in (0, +\infty)^{I}\), and let \(f \in \mathcal{F}_{R,2}\). Then for every \(\epsilon > 0\) there exists a finite \(F \subseteq I\) and a \(\delta > 0\) such that for every \((\mathcal{M}, \tau)\) and \(x, y \in \mathcal{M}_a^i\) with \(\|x_i\|, \|y_i\| \leq R_i\), if \(\|x_i - y_i\|_2 < \delta\) for all \(i \in F\), then \(\|f(x) - f(y)\|_2 < \epsilon\).

Proof. First, suppose that \(f\) has the form \(1 \otimes p\) where \(p\) is a non-commutative polynomial (or in other words, \(f(x) = p(x)\)). Then it is a straightforward exercise to check this uniform continuity property (for instance, by handling each monomial explicitly).

Second, consider the case where \(f\) has the form \(\phi \otimes 1\) for some \(\phi \in C(\Sigma_R)\). Let \(A\) be the set of all \(\phi \in C(\Sigma_R)\) such that \(\phi \otimes 1\) satisfies the above uniform continuity property. One checks easily that \(A\) is a \(*\)-subalgebra of \(C(\Sigma_R)\). Moreover, since \(\|\phi \otimes 1\|_{R,2}\) is simply \(\|\phi\|_{C(\Sigma_R)}\), we see that \(A\) is closed in \(C(\Sigma_R)\).

In light the first case, \(A\) contains every function of the form \(\phi(\lambda_X) = \tau(p(x))\) for \(p \in \mathbb{C}(t_i : i \in I)\). This means that \(A\) separates points in \(\Sigma_R\), by definition of \(\Sigma_R\). Therefore, by the Stone-Weierstrass Theorem (since \(A\) contains 1), we have \(A = C(\Sigma_R)\). Thus, we have the desired continuity property for \(\phi \otimes 1\).

Finally, we check that every function of the form \(\phi \otimes p\) in \(\mathcal{F}_{R,2}\) satisfies the desired continuity property by using Cases 1 and 2 and the inequality

\[
\|\phi(\lambda_X)p(x) - \phi(\lambda_Y)p(y)\|_2 \leq \|\phi\|_{C(\Sigma_R)}\|p(x) - p(y)\|_2 + \|\phi(\lambda_X) - \phi(\lambda_Y)\||1 \otimes p||_{R,2}.
\]

Linear combinations of these functions are dense by definition of \(\mathcal{F}_{R,2}\), and hence the proof is complete. \(\square\)
The $L^2$-uniform continuity has several consequences that we will use in our handling of microstate spaces. The first is well known to experts. We will deduce it from the previous lemma, although it is also easy to prove directly.

**Lemma 2.9.** Let $I$ be an index set and $R \in (0, +\infty)^I$. Let $\mathcal{U} \subseteq \Sigma_R$ and $\mu \in \mathcal{U}$ be open. Then there exists an open $V \ni \mu$, a finite $F \subseteq I$, and an $\delta > 0$ such that for any tracial von Neumann algebra $(\mathcal{M}, \tau)$ and $\lambda \in \mathcal{M}_{sa}$ with $\|x_i\|, \|y_i\| \leq R_i$, if $\lambda_x \in V$ and $\|x_i - y_i\|_2 < \delta$ then $\lambda_y \in V$.

**Proof.** By Urysohn’s lemma, there exists $\phi \in C(\Sigma_R)$ such that $\phi(\mu) = 1$ and $\phi$ is supported in $\mathcal{U}$. Let $V = \{\lambda : \phi(\lambda) > 1/2\}$. By the previous lemma, there exist $F \subseteq I$ finite and $\delta > 0$ such that for any $(\mathcal{M}, \tau)$ and $x, y \in \mathcal{M}_{sa}$ with $\|x_i\|, \|y_i\| \leq R_i$, if $\|x_i - y_i\|_2 < \delta$ for $i \in I$ implies that $|\phi(\lambda_x) - \phi(\lambda_y)| < 1/2$. This choice of $V$, $F$, and $\delta$ works. \qed

The analogous statement below for sequences of laws follows immediately.

**Corollary 2.10.** Given $I$ and $R \in (0, +\infty)^I$. Let $(\mathcal{M}(k), \tau'(k))$ for $k \in \mathbb{N}$ be a sequence of tracial von Neumann algebras, and $x^{(k)}, y^{(k)} \in (\mathcal{M}(k))_{sa}$ supported on matrix tuples with $\|x_i^{(k)}\|, \|y_i^{(k)}\| \leq R$. If $\lambda^{(k)} \rightarrow \lambda$ in $\Sigma_R$ and if $\|x_i^{(k)} - y_i^{(k)}\|_2 \rightarrow 0$ as $k \rightarrow \infty$ for each $i \in I$, then $\lambda_y \rightarrow \lambda$.

Another consequence of the $L^2$-uniform continuity is that exponential concentration of measure is preserved when we push forward a sequence of probability measures on matrix tuples through a function $f \in \mathcal{F}_{R, R'}$. The push-forward here is technically different from that of Proposition 2.6. If $\mu(k)$ is a probability measure on $M_{n(k)}(\mathbb{C})_{sa}$ supported on matrix tuples $A$ with $\|A_i\| \leq R_i$, then $f_*\mu(k)$ is defined by using the evaluation of $f$ on tuples $A$ of $n(k) \times n(k)$ self-adjoint matrices satisfying the given operator norm bounds. The following result on preservation of exponential concentration will of course be used later for the measures coming from the random matrix models in Theorem A.9.

**Corollary 2.11.** Let $R \in (0, +\infty)^I$ and $R' \in (0, +\infty)^{I'}$. Let $\mu(k)$ be a probability measure on $M_{n(k)}(\mathbb{C})_{sa}$ supported on $A_i : \|A_i\| \leq R_i$, and let $f \in \mathcal{F}_{R, R'}$. Moreover, if $(\mu(k))_{k \in \mathbb{N}}$ has exponential concentration, then so does $(f_*\mu(k))_{k \in \mathbb{N}}$.

**Proof.** Fix a finite $F' \subseteq I$ and $\varepsilon > 0$. Applying the previous proposition to each $f_i$ for $i \in F'$, we see that there exists $F \subseteq I$ and $\delta > 0$ such that for tuples of matrices $A$ and $B$ with $\|A_i\| \leq R_i$ and $\|B_i\| \leq R_i$, we have $\|A_i - B_i\|_2 < \delta$ for $i \in F \implies \|f_i(A) - f_i(B)\|_2 < \varepsilon$ for $i \in F'$. Therefore, if $\Omega \subseteq M_{n(k)}(\mathbb{C})_{sa}$ with $f_*(\mu(k))(\Omega) = \mu(k)(f^{-1}(\Omega)) \geq 1/2$, then $N_{F, \delta}(f^{-1}(\Omega)) \subseteq f^{-1}(N_{F', \varepsilon}(\Omega))$, and consequently $\mu(k)(N_{F, \delta}(f^{-1}(\Omega))^c) \geq (f_*\mu(k))(N_{F', \varepsilon}(\Omega)^c)$. Since this holds for all Borel $\Omega$, we get $\alpha_{f_*\mu(k)}(F', \varepsilon) \leq \alpha_{\mu(k)}(F, \delta)$, and hence

$$\limsup_{k \to \infty} \frac{1}{n(k)^2} \log \alpha_{f_*\mu(k)}(F', \varepsilon) \leq \limsup_{k \to \infty} \frac{1}{n(k)^2} \log \alpha_{\mu(k)}(F, \delta) < 0.$$ \qed

**Remark 2.12.** The space $\mathcal{F}_{R, R'}$ can be used to slightly simplify the proof that 1-bounded entropy is a $W^*$-algebra invariant [Hay18, Theorem A.9]. By a variant of Corollary 2.7, we can arrange a function $f$ that maps one microstate space to the other, and the $L^2$-uniform continuity of $f$ allows us to push forward a $\delta$-dense subset of the first microstate space to an $\varepsilon$-dense subset of the second one.

### 3. Proof of Theorem A.9

3.1. **Setup.** We remind the reader of the setup of Theorem A.9. Let $(\mathcal{M}, \tau)$ be a tracial $W^*$-algebra and $x^{(k)} \in \mathcal{M}_{sa}^I$ be a set of self-adjoint generators indexed by $I$, and let $R \in (0, +\infty)^I$ with $\|x_i\| < R_i$. Suppose that $\mathcal{P}$ is a $W^*$-subalgebra of $\mathcal{M}$. We assume $n(k) \to \infty$ and that $X^{(k)}$ is an $I$-tuple of random $n(k) \times n(k)$ matrices, satisfying (1) - (4) below. For the reader’s convenience, we restate these conditions also in terms of the probability distribution $\mu(k)$ of $X^{(k)}$ (which is a probability measure on $M_{n(k)}(\mathbb{C})_{sa}$).

1. $\|X_i^{(k)}\|_\infty \leq R_i$, or equivalently $\mu(k)$ is supported on $\{A \in M_{n(k)}(\mathbb{C})_{sa} : \|A\| \leq R_i\}$. 


(2) \( \tau_{n(k)}(p(X^{(k)})) \to \tau(p(x)) \) in probability for every non-commutative polynomial \( p \). Equivalently, \( \mu^{(k)} \) is asymptotically supported on the microstate spaces for \( x \), meaning that

\[
\mu^{(k)}(U_{R,n(k)}(x;U)) \to 1
\]

for every neighborhood \( U \) of \( \lambda_x \) in \( \Sigma_R \).

(3) The measures \( \mu^{(k)} \) exhibit exponential concentration in the sense of Definition \ref{dfn:C}. This is equivalent to the conclusion of the theorem. Let \( f \in C(t_i : i \in I) \), and note that

\[
\lim_{k \to \infty} \| \mathbb{E}[p(X^{(k)})] \|_2 = \lim_{k \to \infty} \left\| \int p(A) \, d\mu^{(k)}(A) \right\|_2 = \| \mathbb{E}_P[p(x)] \|_2.
\]

We refer to this condition as the external averaging property.

The conclusion to the theorem is that if \( N \leq M \) such that \( N \cap P \) is diffuse and \( h(N : M) = 0 \), then \( N \subseteq P \).

### 3.2. The External Averaging Property

As the first ingredient in the proof, we provide several equivalent interpretations of the external averaging property.

**Lemma 3.1.** Let \( P \leq M = W^*(x) \), let \( R \in (0, +\infty)^I \) with \( \| x_i \| \leq R_i \), and let \( \mu^{(k)} \) be a sequence of probability measures on \( M_n(k)(\mathbb{C})_{t_i}^I \) satisfying conditions (1) and (2) of \ref{dfn:C}. Then the following are equivalent:

1. For every non-commutative polynomial \( q \in C(t_i : i \in I) \),
   
   \[
   \lim_{k \to \infty} \left\| \int q \, d\mu^{(k)} \right\|_2 = \| \mathbb{E}_P(q(x)) \|_2.
   \]

2. For every non-commutative polynomial \( q \),

   \[
   \lim_{k \to \infty} \int \left\| q(A) - \int q \, d\mu^{(k)} \right\|_2^2 \, d\mu^{(k)}(A) = \| q(x) - \mathbb{E}_P(q(x)) \|_2^2.
   \]

3. For every \( f \in \mathcal{F}_R, \) we have

   \[
   \lim_{k \to \infty} \left\| \int f \, d\mu^{(k)} \right\|_2 = \| \mathbb{E}_P(f(x)) \|_2.
   \]

4. For every \( f \in \mathcal{F}_R, \) we have

   \[
   \lim_{k \to \infty} \int \left\| f(A) - \int f \, d\mu^{(k)} \right\|_2^2 \, d\mu^{(k)}(A) = \| f(x) - \mathbb{E}_P(f(x)) \|_2^2.
   \]

**Proof.** First, we show that (1) \( \iff \) (2) and (3) \( \iff \) (4). It suffices to show that for each individual \( f \in \mathcal{F}_R, \) we have

\[
\lim_{k \to \infty} \left\| \int f \, d\mu^{(k)} \right\|_2 = \| \mathbb{E}_P(f(x)) \|_2 \iff \lim_{k \to \infty} \int \left\| f(A) - \int f \, d\mu^{(k)} \right\|_2^2 \, d\mu^{(k)}(A) = \| f(x) - \mathbb{E}_P(f(x)) \|_2^2.
\]

This is because of course a non-commutative polynomial \( q \) is also in \( \mathcal{F}_R \).

So suppose that \( f \in \mathcal{F}_R, \) and note that \( \| f(x) \|_2 \) is uniformly bounded when \( \| x_i \| \leq R_i \), so that the integrals and conditional expectations above are well-defined. Now we observe that

\[
\| f(x) - \mathbb{E}_P(f(x)) \|_2^2 = \| f(x) \|_2^2 - \| \mathbb{E}_P(f(x)) \|_2^2
\]

and

\[
\int \left\| f(A) - \int f \, d\mu^{(k)} \right\|_2^2 \, d\mu^{(k)}(A) = \int \| f \|_2^2 \, d\mu^{(k)} - \int \left\| \int f \, d\mu^{(k)} \right\|_2^2.
\]

Therefore, the claim \ref{dfn:C} reduces to showing that

\[
\lim_{k \to \infty} \int \| f \|_2^2 \, d\mu^{(k)} = \| f(x) \|_2^2.
\]

Now \( \| f(y) \|_2^2 = \tau(f(y)^2) \) depends continuously on the law of \( f(y) \) and hence depends continuously on the law of \( y \) by Proposition \ref{prop:law}. We assumed that the measures \( \mu^{(k)} \) are asymptotically supported on microstates of \( x \). In other words, for the associated random matrix tuple \( X^{(k)} \), the non-commutative law \( \lambda_{X^{(k)}} \) converges
in probability to \(\lambda_x\). Hence, \(\|f(X^{(k)})\|_2^2 \rightarrow \|f(x)\|_2^2\) in probability. But it is also uniformly bounded, so convergence in probability implies convergence in expectation.

To complete the proof of the lemma, note that (3) \(\implies\) (1) trivially, so it remain to show (1) \(\implies\) (3). First, suppose that \(f\) is given by a sum of simple tensors, that is, \(f(y) = \sum_{j=1}^N \phi_j(\lambda_y)q_j(y)\) for some non-commutative polynomials \(q_j\) and \(\phi_j \in C(\Sigma_R)\). Define \(\tilde{f}(y) = \sum_{j=1}^N \phi_j(\lambda_x)q_j(y)\), which is a non-commutative polynomial since \(\phi_j\) has been replaced by the constant \(\phi_j(\lambda_x)\). Since \(\lambda_{X^{(k)}} \rightarrow \lambda_x\) in probability, we have

\[
\left\|f(X^{(k)}) - \tilde{f}(X^{(k)})\right\|_2 \rightarrow 0
\]

in probability. But because \(\|f - \tilde{f}\|_2\) is uniformly bounded, it also converges to zero in expectation, and hence

\[
\lim_{k \rightarrow \infty} \mathbb{E}(f(X^{(k)})) = \lim_{k \rightarrow \infty} \mathbb{E}(\tilde{f}(X^{(k)})) = \|E_P(\tilde{f}(x))\|_2 = \|E(f(x))\|_2,
\]

since \(f(x) = \tilde{f}(x)\). Thus, (3) holds when \(f\) is a finite linear combination of simple tensor. But every \(f \in \mathcal{F}_{\mathcal{R},\infty}\) (hence every \(f \in \mathcal{F}_{\mathcal{R},\infty}\)) can be approximated in \(\|\cdot\|_{\mathcal{R},\infty}\) by linear combinations of simple tensors, and thus by a straightforward approximation argument (3) extends to this case as well.

This lemma has the following interpretation in terms of Hilbert space geometry. Let us denote by \(L^2(\mu^{(k)}, M_{n(k)}(\mathbb{C}))\) the \(L^2\)-space of functions \((M_{n(k)}(\mathbb{C}), \mu^{(k)}) \rightarrow (M_{n(k)}(\mathbb{C}), \|\cdot\|_2)\) (random matrices defined by the measure \(\mu^{(k)}\)). The expectation (or integral with respect to \(\mu^{(k)}\)) is the orthogonal projection onto the subspace of deterministic matrices, and of course the \(W^*\)-algebraic conditional expectation \(E_P\) is the orthogonal projection on \(L^2(\mathcal{P})\). Thus, the external averaging property says that for \(f \in \mathcal{F}_{\mathcal{R},\infty}\), the distance of \(f(X^{(k)})\) from deterministic matrices in \(L^2(\mu^{(k)}, M_{n(k)}(\mathbb{C}))\) converges to the distance from \(f(x)\) to \(L^2(\mathcal{P})\) in \(L^2(\mathcal{M})\). Hence, a function \(f(x)\) will be in \(\mathcal{P}\) if and only if \(f(X^{(k)})\) is well-approximated by deterministic matrices in \(L^2(\mu^{(k)}, M_{n(k)}(\mathbb{C}))\), which we state precisely in the next proposition.

**Proposition 3.2.** Let \(\mathcal{P} \subseteq \mathcal{M} = W^*(X)\), and let \(\mu^{(k)}\) be a sequence of measures satisfying (1), (2), and (4) from \(\mathcal{F}\). Let \(z \in W^*(x)_{sa}\) and recall \(z = f(x)\) for some \(f \in (\mathcal{F}_{\mathcal{R},\infty})_{sa}\) by Proposition \(\mathcal{F}\). Then we have \(z \in \mathcal{P}\) if and only if there exists a sequence \((C^{(k)})_{k \in \mathbb{N}}\) of deterministic matrices such that \(\|f - C^{(k)}\|_{L^2(\mu^{(k)}, M_{n(k)}(\mathbb{C}))} \rightarrow 0\). Moreover, in this case, we can choose \((C^{(k)})_{k \in \mathbb{N}}\) to be bounded in operator norm.

**Proof.** Suppose that \(z = f(x) \in \mathcal{P}\). Then we have

\[
\lim_{k \rightarrow \infty} \int \left\|f(A) - \int f \, d\mu^{(k)}(A)\right\|_2^2 \, d\mu^{(k)}(A) = \|f(x) - E_N(f(x))\|_2^2 = 0.
\]

Therefore, we can set \(C^{(k)} = \int f(A) \, d\mu^{(k)}(A)\), which is bounded in operator norm because \(f\) is bounded in operator norm.

Conversely, suppose that such a sequence of matrices \((C^{(k)})_{k \in \mathbb{N}}\) exists. Since the expectation of a random matrix is the orthogonal projection onto the subspace of deterministic matrices, we have

\[
\int \left\|f(A) - \int f \, d\mu^{(k)}(A)\right\|_2^2 \, d\mu^{(k)}(A) \leq \int \left\|f(A) - C^{(k)}\right\|_2^2 \, d\mu^{(k)}(A) \rightarrow 0.
\]

Therefore,

\[
\|f(x) - E_P(f(x))\|_2^2 = \lim_{k \rightarrow \infty} \int \left\|f(A) - \int f \, d\mu^{(k)}(A)\right\|_2^2 \, d\mu^{(k)}(A) = 0.
\]

**3.3. Microstate Collapse.** To motivate our next main ingredient (Proposition \(\mathcal{F}\)), let us sketch our strategy for proving Theorem \(\mathcal{F}\). Consider \(\mathcal{P} \subseteq \mathcal{M} = W^*(x)\) as in the theorem. We need to show that if \(\mathcal{N} \cap \mathcal{P}\) is diffuse and \(h(N : \mathcal{M}) = 0\), then \(\mathcal{N} \subseteq \mathcal{P}\). If \(y\) is a tuple of self-adjoint generators for \(\mathcal{N}\) and \(y = f(x)\) for some \(f \in \mathcal{F}_{\mathcal{R},\mathcal{R}}\), we want to show that each \(y_i\) must be in \(\mathcal{P}\). By Proposition \(\mathcal{F}\), it suffices to show that \(f_1(X^{(k)})\) is well-approximated in \(L^2(\mu^{(k)}, M_{n(k)}(\mathbb{C}))\) by a sequence deterministic matrices \(C_i^{(k)}\). In other words, we want to show that most of the mass of \(f_1\) supported on the microstate space of \(y\) in the presence of \(x\) (or at least those induced from microstates of \(x\)) are
close together. Thus, the second main ingredient in the theorem is to show that exponential concentration of measure for $\mu^{(k)}$ together with $h(\mathcal{N} : \mathcal{M}) = 0$ causes such a “collapse” of the microstate space, which is the conclusion of the following proposition. We state the proposition is a slightly more general setting, since we believe it has independent interest.

Proposition 3.3. Let $(\mathcal{M}, \tau)$ be generated by the self-adjoint tuple $x \in \mathcal{M}_sa'$. Let $y \in \mathcal{M}_sa'$, where $I'$ is an arbitrary index set. Let $z \in \mathcal{M}$ be a self-adjoint operator with diffuse spectrum. Let $\mathcal{N} = W^*(y, z)$ and suppose that $h(\mathcal{N} : \mathcal{M}) = 0$.

Let $R \in (0, +\infty)^I$, $R' \in (0, +\infty)^{I'}$, and $R_0 \in (0, +\infty)$ be such that $\|x_i\| < R_i$ for $i \in I$, $\|y_i\| < R'_i$ for $i \in I'$, and $\|z\| < R_0$. Let $y = f(x)$ for some $f \in \mathcal{F}_{R,R'}$.

Let $(C(k))_{k \in \mathbb{N}}$ be a sequence of microstates for $z$ with $\|C(k)\| < R_0$. Let $n(k) \to \infty$, and let $\mu^{(k)}$ be a probability measure on $M_{n(k)}(\mathbb{C})$ supported on $\{A : \|A\| \leq R_i\}$. Suppose that $\mu^{(k)}$ is asymptotically supported on the microstates of $x$ relative to $C(k) \rightsquigarrow z$; that is,

$$\mu^{(k)}(\Gamma_{(R,R_0),n(k)}(x|C(k) \rightsquigarrow z; U)) \to 1$$

for every neighborhood $U$ of the non-commutative law of $x$. Furthermore, suppose that $\mu^{(k)}$ has exponential concentration.

Then there exist tuples $B^{(k)} \in M_{n(k)}(\mathbb{C})_{sa}'$ with $\|B^{(k)}_i\| \leq R'_i$ such that for every finite $F \subseteq I'$ and every $\varepsilon > 0$, we have

$$\lim_{k \to \infty} (f, \mu^{(k)})(N_{F,\varepsilon}(B^{(k)})) = 1.$$  

Proof. First, observe that by Corollary 2.7, $f, \mu^{(k)}$ is asymptotically supported on the microstate spaces for $y$ in the presence of $x$ relative to $C(k) \rightsquigarrow z$, that is for every neighborhood $U$ of the law of $(x,y,z)$, we have

$$\lim_{k \to \infty} (f, \mu^{(k)})(\Gamma_{R,k}(y : x|C(k) \rightsquigarrow z; U)) = 1.$$  

Second, observe that $f, \mu^{(k)}$ has exponential concentration by Corollary 2.11.

Now fix $\varepsilon > 0$ and a finite index set $F \subseteq I'$. Let

$$\eta := -\limsup_{k \to \infty} \frac{1}{n(k)^2} \log \alpha_{f,\mu^{(k)}}(F, \varepsilon/3),$$

which is strictly positive because $f, \mu^{(k)}$ has exponential concentration. Because $h(\mathcal{N} : \mathcal{M}) = 0$, there exists a neighborhood $U$ of the law of $(x,y,z)$ such that

$$\limsup_{k \to \infty} \frac{1}{n(k)^2} \log K_{F,\varepsilon}(\Gamma_{(R,R'),R_0},n(k)(y : x|C(k) \rightsquigarrow z; U)) < \frac{\eta}{4}.$$  

Thus, for sufficiently large $k$, we have

$$K_{F,\varepsilon/3}(\Gamma_{(R,R'),R_0}, n(k)(y : x|C(k) \rightsquigarrow z; U)) < e^{n(k)^2 \eta/4}$$

and at the same time since $f, \mu^{(k)}$ is asymptotically supported on these microstate spaces, we have

$$(f, \mu^{(k)})(\Gamma_{(R,R'),R_0}, n(k)(y : x|C(k) \rightsquigarrow z; U)) \geq 1/2.$$  

Thus, the microstate space can be covered by the $(F, \varepsilon/3)$-neighborhoods of $e^{n(k)^2 \eta/4}$ many points from the microstate space, while the measure of the whole is at least $1/2$. So by the pigeonhole principle, there is some $B^{(k,F,\varepsilon)} \in M_{n(k)}(\mathbb{C})_{sa}'$ in the microstate space such that

$$(f, \mu^{(k)})(N_{F,\varepsilon/3}(B^{(k,F,\varepsilon)})) \geq \frac{1}{2} e^{-n(k)^2 \eta/4}.$$  

For sufficiently large $k$,

$$(f, \mu^{(k)})(N_{F,\varepsilon/3}(B^{(k,F,\varepsilon)})) > e^{-n(k)^2 \eta/2} > \alpha_{f,\mu^{(k)}}(F, \varepsilon/3),$$
where the last inequality follows from exponential concentration of \( f_i \mu^{(k)} \). Therefore, by Lemma 1.5 we have

\[
(f_i \mu^{(k)})(N_{F, \varepsilon}(B^{(k,F,\varepsilon)})) = (f_i \mu^{(k)})(N_{F, 2\varepsilon/3}(N_{F, \varepsilon/3}(B^{(k,F,\varepsilon)}))) \\
\geq 1 - \alpha_{f_i \mu^{(k)}}(F, \varepsilon/3) \\
\geq 1 - e^{-n(k)^2 \eta/2}.
\]

To complete the proof, we must arrange that the same \( B^{(k)} \) works for all \((F, \varepsilon)\). Let \( X^{(k)} \) be a random matrix tuple given by the probability distribution \( \mu^{(k)} \) and let \( B_i^{(k)} = \mathbb{E}[f_i(X^{(k)})] \). Then for every \((F, \varepsilon)\), the probability that \( \|f_i(X^{(k)}) - B_i^{(k,F,\varepsilon)}\|_2 \leq \varepsilon \) for \( i \in F \) tends to 1. Since this random variable is also uniformly bounded, we have

\[
\limsup_{k \to \infty} \mathbb{E}[f_i(X^{(k)})] - B_i^{(k,F,\varepsilon)}\|_2 \leq \varepsilon.
\]

Thus, since \( B_i^{(k)} = \mathbb{E}[f_i(X^{(k)})] \), the probability that \( \|f_i(X^{(k)}) - B_i^{(k)}\|_2 \leq 2\varepsilon \) for \( i \in F \) tends to 1. Since \((F, \varepsilon)\) was arbitrary, we are done. \( \square \)

Now we finish the proof of our first main theorem.

**Proof of Theorem B.** Assume the setup of 3.4. In particular, suppose that \( P \leq M = W^*(x) \), and let \( \mu^{(k)} \) is a sequence of random matrix measures satisfying (1) - (4). Let \( \mathcal{N} \) be a subalgebra of \( M \) with \( \mathcal{N} \cap P \) diffuse and \( h(\mathcal{N} : M) = 0 \). Let \( y \in \mathcal{N}^{(1)}_{sa} \) be a set of generators for \( \mathcal{N} \) and let \( R' \in (0, +\infty)^I \) satisfy \( \|y_i\| < R'_i \) for \( i \in I' \).

In order to evaluate 1-bounded entropy and apply Proposition 3.3, we fix \( z \in \mathcal{N} \cap P \) self-adjoint with diffuse spectrum, and we obtain a sequence of microstates for \( z \) as follows. By Proposition 2.4, we can write \( z = g(x) \) for some \( g \in (F_{P, \infty})_{sa} \). Let \( R_0 > \|g\|_{P, \infty} \). By Proposition 3.3, since \( z \in P \), there is a sequence \( (C(k))_{k \in \mathbb{N}} \) of matrices such that \( \|g(X^{(k)}) - C^{(k)}\|_{L_2(\mu^{(k)}, M_1(\mathbb{C}))} \to 0 \), where \( X^{(k)} \) is again a random matrix tuple given by \( \mu^{(k)} \).

In order to apply Proposition 3.3, we want to check that \( \mu^{(k)} \) is asymptotically supported on the microstate spaces for \( x \) relative to \( C^{(k)} \). This is equivalent to saying that the non-commutative law of \( (X^{(k)}, C^{(k)}) \) converges to \( \lambda_{(x,z)} \) in probability. But we know by Proposition 2.6 that the non-commutative law of \( (X^{(k)}, g(X^{(k)})) \) converges to \( \lambda_{(x,g(x))} = \lambda_{(x,z)} \) in probability. Also, \( \|g(X^{(k)}) - C^{(k)}\|_2 \to 0 \) in probability. Hence, by Corollary 1.1, we have \( \lambda_{(x,g(x)), C^{(k)}} \to \lambda_{(x,z)} \) in probability.

Therefore, \( \mu^{(k)} \) satisfies the conditions needed for Proposition 3.3. Choose some tuple \( y \in \mathcal{N}^{(1)}_{sa} \) such that \( (y, z) \) generates \( \mathcal{N} \). Let \( R' \) be a tuple of bounds for the operator norms of \( y_i \), and let \( f \in F_{P, \infty} \) such that \( y = f(x) \). Applying the proposition to this \( y \), there exists a sequence \( B^{(k)} \) such that for every finite \( F \subseteq I' \) and \( \varepsilon > 0 \),

\[
\lim_{k \to \infty} (f_i \mu^{(k)})(N_{F, \varepsilon}(B^{(k)}) = 1.
\]

In particular, \( \|f_i(X^{(k)}) - B_i^{(k)}\|_2 \to 0 \) for every \( i \in I' \). So \( y_i = f_i(x) \) is in \( P \) by Proposition 3.2. But \( (y, z) \) generates \( \mathcal{N} \), hence \( \mathcal{N} \subseteq P \). \( \square \)

4. **Examples from Free Probability and Random Matrix Theory**

4.1. **The Generator MASA in** \( L(\mathbb{F}_d) \) **is a Pinsker Algebra.** In this section, we will deduce from Theorem B the following proposition.

**Proposition 4.1.** Let us express the von Neumann algebra \( L(\mathbb{F}_d) \) associated to the free group as the tracial free product \( L(\mathbb{Z}) * L(\mathbb{F}_{d-1}) \). Then \( L(\mathbb{Z}) \) is a Pinsker algebra in \( L(\mathbb{F}_d) \), that is, a maximal subalgebra with 1-bounded entropy zero in the presence of \( L(\mathbb{F}_d) \).

In particular, this implies that \( L(\mathbb{Z}) \) is a maximal amenable subalgebra and a maximal subalgebra with property Gamma. Indeed, if \( \mathcal{N} \geq L(\mathbb{Z}) \) is amenable or Gamma, then \( h(\mathcal{N} : \mathcal{N}) = 0 \), and hence \( h(\mathcal{N} : L(\mathbb{F}_d)) \leq h(\mathcal{N} : \mathcal{N}) = 0 \). So if \( L(\mathbb{Z}) \) is a Pinsker algebra, then \( \mathcal{N} \subseteq L(\mathbb{Z}) \).

Now \( L(\mathbb{Z}) \) is diffuse abelian and has \( h(L(\mathbb{Z}) : L(\mathbb{F}_d)) = 0 \). So for \( L(\mathbb{Z}) \) to be a Pinsker algebra, it would be sufficient to show that if \( \mathcal{N} \cap L(\mathbb{Z}) \) is diffuse and \( h(\mathcal{N} : L(\mathbb{F}_d)) = 0 \), then \( \mathcal{N} \subseteq L(\mathbb{Z}) \), which is precisely the conclusion of Theorem B. Actually, it is no more difficult to establish the conclusion of Theorem B for
the free product of two arbitrary finitely generated Connes-embeddable von Neumann algebras $\mathcal{M}_1$ and $\mathcal{M}_2$ rather than only $L(\mathbb{Z})$ and $L(\mathbb{F}_{d-1})$.

**Proposition 4.2.** Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be Connes-embeddable tracial $W^*$-algebras, and suppose that $\mathcal{M}_1$ is generated by $x = (x_1, \ldots, x_n)$ and $\mathcal{M}_2$ is generated by $y = (y_1, \ldots, y_n)$. Let $(\mathcal{M}, \tau)$ be the (tracial) free product of $\mathcal{M}_1$ and $\mathcal{M}_2$. Then there exist random matrix models for $x$ and $y$ satisfying the hypothesis of Theorem B with $\mathcal{P} = \mathcal{M}_1$. Hence, by that theorem, if $\mathcal{N} \leq \mathcal{M}$, and if $\mathcal{N} \cap \mathcal{P}$ is diffuse with $h(\mathcal{N} : \mathcal{M}) = 0$, then $\mathcal{N} \subseteq \mathcal{P}$.

This proposition is actually a special case of Theorem A which we prove in the next section. However, the proof of Proposition 4.2 requires less technical preparation and already covers the case of the generator MASA in a free group factor. Therefore, we will prove the Proposition here directly, both as an application of Theorem B and as motivation for the proof of Theorem A.

**Proof of Proposition 4.2.** Because $\mathcal{M}_1$ and $\mathcal{M}_2$ are Connes-embeddable, there exist (deterministic) tuples $X^{(k)} = (X_1^{(k)}, \ldots, X_n^{(k)})$ and $Y^{(k)} = (Y_1^{(k)}, \ldots, Y_n^{(k)})$ of $k \times k$ self-adjoint matrices satisfying:

- $\|X_i^{(k)}\| \leq \|x_i\|$ and $\|Y_i^{(k)}\| \leq \|y_i\|$.
- $\lambda_{X^{(k)}} \rightarrow \lambda_x$ and $\lambda_{Y^{(k)}} \rightarrow \lambda_y$.

Let $U^{(k)}$ be a $k \times k$ Haar random unitary matrix, and consider the random matrix tuple $(X^{(k)}, U^{(k)}Y^{(k)}(U^{(k)})^*)$. We claim that these random matrix models satisfy the hypothesis of Theorem B with respect to the generating set $(x, y)$ and the subalgebra $\mathcal{M}_1 \leq \mathcal{M}$.

(1) The random matrices are bounded in operator norm by construction.

(2) We must show that $\lambda_{X^{(k)}, U^{(k)}Y^{(k)}(U^{(k)})^*} \rightarrow \lambda_{(x,y)}$ in probability. Using Voiculescu’s asymptotic freeness theorem, specifically [Voiculescu 1998, Corollary 2.13], if $W^{(k)}$ is another independent Haar unitary, then the non-commutative law of $(W^{(k)}X^{(k)}(W^{(k)})^*, U^{(k)}Y^{(k)}(U^{(k)})^*)$ converges in probability to that of $(x, y)$. But of course, the probability distribution of $\lambda_{X^{(k)}, U^{(k)}Y^{(k)}(U^{(k)})^*}$ is the same as that of $\lambda_{W^{(k)}X^{(k)}(W^{(k)})^*, U^{(k)}Y^{(k)}(U^{(k)})^*}$ by unitary invariance of non-commutative law and the fact that $(W^{(k)}X^{(k)}(W^{(k)})^*)$ is also a Haar unitary.

(3) It is well known in random matrix theory that the Haar measures on the unitary group satisfy exponential concentration of measure. In particular, Meckes and Meckes gave explicit constants for the log-Sobolev inequality on the unitary group in [Meckes and Meckes, Theorem 15], which implies exponential concentration of measure for $U^{(k)}$ with rate $k^2$ with respect to the metric given by $\|\cdot\|_2$. But $(X^{(k)}, U^{(k)}Y^{(k)}(U^{(k)})^*)$ is a Lipschitz function of $U^{(k)}$ and hence also has exponential concentration of measure in the sense of Definition 1.4. (For details, refer to [5.3] below.)

(4) It remains to show the external averaging property. That is, for every non-commutative polynomial $p$ in $m + n$ variables, we have

$$\lim_{k \to \infty} \|E[p(X^{(k)}, U^{(k)}Y^{(k)}(U^{(k)})^*)]\|_2 = \|E_{\mathcal{M}_1}[p(x, y)]\|_2,$$

where $E$ denotes the classical expectation and $E_{\mathcal{M}_1}$ denotes the $W^*$-algebraic conditional expectation.

Let $V^{(k)}$ be another Haar unitary independent from $U^{(k)}$. Then we can rewrite

$$\|E[p(X^{(k)}, U^{(k)}Y^{(k)}(U^{(k)})^*)]\|_2 = \tau_{(k)}(E[p(X^{(k)}, U^{(k)}Y^{(k)}(U^{(k)})^*)]*E[p(X^{(k)}, V^{(k)}Y^{(k)}(V^{(k)})^*)])$$

$$= \tau_{(k)}(p(X^{(k)}, U^{(k)}Y^{(k)}(U^{(k)})^*)*p(X^{(k)}, V^{(k)}Y^{(k)}(V^{(k)})^*)].$$

Similar to the proof of (2), it follows from [Voiculescu 1998, Corollary 2.13] that the non-commutative law of $(X^{(k)}, U^{(k)}Y^{(k)}(U^{(k)})^*)$, $V^{(k)}Y^{(k)}(V^{(k)})^*)$ converges in probability to the non-commutative law of $(x, y, \tilde{y})$, where $\tilde{y}$ is a copy of $y$ freely independent from $x$ and $y$. In other words, $(x, y, \tilde{y})$ are natural generators for $\mathcal{M}_1 \ast \mathcal{M}_2 \ast \mathcal{M}_2$, where $\mathcal{M}_2$ is a copy of $\mathcal{M}_2$.

Therefore,

$$\lim_{k \to \infty} \|E[p(X^{(k)}, U^{(k)}Y^{(k)}(U^{(k)})^*)]\|_2 = \tau(p(x, y)^*p(x, \tilde{y})).$$

Now note that in the algebra $\mathcal{M}_1 \ast \mathcal{M}_2 \ast \mathcal{M}_2$, the two subalgebras $\mathcal{M}_1 \ast \mathcal{M}_2$ and $\mathcal{M}_1 \ast \tilde{\mathcal{M}}_2$ are freely independent with amalgamation over $\mathcal{M}_1$ (see e.g. [Hou 2007, Proposition 4.1]). Therefore, we have

$$E_{\mathcal{M}_1}[p(x, y)]*E_{\mathcal{M}_1}[p(x, \tilde{y})] = E_{\mathcal{M}_1}[p(x, y)]*E_{\mathcal{M}_1}[p(x, \tilde{y})].$$
Given that $\mathcal{M}_1 * \mathcal{M}_2 \cong \mathcal{M}_1 * \hat{\mathcal{M}}_2$, we have $E_{\mathcal{M}_1}[p(x, y)] = E_{\mathcal{M}_1}[p(x, \hat{y})]$. Thus,
\[
\lim_{k \to \infty} \|E[p(X^{(k)}, U^{(k)} Y^{(k)})(U^{(k)})^*])\|^2 = \tau(E_{\mathcal{M}_1}[p(x, \hat{y})p(x, \hat{y})]) \\
= \tau(E_{\mathcal{M}_1}[p(x, y)^* E_{\mathcal{M}_1}[p(x, y)]) \\
= \|E_{\mathcal{M}_1}[p(x, y)]\|^2,
\]
which completes the proof of the external averaging property.

4.2. Random Matrix Models with Convex Interaction. A standard random matrix model for the free group factor $L(F_d)$ is the Gaussian unitary ensemble. One can consider a tuple of $k \times k$ self-adjoint random matrices $S^{(k)} = (S_1^{(k)}, \ldots, S_d^{(k)})$ with the probability density
\[
\frac{1}{\zeta(k)} e^{-k^2 \sum_{j=1}^d \tau_s(A_j^2)/2} \, dA,
\]
on $M_k(\mathbb{C})_{sa}$ where $dA$ is Lebesgue measure and $\zeta(k)$ is a normalizing constant. Then $\lambda_{S^{(k)}}$ converges in probability to the non-commutative law of $s = (s_1, \ldots, s_d)$ where the $s_j$’s are freely independent and each have a semicircular spectral distribution $(1/(2\pi))\sqrt{4 - \tau^2_1[-2, 2]}(t) \, dt$; see [Voia98 Theorem 2.2].

A natural generalization is a random matrix tuple $X^{(k)}$ given by density
\[
\frac{1}{\zeta(k)} e^{-k^2 V^{(k)}(A)} \, dA,
\]
where $V^{(k)}(A) = \tau_k(p(A))$ for some non-commutative polynomial $p$, such that $e^{-k^2 V^{(k)}}$ is integrable. These models are much better understood when $V^{(k)}$ is close to the quadratic case, or is at least convex; see [GMS06 GS09 GS14]. In particular, sufficient assumptions on $V^{(k)}$ will guarantee that $\lambda_{X^{(k)}}$ converges in probability to $\lambda_{x}$ for non-commutative tuple $x$, and that $W^*(x) \cong L(F_d)$.

Our present goal is to explain how the conditional densities of these random matrix models naturally give rise to random matrix models satisfying the hypotheses of Theorem [B] for the subalgebra generated by our original generators (with $n(k) = k$). Changing notation slightly, we consider a function $V^{(k)} : M_k(\mathbb{C})_{sa}^{m+n} \to \mathbb{R}$ denoted $V^{(k)}(A, B)$ for $A \in M_k(\mathbb{C})_{sa}^m$ and $B \in M_k(\mathbb{C})_{sa}^n$. Let $(X^{(k)}, Y^{(k)})$ be the corresponding random matrices, and suppose that $(x, y)$ is a tuple of non-commutative random variables describing the large $k$ limit.

The idea is to consider the probability distribution of $(X^{(k)}, Y^{(k)})$ conditioned on $X^{(k)}$ being equal to some $A^{(k)}$. In other words, fix a particular sequence $A^{(k)}$ in $M_k(\mathbb{C})_{sa}^m$ with $\lambda_{A^{(k)}} \to \lambda_x$ and then choose $B^{(k)}$ randomly according to the conditional distribution
\[
\frac{1}{\zeta(k)}(A^{(k)}) e^{-k^2 V^{(k)}(A^{(k)}, B)} \, dB.
\]
In order to obtain matrix models that are bounded in operator norm, we will replace $B^{(k)}$ with $\psi(B^{(k)})$ where $\psi : \mathbb{R} \to \mathbb{R}$ is a smooth bounded function satisfying $\psi(t) = t$ for $|t| < \|y_i\|$. This should not affect the convergence in law because $\psi(y) = y$.

Since we will invoke the results of [Jek19], we require the same technical setup. Rather than assuming $V^{(k)}(A, B) = \tau_k(p(A, B))$, we assume that $V^{(k)} : M_k(\mathbb{C})_{sa}^{m+n} \to \mathbb{R}$ is differentiable and the gradient $DV^{(k)}$ is asymptotically approximable by trace polynomials. In the notation of this paper, the asymptotic approximation condition means that for all $R \in (0, +\infty)^{m+n}$, there exists $f \in (\mathcal{F}_{R, 2})_{sa}^{m+n}$ such that
\[
\lim_{k \to \infty} \sup_{(A, B) \in M_k(\mathbb{C})_{sa}^{m+n}} \|DV^{(k)}(A, B) - f(A, B)\|_2 = 0.
\]
We further assume that $V^{(k)}$ is invariant under unitary conjugation and that for some $0 < c < C$, the Hessian $HV^{(k)}$ satisfies $c \leq HV^{(k)} \leq C$ (see Jek19 §1.3 and 2.1 for precise details). Under these conditions, Jek18 Theorem 4.1] shows that the random matrix models converge in non-commutative moments in probability to some $x$.

Remark 4.3. Although the assumption $HV^{(k)} \leq C$ rules out the case where $V^{(k)}(A, B) = \tau_k(p(A, B))$ and $p$ is a non-commutative polynomial of degree greater than 2, the condition ($\| \cdot \|_2$ gives us the flexibility to modify our $V^{(k)}$ outside an operator-norm ball, so that the random matrix models can still achieve the same
limiting distribution as if \( V^{(k)}(A, B) = \tau_k(p(A, B)) \) when \( p \) is non-commutative polynomial that is a small perturbation of a quadratic; see [Jek18 §8.3] for details.

Now the following result from [Jek19] describes the large \( k \) behavior of the classical conditional expectation of a certain functions \( f(X^{(k)}, Y^{(k)}) \) given \( X^{(k)} \), and shows that it approximates the non-commutative conditional expectation \( E_{\mathcal{W}^*(\mathcal{X})}[f(x, y)] \). (Beware that \( X \) and \( Y \) are switched in that paper.)

**Theorem 4.4 ([Jek19 Theorem 5.9]).** Let \( V^{(k)} : M_k(\mathbb{C})^{d+n} \to \mathbb{R} \) satisfy the hypotheses described above, and let \( (X^{(k)}, Y^{(k)}) \) be the random matrix models. Let \( f^{(k)} : M_k(\mathbb{C})^{d+n} \to M_k(\mathbb{C}) \) and suppose that \( f^{(k)} \) is uniformly Lipschitz in \( \| \cdot \|_2 \) and that \( f^{(k)} \) is asymptotically approximable by trace polynomials. Let \( g^{(k)}(A) \) be the function defined by

\[
g^{(k)}(X^{(k)}) = \mathbb{E}[f^{(k)}(X^{(k)}, Y^{(k)})|X^{(k)}].
\]

Then \( g^{(k)} \) is also asymptotically approximable by trace polynomials. In fact, if \( R \) is a tuple of bounds for the operator norms of \( x \), then there exists \( g \in \mathcal{F}_{R, 2} \) satisfying

\[
\lim_{k \to \infty} \sup_{\|A, B\| \leq R_k} \|g^{(k)}(A, B) - g(A, B)\|_2 = 0
\]

and \( g(x) = E_{\mathcal{W}^*(\mathcal{X})}[f(x, y)] \).

This is the key to establishing the external averaging property and thus getting random matrix models satisfying the assumptions of Theorem 13.

**Proposition 4.5.** Let \( V^{(k)} : M_k(\mathbb{C})^d \to \mathbb{R} \) satisfy the assumptions of [Jek19] (explained above). Let \( (x, y) \) be the \( (m + n) \)-tuple non-commutative random variables describing the large \( k \) limit of the corresponding random matrix tuples \( (X^{(k)}, Y^{(k)}) \), and let \( R \) and \( S \) satisfy \( \|x_i\| < R_i \) and \( \|y_i\| < S_i \).

Let \( A^{(k)} \) be a deterministic tuple with \( \|A_i^{(k)}\| \leq R_i \) and \( \lambda_{A^{(k)}} \to \lambda_x \). Let \( B^{(k)} \) be a random matrix tuple chosen according to the conditional distribution \( \mathcal{L} \) of \( Y^{(k)} \) given \( X^{(k)} = A^{(k)} \). For \( i = 1, \ldots, n, \) pick \( \psi_i \in C_c^\infty(\mathbb{R}^d; \mathbb{R}) \) with \( |\psi_i| \leq S_i \) and \( \psi_i(t) = t \) for \( |t| \leq \|y_i\| \), and denote \( \psi(B^{(k)}) = (\psi_1(B_1^{(k)}), \ldots, \psi_n(B_n^{(k)})) \).

Then the matrix models \( (A^{(k)}, \psi(B^{(k)})) \) satisfy the hypotheses of Theorem 13 with respect to \( \mathcal{W}^*(\mathcal{X}) \subseteq \mathcal{W}^*(\mathcal{X}, y) \), the generating set \( (x, y) \), and the operator norm bounds \( (R, S) \).

**Proof.** Condition (1) of the theorem holds because we chose \( A_i^{(k)} \) to be bounded in operator norm by \( R_i \), and the function \( \psi_i \) to be bounded by \( S_i \).

To check the exponential concentration hypothesis (3), recall that \( HV^{(k)} \geq c \). It follows by restriction that also \( H[V^{(k)}(A^{(k)}, \cdot)] \geq c \). Then using the standard machinery of the Bakry-Emery condition, the log-Sobolev inequality, and Herbst’s argument, we obtain exponential concentration for the conditional distribution \( \mathcal{L} \). For details, see [ACTZ09 §2.3 and §4.4].

Next, we check (2) that \( \lambda_{(A^{(k)}, \psi(B^{(k)}))} \) converges in probability to \( \lambda_{x, y} \). We want to show that for every non-commutative polynomial \( p \) in \( m + n \) variables, we have \( \tau_k[p(A^{(k)}, \psi(B^{(k)}))] \to \tau[p(x, y)] \) in probability. But due to the concentration of measure, it suffices to show that \( \mathbb{E}\tau_k[p(A^{(k)}, \psi(B^{(k)}))] \to \tau[p(x, y)] \).

Choose a smooth cut-off function \( \phi \) with \( \phi(t) = t \) for \( |t| \leq R_i \) and let \( \phi(A) = (\phi_1(A_1), \ldots, \phi_m(A_m)) \) (analogous to the choice of \( \psi \)). Define \( f^{(k)}(A, B) := \tau_k(p(\phi(A), \psi(B))) \). It follows from the discussion in [Jek18 §8.3] and [Jek19 Lemma 3.14] that \( \phi \) and \( \psi \) are \( \| \cdot \|_2 \)-Lipschitz and asymptotically approximable by trace polynomials, and hence the same is true for \( f^{(k)} \) (since the images of \( \phi \) and \( \psi \) are contained in an operator norm ball, and \( p \) is \( \| \cdot \|_2 \)-Lipschitz on the operator norm ball). Thus, we can apply Theorem 4.4 to \( f^{(k)} \). If \( g^{(k)} \) and \( g \) are as in the theorem, then

\[
g^{(k)}(A^{(k)}) = \mathbb{E}[f^{(k)}(A^{(k)}, \psi(B^{(k)}))].
\]

(As the fact that \( \phi(A^{(k)}) = A^{(k)} \) and

\[
g(x) = E_{\mathcal{W}^*(\mathcal{X})}[p(x, y)]
\]

(As the fact that \( \phi(x) = x \) and \( \psi(y) = y \)). We also know from the theorem that \( \|g^{(k)}(A^{(k)}) - g(A^{(k)})\|_2 \to 0 \). Therefore, using the convergence of \( \lambda_{A^{(k)}} \to \lambda_x \),

\[
\mathbb{E} \circ \tau_k[f^{(k)}(A^{(k)}, \psi(B^{(k)}))] = \tau_k[|g^{(k)}(A^{(k))}|^{1/2}] \to \tau[|g(x)|^{1/2}] = \tau \circ E_{\mathcal{W}^*(\mathcal{X})}[p(x, y)] = \tau[p(x, y)].
\]

This completes the proof of (2).
The external averaging property (4) follows by the same token because with the notation as above, we have
\[ \|E[f^{(k)}(A^{(k)}, \psi(B^{(k)})])\|_2 = \tau_k[g^{(k)}(A^{(k)})*g^{(k)}(A^{(k)})]^{1/2} \to \tau[g(x)^*g(x)]^{1/2} = \|E_W\cdot(x)(p,x,y)\|_2. \]

Therefore, by Theorem [\ref{thm:main}] we have the following result.

**Corollary 4.6.** Let \((x, y)\) be the tuple of non-commutative random variables given above. Let \( \mathcal{M} = W^*(x, y) \) and \( \mathcal{P} = W^*(x) \). Then if \( N \leq M, N \cap \mathcal{P} \) is diffuse and \( h(N : M) = 0, \) then \( N \subseteq \mathcal{P}. \)

In particular, this provides another proof for the case when \((x, y)\) is a free semicircular family, that is, the case of the free group factor which we discussed in the previous subsection. Perhaps disappointingly, the application of this corollary turns out to not be any wider than the free group factor case. Indeed, it was shown in [\ref{Jek19} Theorem 8.11] that under the same set of hypotheses, there is an isomorphism \( \phi : W^*(x, y) \to L(F_{m+n}) \) that maps \( W^*(x) \) to the canonical copy of \( L(F_m) \) inside \( L(F_{m+n}) \). Nonetheless, given that the result about conditional expectations was somewhat easier to prove in [\ref{Jek19}] than the isomorphism result, we suspect that the technique described in this section will have applications to other situations in random matrix theory where the isomorphism to the free group factor setting is not true (or at least not known).

For instance, does this result about conditional expectation extend to the case where \( V^{(k)} \) is not semi-concave, or even not convex? We conjecture that to establish the convergence of the non-commutative law in the large \( k \) limit and the external averaging property (which is weaker than the conclusion of Theorem \[\ref{thm:main}\]), it should be sufficient to assume that \( DV^{(k)} \) is asymptotically approximable by trace polynomials and globally \( \|\cdot\|_2 \)-Lipschitz, and that the distribution of \( X^{(k)} \) and the conditional distribution of \( Y^{(k)} \) given \( X^{(k)} \) satisfy the log-Sobolev inequality with dimension-independent constants (after the appropriate normalization). If this is true, then Proposition \[\ref{prop:main}\] would generalize to this case, hence Theorem \[\ref{thm:main}\] could be applied to produce examples of subalgebras \( \mathcal{P} \subset \mathcal{M} \) that absorb other subalgebras of \( 1 \)-bounded entropy zero.

Another question is whether Proposition \[\ref{prop:main}\] can generalize to measures that are cut off to an operator norm ball. Specifically, suppose that \( V^{(k)} \) is defined on the operator norm ball \( \{A : \|A_j\| \leq R\} \), that \( V^{(k)}(A) - (c/2)\|A\|_2^2 \) is convex, and that \( DV^{(k)} \) is asymptotically approximable by trace polynomials on that ball. Let \( \mu^{(k)} \) be the measure supported on \( \{A : \|A_j\| \leq R\} \) given by the density proportional to \( e^{-k^2V^{(k)}(A)} \) \( dA \). Then if there is a limiting non-commutative law \( \lambda_x \) such that \( \|x_i\| \) is strictly less than \( R \), we conjecture that something similar to Theorem \[\ref{thm:main}\] will hold for the conditional expectation, perhaps after restricting to a smaller operator-norm ball, and hence Proposition \[\ref{prop:main}\] would generalize to this case, without the need for the cut-off function \( \psi \).

5. Proof of Theorem [\ref{thm:main}]

5.1. Overview. To establish Theorem \[\ref{thm:main}\] in the separable case, we consider the amalgamated free product \((\mathcal{M}, \tau) = (\mathcal{M}_1, \tau_1) *_{(\mathcal{D}, \tau_\mathcal{D})} (\mathcal{M}_2, \tau_2) \), where \((\mathcal{M}_1, \tau_1) \) and \((\mathcal{M}_2, \tau_2) \) are tracial \( W^* \)-algebras with separable predual that can embed into \( R^n \), and where \((\mathcal{D}, \tau_\mathcal{D}) \) is a totally atomic \( W^* \)-algebra which is a common tracial \( W^* \)-subalgebra of both \( \mathcal{M}_1 \) and \( \mathcal{M}_2 \).

Choose self-adjoint tuples \( x_1 \in (\mathcal{M}_1)_{1}^{(k)} \) and \( x_2 \in (\mathcal{M}_2)_{2}^{(k)} \) which generate \( \mathcal{M}_1 \) and \( \mathcal{M}_2 \) respectively, where \( I_1 \) and \( I_2 \) are countable index sets. We use the following construction of random matrix models for \( x = (x_1, x_2) \), due to Brown–Dykema–Jung [\ref{BDJ08}]:

1. We choose integers \( n(k) \to \infty \) and unital \(*\)-homomorphisms \( \pi^{(k)} : \mathcal{D} \to M_{n(k)}(\mathbb{C}) \) which are asymptotically trace preserving as \( k \to \infty \).
2. We choose a sequence of (deterministic) microstates \( X^{(k)}_\ell \) for \( x_\ell \) for each \( \ell = 1, 2 \) and arrange that they are compatible with our chosen maps \( \pi^{(k)} : \mathcal{D} \to M_{n(k)}(\mathbb{C}) \) (see Lemma \[\ref{lem:compatibility}\]).
3. Let \( \pi^{(k)}(\mathcal{D})' \) be the commutant of \( \pi^{(k)}(\mathcal{D}) \) in \( M_{n(k)}(\mathbb{C}) \). Let \( U^{(k)} \) be a Haar random unitary from \( \pi^{(k)}(\mathcal{D})' \).
4. Then our random matrix models for generating tuple \( x = (x_1, x_2) \) in the free product will be \( X^{(k)} = (X^{(k)}_1, U^{(k)}X^{(k)}_2(U^{(k)})^*) \).

Adapting the arguments of [\ref{BDJ08}], we will verify that the random matrix models satisfy all the hypotheses of Theorem \[\ref{thm:main}\] for \( \mathcal{M}_1 \leq \mathcal{M} \), that is, boundedness in operator norm, convergence of the non-commutative

moments in probability, exponential concentration, and the external averaging property. At the end of the discussion, we will describe how to deduce the general case of the theorem from the separable case.

5.2. Construction of *-homomorphisms on \( D \) and Compatible Microstates. Now we construct the *-homomorphisms \( \pi(k) : \mathcal{D} \to M_{n(k)} \) and compatible microstates for \( x_1 \) and \( x_2 \).

It is well known that if \( \mathcal{D} \) is totally atomic, then there exist natural numbers \( n(k) \) and unital *-homomorphisms \( \pi(k) : \mathcal{D} \to M_{n(k)}(\mathbb{C}) \) such that \( \tau_n(k) \circ \pi(k) \to \tau_D \) pointwise. We recall an explicit construction of \( \pi(k) \) here, because the decomposition of \( \pi(k)(\mathcal{D}) \) and its commutant in \( M_{n(k)}(\mathbb{C}) \) will be used in the next section to check exponential concentration for Haar unitaries in \( \pi(k)(\mathcal{D})' \).

We assumed \( \mathcal{D} \) is totally atomic, that is, a direct sum of matrix algebras. So there are natural numbers \( \{r(j)\}_{j \in J} \), where \( J = \{1, \ldots, N\} \) or \( J = \mathbb{N} \), and positive weights \( \{\gamma(j)\}_{j \in J} \) with \( \sum_{j \in J} \gamma(j) = 1 \) such that

\[
(D, \tau_D) = \left( \bigoplus_{j \in J} M_{r(j)}(\mathbb{C}), \bigoplus_{j \in J} \gamma(j) \tau_{r(j)} \right).
\]

For each \( k \), we define

\[
m(j, k) = \begin{bmatrix} k \gamma(j) \\ \tau(j) \end{bmatrix}.
\]

Note that for each \( k \), \( m(j, k) \) is nonzero for only finitely many values of \( j \), because in the case where \( J \) is infinite, we have \( \gamma(j) \to 0 \) as \( j \to \infty \). We define

\[
n(k) = \sum_{j \in J} r(j)m(j, k),
\]

Then we have \( r(j)m(j, k)/k \leq \gamma(j) \) and \( r(j)m(j, k)/k \to \gamma(j) \) as \( k \to \infty \). It follows that \( n(k)/k \leq 1 \) and \( n(k)/k \to 1 \) as \( k \to \infty \), and in particular \( n(k) \to \infty \).

We define \( \pi(k) : \mathcal{D} \to M_{n(k)}(\mathbb{C}) \) by

\[
\pi(k) \left( \bigoplus_{j \in J} z_j \right) = \bigoplus_{m(j, k) > 0} z_j \otimes I_{m(j, k)},
\]

or more explicitly, it is the block diagonal matrix

\[
\pi(k) \left( \bigoplus_{j \in J} z_j \right) = \text{diag}\left(z_1, \ldots, z_1, z_2, \ldots, z_2, \ldots, z_3, \ldots\right),
\]

where the \( r(j) \times r(j) \) block \( z_j \) is repeated \( m(j, k) \) times. Since only finitely many \( m(j, k) \)'s are positive for a given \( k \), there are finitely many terms, and overall the matrix has size \( n(k) = \sum_{j \in J} r(j)m(j, k) \).

To check that \( \tau_n(k) \circ \pi(k) \to \tau_D \) pointwise, it suffices to verify this on dense subset of \( \mathcal{D} \). But if \( z = \bigoplus_{j \in J} z_j \in \mathcal{D} \) with only finitely many \( z_j \)'s nonzero, then

\[
\tau_n(k) \circ \pi(k) \left( \bigoplus_{j \in J} z_j \right) = \sum_{j \in J} \frac{r(j)m(j, k)}{n(k)} \tau_{r(j)}(z_j) \to \sum_{j \in J} \gamma(j) \tau_{r(j)}(z_j) = \tau_D \left( \bigoplus_{j \in J} z_j \right).
\]

Next, we have to create microstates for \( x_\ell \) compatible with the chosen maps \( \pi(k) \).

**Lemma 5.1.** For each \( \ell = 1, 2 \), let \( x_\ell \) be an \( I_\ell \)-tuple of generators for \( \mathcal{M}_\ell \), which we assumed to be trivially embeddable into \( \mathcal{R}^\omega \). Let \( \mathbf{R} \in (0, +\infty)^{I_1 \cup I_2} \) with \( \|(x_i)\| \leq R_i \) for \( i \in I_\ell \). Let \( \mathbf{d} = (d_j)_{j \in J} \) be a tuple of operators whose span is dense in \( \mathcal{D} \). Then there exists a (deterministic) tuple \( X^{(k)}_\ell \in M_{n(k)}(\mathbb{C})^{I_\ell}_{sa} \) such that \( (X^{(k)}_\ell, \pi(k)(d)) \) converges in non-commutative *-moments to \( (x_\ell, d) \).

**Proof.** Fix \( \ell \). It will be convenient to use the set of generators \( (x_\ell, d) \) for \( \mathcal{M}_\ell \) rather than \( x_\ell \), and we use the tuple of operator norm bounds \( T \in (0, +\infty)^{I_1 \cup I_2} \) given by \( T_i = R_i \) for \( i \in I_\ell \) and \( T_j = 2\|d_j\| \) for \( j \in J \). We claim that there exist tuples \( Z^{(k)} \in M_{n(k)}(\mathbb{C})^{I_\ell \cup I_2}_{sa} \) satisfying \( \|Z^{(k)}_i\| \leq T_i \) for all \( i \in I_\ell \cup J \) and \( \lambda_{Z^{(k)}} \to \lambda_{(x_\ell, d)} \).

We assumed that \( \mathcal{M}_\ell \) is embeddable into \( \mathcal{R}^\omega \), and it is well-known that a separable tracial von Neumann algebra can be embedded into \( \prod_{n \to \omega} \mathcal{M}_n(\mathbb{C}) \). Indeed, both
conditions are equivalent to the non-commutative law of the generators, in this case \((x_t, d)\), being in the closure in \(\Sigma_T\) of the set of non-commutative laws of matrix tuples. Thus, there exist tuples \(\mathbf{Y}(j)\) of \(n(j) \times n(j)\) matrices, \(j \in \mathbb{N}\), such that \(\|Y_{ij}^{(j)}\| \leq R_i\) and \(\lambda_{R_i}^{(j)} \to \lambda_{(x_t, d)}\) as \(j \to \infty\).

But we can modify our sequence as follows to obtain another sequence of matrix tuples (indexed by \(k\) rather than \(j\)) with the prescribed size \(n(k) \times n(k)\). Choose \(j(k)\) such that \(j(k) \to \infty\) and \(\frac{n(j(k))}{n(k)} \to 0\). Then by integer division, we write \(n(k) = q(k)n(j(k)) + r(k)\) with \(0 \leq r(k) < j(k)\), and we set

\[
Z_{i}^{(k)} := (Y_{i}^{(k)})^{\otimes q(k)} \oplus 0_{r(k) \times r(k)} \in M_{n(k)}(\mathbb{C})_{sa}.
\]

It is an exercise to verify that \(\lambda_{Z^{(k)}} \to \lambda_{(x_t, d)}\).

In particular, the non-commutative law of \(Z^{(k)}|_{J}\) converges to \(\lambda_{d}\). Meanwhile, the non-commutative law of \(\pi^{(k)}(d)\) also converges to \(\lambda_{d}\), and the algebra \(D = W^{*}(d)\) is hyperfinite. So by Lemma 5.6 there exists a sequence \(V^{(k)}\) of unitaries such that

\[
\lim_{k \to \infty} \|V^{(k)}Z_{j}^{(k)}(V^{(k)})^{*} - \pi^{(k)}(d_{j})\|_{2} = 0 \quad \text{for } j \in J.
\]

Now let \(X_{i}^{(k)} = V^{(k)}Z_{j}^{(k)}V^{(k)}\). The non-commutative law of \(V^{(k)}Z^{(k)}(V^{(k)})^{*}\) converges to that of \((x_t, d)\), and also each entry of the tuple \(V^{(k)}Z^{(k)}(V^{(k)})^{*} - (X_{i}^{(k)}, \pi^{(k)}(d))\) goes to zero in \(\|\cdot\|_{q}\) (the \(I\)-indexed entries are already zero). Thus, by Corollary 2.13, the non-commutative law of \((X_{i}^{(k)}, \pi^{(k)}(d))\) converges to \(\lambda_{(x_t, d)}\). \(\square\)

5.3. Exponential Concentration. We want to show that the Haar random unitary from \(\pi^{(k)}(D)^{1/2}\) exhibits exponential concentration as \(k \to \infty\) (Lemma 5.5 below). The finite-dimensional algebra \(\pi^{(k)}(D)\) is a direct sum of matrix algebras, and hence its unitary group is a direct product of unitary groups. Our proof of concentration relies on the log-Sobolev inequality, its behavior under products, and the log-Sobolev constants of the unitary groups.

To state the log-Sobolev inequality, we treat the unitary groups as Riemannian manifolds. Since a unitary matrix can be written as \(e^{iA}\) for \(A\) self-adjoint, the tangent space at each point of \(U(M_n(\mathbb{C}))\) can be identified with \(M_n(\mathbb{C})_{sa}\). The standard Riemannian metric on \(U(M_n(\mathbb{C}))\) is the one defined by using the inner product \(\text{Tr}(A^*B)\) for \(A, B \in M_n(\mathbb{C})_{sa}\) with the standard trace (so \(\text{Tr}(I) = n\)).

**Definition 5.2.** A probability measure \(\mu\) on a Riemannian manifold \(M\) is said to satisfy the log-Sobolev inequality with constant \(c > 0\) if for all \(f \in C^{\infty}_c(M, \mathbb{R})\), we have

\[
\int_M f(x)^2 \log \frac{f(x)^2}{\int f^2 \, d\mu} \, d\mu(x) \leq 2c \int_M \|Df(x)\|^2 \, d\mu(x),
\]

where \(Df\) and \(\|Df\|\) denote the gradient of \(f\) and its norm with respect to the Riemannian metric.

**Theorem 5.3** (see [MM13, Theorem 15]). The unitary group \(U(M_n(\mathbb{C}))\) with the standard Riemannian metric satisfies the log-Sobolev inequality with constant \(c = 6/n\).

**Remark 5.4.** The geodesic distance function \(d(U, V)\) on the unitary group that arises from this Riemannian metric is not the same as the Hilbert-Schmidt distance \(d_{HS}(U, V) = \text{Tr}((U - V)^*(U - V))^{1/2}\). However, we have

\[
d_{HS}(U, V) \leq d(U, V) \leq \frac{\pi}{2} d_{HS}(U, V).
\]

See for instance [Ble09, Lemma 3.9.1]. We will continue to work with the geodesic distance.

We need the following facts about the log-Sobolev inequality and concentration.

**Observation 5.5.** Let \(M\) be a Riemannian manifold and suppose \(\mu\) is a probability measure satisfying the log-Sobolev inequality with constant \(c\). Fix \(t > 0\), and let \(\hat{M}\) be the same manifold with the rescaled Riemannian metric \((x, y)|\hat{M} = t(x, y)|M\), where \(x\) and \(y\) are tangent vectors at a point \(p \in M\). Then \(\mu\) satisfies the log-Sobolev inequality with constant \(ct\).

**Lemma 5.6** (see [Led01, Corollary 5.7]). Let \(M_1, \ldots, M_n\) be Riemannian manifolds, and let \(\mu_j\) be a probability measure on \(M_j\) satisfying the log-Sobolev inequality with constant \(c_j\). Let \(M = M_1 \times \cdots \times M_n\) with product Riemannian metric. Then the product measure \(\mu = \mu_1 \otimes \cdots \otimes \mu_n\) on \(M\) satisfies the log-Sobolev inequality with constant \(c = \max(c_1, \ldots, c_n)\).
Lemma 5.7 (see [Led01 Corollary 5.4]). Suppose $\mu$ is a probability measure on the Riemannian manifold $M$. Let $d$ be the geodesic distance on $M$, that is the metric obtained from infimizing the lengths of paths (measured using the Riemannian metric), and let $N_\varepsilon(\Omega)$ denote the $\varepsilon$-neighborhood of a set. Let $\alpha_\mu$ be the metric concentration function defined by

$$ \alpha_\mu(\varepsilon) = \sup\{\mu(N_\varepsilon(\Omega)) : \Omega \text{ Borel with } \mu(\Omega) \geq 1/2\}.$$ 

If $\mu$ satisfies the log-Sobolev inequality with constant $c$, then

$$ \alpha_\mu(\varepsilon) \leq e^{-\varepsilon^2/8c}. $$

Sketch of proof. First, the argument of Herbst (see [Led01 Theorem 5.3] or [ACZ00 Lemma 2.3.3]) shows that for the Lipschitz $f : M \to \mathbb{R}$ and $\lambda \geq 0$,

$$ A_\lambda := \int_M e^{\lambda(f(x)-\int f \, d\mu)} \, d\mu(x) \leq e^{\varepsilon^2/2} e^{-\varepsilon^2/8c}. $$

Second, one shows that for $\delta > 0$,

$$ \mu(\{x : f(x) - \int f \, d\mu \geq \delta\}) \leq e^{-\delta^2/2c} e^{-\varepsilon^2/8c}. $$

This follows from the previous estimate by substituting $\lambda = \delta/c$ and using Markov’s inequality.

Finally, to estimate the concentration function $\alpha_\mu$, fix a Borel set $\Omega$ with $\mu(\Omega) \geq 1/2$. Define

$$ f(x) = \min \left( \frac{d(x,\Omega)}{\varepsilon}, 1 \right). $$

Since $f \leq 1$ and $\mu(\Omega) \geq 1/2$, we have $\int f \, d\mu \leq 1/2$. Also, $f$ is $1/\varepsilon$-Lipschitz. Thus,

$$ \mu(N_\varepsilon(\Omega)^c) = \mu(\{x : f(x) = 0\}) \leq \mu(\{x : f(x) - \int f \, d\mu \geq 1/2\}) \leq e^{-(1/2)^2/2c} e^{-\varepsilon^2/8c}. $$

□

Lemma 5.8. Let $G^{(k)}$ be the unitary group of $\pi^{(k)}(D)'$, equipped with the Riemannian metric associated to the normalized trace $\tau_{n(k)}$ (in other word, identify the tangent space at a point with the self-adjoints of $\pi^{(k)}(D)'$ and use the inner product associated to the normalized trace from $M_{n(k)}(\mathbb{C})$). Let $\nu^{(k)}$ be the Haar measure on $G^{(k)}$, and let $\alpha_{\nu^{(k)}}$ denote the concentration function. Then for every $\varepsilon > 0$,

$$ \limsup_{k \to \infty} \frac{1}{n(k)} \log \alpha_{\nu^{(k)}}(\varepsilon) < 0. $$

Proof. Let us first compute the unitary group of $\pi^{(k)}(D)'$ more explicitly. The image of $\pi^{(k)}(D)$ can be expressed as

$$ \pi^{(k)}(D) = \bigoplus_{j \in J} M_{r(j)} \otimes I_{m(j,k)} \subseteq \bigoplus_{j \in J} M_{r(j)}(\mathbb{C}) \otimes M_{m(j,k)}(\mathbb{C}) \cong \bigoplus_{j \in J} M_{r(j)m(j,k)}(\mathbb{C}) \subseteq M_{n(k)}(\mathbb{C}), $$

where the inclusions and identifications at each step are the standard ones. It is well known (and easy to verify) that the commutant of this subalgebra is

$$ \pi^{(k)}(D)' = \bigoplus_{j \in J} I_{r(j)} \otimes M_{m(j,k)}(\mathbb{C}) \subseteq \bigoplus_{j \in J} M_{r(j)m(j,k)}(\mathbb{C}) \subseteq M_{n(k)}(\mathbb{C}). $$

Hence,

$$ G^{(k)} \cong \prod_{j \in J} U(M_{m(j,k)}(\mathbb{C})), $$

and of course the Haar measure on $G^{(k)}$ is the product of the Haar measures on the individual factors. As in the case of the unitary groups, we identify the tangent space with the self-adjoints of $\pi^{(k)}(D)'$, or the direct sum of $M_{m(j,k)}(\mathbb{C})_{sa}$ over $j$ with $m(j,k) > 0$. Then the Riemannian metric is given by

$$ \left\langle \bigoplus_j A_j, \bigoplus_j B_j \right\rangle = \frac{1}{n(k)} \sum_{j \in J, m(j,k) > 0} r(j) \text{Tr}(A_j B_j). $$
where the associated Riemannian metrics are given by the weights

\[ m(\nu) = \sum_{\epsilon \in \pi_r} \epsilon \, U \]

The Riemannian metric, if we rescale the metric by \( \gamma \), we have \( \gamma' \). Here \( \Omega \) times the normalized Hilbert-Schmidt distance

As in Remark 5.4, the corresponding geodesic distance on \( G^{(k)} \) is bounded above and below by constants times the normalized Hilbert-Schmidt distance

\[ \| (U_j - V_j)_j \|_2 = d_{\text{HS}}((U_j)_j, (V_j)_j) = \left( \frac{1}{n(k)} \sum_{j \in J, m(j,k) > 0} r(j) \text{Tr}((U_j - V_j)^*(U_j - V_j)) \right)^{1/2}. \]

Since the number of direct summands is unbounded as \( k \to \infty \) if \( J \) is infinite, we will not study the log-Sobolev inequality on \( G^{(k)} \) itself, but rather we will truncate to a fixed number of summands. Fix \( \varepsilon > 0 \). Because the weights \( \gamma(j) \) in our direct sum decomposition sum to 1, there exists a finite \( N \in \mathbb{N} \) such that

\[ \sum_{j=1}^{N} \gamma(j) > 1 - \varepsilon^2 4 \left( \frac{\pi}{2} \right)^2. \]

Moreover, since \( m(j,k) r(j)/n(k) \to \gamma(j) \) as \( k \to \infty \), we know that for sufficiently large \( k \),

\[ \sum_{j=1}^{N} \frac{m(j,k) r(j)}{n(k)} > 1 - \varepsilon^2 4 \left( \frac{\pi}{2} \right)^2. \]

It follows that for \( U = (U_j)_j \) and \( V = (V_j)_j \) in \( G^{(k)} \), if \( U_j = V_j \) for \( j \leq N \), then \( d(U,V) \leq 2/\pi \| U - V \|_2 < \varepsilon/2 \). Define

\[ G^{(k)}_{\leq N} = \prod_{j=1}^{N} U(M_{m(j,k)}(\mathbb{C})), \quad G^{(k)}_{> N} = \prod_{j>N} U(M_{m(j,k)}(\mathbb{C})) \]

with the associated Riemannian metrics given by the weights \( m(j,k) r(j)/n(k) \), and denote the associated Haar measures by \( \nu^{(k)}_{\leq N} \) and \( \nu^{(k)}_{> N} \).

Now, since \( U(M_{m(j,k)}(\mathbb{C})) \) with the satisfies log-Sobolev with constant \( 6/m(j,k) \) with the standard Riemannian metric, if we rescale the metric by \( r(j)/n(k) \), then the log-Sobolev inequality holds with constant \( 6r(j)/n(k)m(j,k) \). Then the direct product \( G^{(k)}_{\leq N} \) satisfies the log-Sobolev inequality with constant

\[ c^{(k)}_N := \max \left\{ \frac{6r(j)}{n(k)m(j,k)} : j = 1, \ldots, N \right\}. \]

Fix a Borel set \( \Omega \subseteq G^{(k)} \) with \( \nu^{(k)}(\Omega) \geq 1/2 \). Let \( \Omega' \) be the projection of \( \Omega \) onto \( G^{(k)}_{\leq N} \). By our choice of \( N \), we have

\[ N_{\varepsilon/2}(\Omega) \supseteq \Omega' \times G^{(k)}_{> N}. \]

Here \( \Omega' \) is a subset of \( G^{(k)}_{\leq N} \) with Haar measure at least \( 1/2 \).

On the other hand, by applying Lemma 5.7 on \( G^{(k)}_{\leq N} \), we have

\[ \nu^{(k)}(N_{\varepsilon}(\Omega)^c) \leq \nu^{(k)}_{\leq N}(N_{\varepsilon/2}(\Omega')^c) \leq \alpha^{(k)}_{\nu^{(k)}_{\leq N}}(\varepsilon/2) \leq e^{-\varepsilon^2/32^{(k)}_N}. \]

Hence,

\[ \frac{1}{n(k)^2} \log \alpha^{(k)}_{\nu^{(k)}}(\varepsilon) \leq -\varepsilon^2 \frac{32^{(k)}_N}{32^{(k)}_N}. \]

But

\[ n(k)^2 c^{(k)}_N = \max \left\{ \frac{6r(j)n(k)}{m(j,k)} : j = 1, \ldots, N \right\}, \]

and

\[ \frac{6r(j)n(k)}{m(j,k)} = \frac{6r(j)^2 n(k)}{r(j)m(j,k)} \to \frac{6r(j)^2}{\gamma(j)} < +\infty, \]

so that \( \limsup_{k \to \infty} n(k)^2 c^{(k)}_N < +\infty \), which implies that \( \limsup_{k \to \infty} n(k)^{-2} \log \alpha^{(k)}_{\nu^{(k)}}(\varepsilon) < 0. \)

\[ \square \]

**Corollary 5.9.** Let \( X^{(k)}_{\ell} \) be the microstates for \( \pi_{\ell}^{(k)} \), where \( \ell = 1, 2 \), chosen in Remark 5.4 and let \( U^{(k)} \) be a Haar random unitary from \( \pi^{(k)}(\mathcal{D})' \). Then the random matrix tuple \( X^{(k)} = (X^{(k)}_1, U^{(k)} X^{(k)}_2 (U^{(k)})^*) \) satisfies the exponential concentration property of Definition 1.4.
Proof. For \(i \in I_2\), we have \(f^{(k)}_i : U^{(k)} \to U^{(k)}(X^{(k)}_2)_{(i)}(U^{(k)})^*\) is \(2R_t\)-Lipschitz with respect to \(\|\cdot\|_2\) because \(\|X^{(k)}_2\|_i \leq R_t\) for \(i \in I_2\). But the normalized Hilbert-Schmidt norm is bounded by the geodesic distance on \(G^{(k)}\), hence the function is also \(2R_t\)-Lipschitz if we use the geodesic distance in the domain. Of course, for \(i \in I_1\), the constant function \(f^{(k)}_i : U^{(k)} \to (X^{(k)}_1)_i\) is trivially Lipschitz. Now, \(U^{(k)}\) has exponential concentration by the previous lemma. Thus, using the same argument as in Corollary 5.11 the exponential concentration is preserved when we push forward by the tuple of functions \((f^{(k)}_i)_{i \in I_1 \cup I_2}\), since each \(f^{(k)}_i\) is \(\|\cdot\|_2\)-uniformly continuous with estimates independent of \(k\). \(\square\)

5.4. Asymptotic freeness with amalgamation. The key point in establishing the convergence in moments and external averaging property (hypotheses (2) and (4) of Theorem 5.9) will be the following theorem. In the case where \(D\) is finite-dimensional, this is a restatement of [BDJ08 Theorem 3.9], and we merely extend it to the totally atomic case by an approximation argument. Note that this theorem considers countably many algebras \(M_{\ell}\) rather than only two algebras \(M_{1}\) and \(M_{2}\) as in 5.4.4.

Theorem 5.10. Let \(\{M_{\ell}, \tau_{M_{\ell}}\}_{\ell \in \mathbb{N}}, (M_{2}, \tau_{M_{2}}), \ldots\) be tracial von Neumann algebras that contain countable totally atomic subalgebras \((D, \tau_D)\) in a trace-preserving way. Let \(M\) be the free product of \(\{M_{\ell}\}_{\ell \in \mathbb{N}}\) with amalgamation over \(D\), and let us view \(M_{\ell}\) as a subalgebra of \(M\) in the canonical way.

For each \(\ell\), let \(x_{\ell} \in \{M_{\ell}\}_{\ell \in \mathbb{N}}\) be a generating tuple for \(M_{\ell}\) with \(I_{\ell}\) an index set. Let \(I = \bigcup_{\ell \in \mathbb{N}} I_{\ell}\) and let \(x\) be a \(I\)-tuple \((x_{\ell})_{\ell \in \mathbb{N}}\). Suppose \(R \in (0, +\infty)^I\) with \(\|x_{\ell}\| \leq R_{\ell}\) for \(\ell \in I_{\ell}\).

Let \(n(k) \to \infty\). Let \(\pi^{(k)} : D \to M_{n(k)}(\mathbb{C})\) be a *-homomorphism satisfying \(\tau_{n(k)} \circ \pi^{(k)} \to \tau_D\). Let \(d \in D\) be a tuple of operators whose span is dense in \(D\), obtained as the union of spanning sets for each of the finite-dimensional matrix algebras that are direct summands of \(D\). For each \(\ell\), let \(X^{(k)}_{\ell} \in M_{n(k)}(\mathbb{C})_{I_{\ell}}\) with \(\|X^{(k)}_{\ell}\| \leq R_{\ell}\) and such that \((X^{(k)}_{\ell}, \pi^{(k)}(d)) \to (x_{\ell}, d)\) in non-commutative *-moments. Let \(\{U^{(k)}_{\ell}\}_{\ell \in \mathbb{N}}\) be a family of independent Haar unitaries from \(\pi^{(k)}(D)\).

Let \(X^{(k)}\) be the \(I\)-tuple \((U^{(k)}_{\ell}X^{(k)}_{\ell}(U^{(k)}_{\ell})^*)_{\ell \in \mathbb{N}}\). Then \(\lambda_{X^{(k)}} \to \lambda_x\) in \(\Sigma_{\mathbb{R}}\) in probability as \(k \to \infty\).

Proof. The case where \(D\) is finite-dimensional and the index sets \(I_{\ell}\) are finite follows from [BDJ08 Theorem 3.9]. The extension to infinite index sets \(I_{\ell}\) is automatic. Indeed, for \(\lambda_{X^{(k)}}\) to converge to \(\lambda_x\) in probability means that for every \(f \in \mathbb{C}(t_i : i \in \bigcup_{\ell} I_{\ell})\), we have \(\tau_{n(k)}(f(X^{(k)})) \to \tau(f(x))\). But \(f\) only depends on finitely many of the \((x_{\ell})_{\ell}\)'s so the convergence follows from the case for finite index sets.

Now we must extend from finite-dimensional \(D\) to totally atomic \(D\). As in the proof of Lemma 5.8 for general totally atomic \(D\), we must approximate by truncating to finitely many direct summands. Given \(\varepsilon > 0\), choose \(N\) large enough that \(\sum_{j=1}^{N} \gamma(j) > 1 - \varepsilon^2\). Let \(p \in D\) be the (central) projection onto the first \(N\) direct summands. Let \(P^{(k)} = \pi^{(k)}(p)\) be the \(k\)th matrix approximation of \(p\). Then \(\tau_D(p) > 1 - \varepsilon^2\) and hence \(\tau_{n(k)}(P^{(k)}) \geq 1 - \varepsilon^2\) for sufficiently large \(k\). It follows that for every matrix \(C \in M_{n(k)}(\mathbb{C})\), we have

\[
\|C - P^{(k)}CP^{(k)}\|_2 \leq \|(1 - P^{(k)})C\|_2 + \|P^{(k)}(1 - P^{(k)})\|_2 \leq 2\varepsilon\|C\|_2.
\]

To show that \(\tau_{n(k)}(f(X^{(k)})) \to \tau(f(x))\) in probability for all \(f \in \mathbb{C}(t_i : i \in \bigcup_{\ell} I_{\ell})\), it suffices to consider the case where \(f\) is a monomial, say

\[
f(t_i : i \in \bigcup_{\ell} I_{\ell}) = \prod_{j=1}^{J} t_{i(j)},
\]

where the terms in the product are indexed from left to right. Let \(J\) be the index with \(i(j) \in I_{\ell(j)}\), so that

\[
f(X^{(k)}) = \prod_{j=1}^{J} U^{(k)}_{\ell(j)}X^{(k)}_{\ell(j)}(U^{(k)}_{\ell(j)})^*,
\]

where \(X^{(k)}_{\ell(j)}\) is the \(i(j)\)-indexed element of the tuple \(X^{(k)}_{\ell(j)}\). Next, we replace each term in the product with its truncation by \(P^{(k)}\) and estimate the error for each “swap” by (5.1) to obtain

\[
\|f(X^{(k)}) - \prod_{j=1}^{k} P^{(k)}U^{(k)}_{\ell(j)}P^{(k)}X^{(k)}_{\ell(j)}P^{(k)}(U^{(k)}_{\ell(j)})^*P^{(k)}\|_2 \leq (2\varepsilon)(3J)\prod_{j=1}^{J} \|X^{(k)}_{\ell(j)}\| \leq 6\varepsilon J\prod_{j=1}^{J} \|\frac{P^{(k)}X^{(k)}_{\ell(j)}P^{(k)}}{1 - P^{(k)}}\|
\]
since the total number of swaps is $3J$ and $\|X_{1(i)}^{(k)}\| \leq R_{i(j)}$. Analogously,

\[(5.3) \quad \left\| f(x) - \prod_{j=1}^{J} u_{\ell(j)} x_{i(j)} u_{\ell(j)}^{*} \right\|_{2} \leq 6\varepsilon J \prod_{j=1}^{J} R_{i(j)}^{(k)}.\]

Next, we will evaluate the limit of the trace of the truncated version of $f(X^{(k)})$ by applying the finite-dimensional case of the theorem to the algebras $pM_{i}\rho$ and the finite-dimensional subalgebra $pD_{\rho}$. We consider each of the compressions $pM_{i}\rho$, $pD_{\rho}$, and $pM_{i}$ as tracial von Neumann algebras by renormalizing the trace from the ambient algebra by $1/\tau(p)$, as is standard. Note that $\{pM_{i}\rho\}_{i\in\mathbb{N}}$ are freely independent in $pM_{\rho}$ with amalgamation over $pD_{\rho}$; this is an exercise to check from the definition of free independence and the fact that $p \in D$.

Similarly, we consider the truncated matrix algebra $P^{(k)}M_{n(k)}(\mathbb{C})P^{(k)}$ with the trace renormalized by $1/\tau_{n(k)}(P^{(k)})$. Note that $\{P^{(k)}U_{\ell}^{(k)}P^{(k)}\}_{\ell \in \mathbb{N}}$ are independent Haar unitaries from $\pi^{(k)}(pD_{\rho})\cap P^{(k)}M_{n(k)}(\mathbb{C})P^{(k)}$. The matrix tuples $(P^{(k)}X_{\ell}^{(k)}P^{(k)}, P^{(k)}\pi^{(k)}(d)P^{(k)})$ converge in non-commutative $*$-moments to $(x_{\ell}, p_{\ell}d)$ with respect to the renormalized traces on the truncated algebras (of course, since $\tau_{n(k)}(P^{(k)}) \to \tau(p)$, it would be equivalent to show this convergence with respect to the traces on the ambient algebras). But we chose $d$ to contain a spanning set for each of the direct summands of $D$, and hence $p \in \text{Span}(d)$. Therefore, any $*$-polynomial $f$ in $(p_{\ell}x_{\ell}, p_{\ell}d)$ can be expressed as a $*$-polynomial $g$ in $(x_{\ell}, d)$ such that we also have $f(P^{(k)}X_{\ell}^{(k)}P^{(k)}, P^{(k)}\pi^{(k)}(d)P^{(k)}) = g(X_{\ell}^{(k)}, \pi^{(k)}(d))$ (this uses the fact that $\pi^{(k)}$ is a $*$-homomorphism).

All these observations mean that the case of the theorem where $D$ is finite-dimensional can be applied to the truncated version of $f(X^{(k)})$, and so we obtain that in probability,

$$\lim_{k \to \infty} \frac{1}{\tau_{n(k)}(P^{(k)})} \tau_{n(k)} \left( \prod_{j=1}^{k} P^{(k)}U_{\ell}^{(k)}P^{(k)}X_{i(j)}^{(k)}P^{(k)}(U_{\ell(j)}^{(k)})^{*}P^{(k)} \right) = \frac{1}{\tau(p)} \tau \left( \prod_{j=1}^{J} p_{\ell(j)}x_{i(j)}u_{\ell(j)}^{*} \right).$$

Of course, since $\tau_{n(k)}(P^{(k)}) \to \tau(p) > 0$, we can remove the terms $1/\tau_{n(k)}(P^{(k)})$ and $1/\tau(p)$ from the equation. We combine this with (5.2) and (5.3) to conclude that in probability,

$$\limsup_{k \to \infty} \left| \tau_{n(k)}[f(X^{(k)})] - \tau[f(x)] \right| \leq 12\varepsilon J \prod_{j=1}^{J} R_{i(j)}^{(k)}.$$

Since $\varepsilon$ was arbitrary, we are finished. \hfill $\square$

**Remark 5.11.** Clearly, the same theorem is true if we only have finitely many algebras $M_{\ell}$ rather than countably many.

**Remark 5.12.** The conclusion of the theorem still holds if we replace the first unitary $U_{1}^{(k)}$ by the identity. This is because the non-commutative law is invariant under unitary conjugation, so we can conjugate $X_{i}^{(k)}$ by $(U_{1}^{(k)})^{*}$ and note that $\{(U_{1}^{(k)})^{*}U_{i}^{(k)}\}_{i \geq 2}$ has the same probability distribution as $\{U_{i}^{(k)}\}_{i \geq 2}$.

### 5.5. Conclusion to the proof.

**Proof of Theorem [4]**

The separable case: First, assume that $M_{1}$ and $M_{2}$ are separable preduals. We retain all the setup and notation from [5.1] and [5.2]. We want to show that random matrix tuples $X^{(k)} = (X_{1}^{(k)}, U^{(k)}X_{2}^{(k)}(U^{(k)})^{*})$ satisfy the hypotheses of Theorem [3] with respect to $M_{1} \subseteq M = M_{1} \oplus_{D} M_{2}$ and the chosen generators $x = (x_{1}, x_{2})$ and operator norm bounds $R$. We check each of the four hypotheses of Theorem [3] in turn.

1. By construction $\|X_{i}^{(k)}\| \leq R_{i}$ for $i \in I_{1}$ and $\|U^{(k)}X_{2}^{(k)}(U^{(k)})^{*}\| = \|X_{2}^{(k)}\| \leq R_{i}$ for $i \in I_{2}$.
2. The fact that $\lambda_{X^{(k)}} \to \lambda_{x}$ follows from Theorem [5.10] and the two remarks given above. In other words, we apply the version of the theorem for two algebras with the first unitary replaced by the identity.
3. Corollary [5.9] showed that these random matrix models have exponential concentration.
4. To check the external averaging property, let $f \in \mathbb{C}\{t_{i} : i \in I_{1} \cup I_{2}\}$, and we will show that

$$\lim_{k \to \infty} \|\mathbb{E}[f(X^{(k)})]\|_{2} = \|\mathbb{E}_{M_{1}}[f(x)]\|_{2}.$$
We write
\[ \|E[f(X^{(k)})]\|_2^2 = \tau_{n(k)} \left[ E[f(X_1^{(k)}, U^{(k)}X_2^{(k)}(U^{(k)})^*)] \cdot E[f(X_1^{(k)}, V^{(k)}X_2^{(k)}(V^{(k)})^*)] \right] = E \circ \tau_{n(k)} \left[ f(X_1^{(k)}, U^{(k)}X_2^{(k)}(U^{(k)})^*) \cdot f(X_1^{(k)}, V^{(k)}X_2^{(k)}(V^{(k)})^*) \right] \]
where $V^{(k)}$ is an independent Haar unitary from $\pi^{(k)}(D)'$.

Let $M_3$ be an isomorphic copy of $M_2$ and let $x_3$ be the tuple corresponding to $x_2$. Let $Q = M_1 \ast_D M_2 \ast_D M_3$. Applying Theorem 6.10 to the free product of $M_1$, $M_2$, and $M_3$ (again using the two remarks following the theorem), we obtain
\[ \lim_{k \to \infty} E \circ \tau_{n(k)} \left[ f(X_1^{(k)}, U^{(k)}X_2^{(k)}(U^{(k)})^*) \cdot f(X_1^{(k)}, V^{(k)}X_2^{(k)}(V^{(k)})^*) \right] = \tau_Q[f(x_1, x_2)^*f(x_1, x_3)]. \]

There is a canonical isomorphism

\[ Q \cong (M_1 \ast_D M_2) \ast_D (M_1 \ast_D M_3), \]

or in other words, $M_1 \ast_D M_2$ and $M_1 \ast_D M_3$ are freely independent in $Q$ with amalgamation over $M_1$ (see [Hou07 Proposition 4.1]). This implies that
\[ E_{M_1}[f(x_1, x_2)^*f(x_1, x_3)] = E_{M_1}[f(x_1, x_2)]^*E_{M_1}[f(x_1, x_3)]. \]

Of course, using the obvious isomorphism $M_1 \ast M_2 \cong M_1 \ast M_3$, we have $E_{M_1}[f(x_1, x_3)] = E_{M_1}[f(x_1, x_2)]$. Thus, we have
\[ \tau_Q[f(x_1, x_2)^*f(x_1, x_3)] = \tau[E_{M_1}[f(x_1, x_2)]^*E_{M_1}[f(x_1, x_2)]], \]
and hence
\[ \lim_{k \to \infty} \|E[f(X_1^{(k)}, U^{(k)}X_2^{(k)}(U^{(k)})^*)]\|_2^2 = \|E_{M_1}[f(x_1, x_2)]\|_2^2. \]

Therefore, the hypotheses of Theorem B are satisfied, and so Theorem B implies Theorem A in the separable case.

**The general case:** Now suppose that $M = M_1 \ast_D M_2$ where $D$ is totally atomic and $M_1$ and $M_2$ are not necessarily separable. Suppose that $N \cap M_1$ is diffuse and $h(N : M) = 0$. It suffices to show that if $N^0$ is a diffuse separable subalgebra of $N$, then $N^0 \subseteq M_1$.

Note that $h(N^0 : M) \leq h(N : M) = 0$ by Property 2 of $h$ from [1, 2]. Letting $x \in (N^0)^I$, let be a generating tuple for $N^0$ and $y$ be a generating tuple for $M$, this means that for every finite $F \subseteq I$ and $\varepsilon > 0$ and $\delta > 0$, there exists a neighborhood $U$ of $\lambda_{x,y} \in \mathcal{U}$ such that
\[ \limsup_{n \to \infty} \frac{1}{n^2} \log K_{F,x}(\Gamma_{N,n}(x : y|C(n) \to z; U)) < \delta. \]

Recall $\inf_{\mathcal{U}}$ of the above expression will increase if $\varepsilon$ becomes smaller and $F$ becomes larger. Thus, by iterating over countably many values of $F$, $\varepsilon$, and $\delta$, we see that only countably many neighborhoods $U$ are needed to witness that the $\sup_{F,x} \inf_{\mathcal{U}}$ of the above expression is zero. And each neighborhood only specifies conditions on finitely many generators of $M$. Thus, there exists a separable subalgebra $M^0$ such that $h(N^0 : M^0) = 0$.

We can choose separable subalgebras $M^0_1 \leq M_1$ and $M^0_2 \leq M_2$ such that $M^0 \leq M^0_1 \ast_D M^0_2$, and hence $h(N^0 : M^0_1 \ast_D M^0_2) = 0$. By making $M^0_1$ larger if necessary, we can arrange that $N^0 \cap M^0_1$ is diffuse. Then it follows from the separable case of the theorem that $N^0 \subseteq M^0_1 \subseteq M_1$ as desired. \(\square\)

**Remark 5.13.** While much of the proof could be adapted to the case of amenable $D$, Theorem A does not hold for $D$ amenable, as we saw in the introduction. The place where the proof would break down is in achieving exponential concentration of measure. Indeed, even using matrix approximations for an amenable algebra, the exponential concentration on the relative commutants of $\pi^{(k)}(D)$ in the matrix models decays in the limit. Thus we lose the exponential concentration of measure when $D$ is amenable and non-atomic.
6. Further Remarks

6.1. The External Averaging Property and Ultraproducts. The external averaging property (hypothesis (4) from Theorem B) can alternatively be stated in terms of a commuting square property [Pop83a, Pop90] for embeddings into ultraproducts.

Let \((M, \tau)\) be a tracial W*-algebra, let \(x\) be an \(I\)-tuple of self-adjoint generators, and let \(R \subseteq (0, +\infty)^I\) with \(\|x_i\| \leq R_i\). Suppose that \(X^{(k)}\) is an \(I\)-tuple of random \(n(k) \times n(k)\) self-adjoint matrices satisfying \(\|X^{(k)}_i\| \leq R_i\) and \(\lambda_{X^{(k)}} \rightarrow \lambda_x\) in probability.

Let us realize the \(X^{(k)}\)'s on the same probability space \(\Omega\). Then \(X^{(k)}_i\) can be viewed as an element of \(L^\infty(\Omega, M_{n(k)}(\mathbb{C})) = L^\infty(\Omega) \otimes M_{n(k)}(\mathbb{C})\). Consider the tracial W*-ultra product

\[
\prod_{k \rightarrow \omega} L^\infty(\Omega, M_{n(k)}(\mathbb{C})).
\]

Let \(y_i\) be the element of the ultraproduct represented by \(\{X^{(k)}_i\}_{k \in \mathbb{N}}\). Then \(y = (y_i)_{i \in I}\) has the same non-commutative law as \(x\), and hence there is an embedding \(\rho_\omega : M \rightarrow \prod_{k \rightarrow \omega} L^\infty(\Omega, M_{n(k)}(\mathbb{C}))\) given by \(\rho_\omega(x_i) = y_i\).

We also remark that there is a canonical inclusion \(M_{n(k)}(\mathbb{C}) \rightarrow L^\infty(\Omega, M_{n(k)}(\mathbb{C}))\), mapping \(A\) to the deterministic function \(A\) on \(\Omega\), and hence we have an inclusion

\[
\iota_\omega : \prod_{k \rightarrow \omega} M_{n(k)}(\mathbb{C}) \rightarrow \prod_{k \rightarrow \omega} L^\infty(\Omega, M_{n(k)}(\mathbb{C})).
\]

**Proposition 6.1.** Let \((M, \tau), x, R, X^{(k)}, y\) be as above, and let \(P \subseteq M\) be a W*-subalgebra. Then the following are equivalent:

1. \(X^{(k)}\) satisfies the external averaging property for \(P \subseteq M\).
2. For every free ultrafilter \(\omega\) on \(\mathbb{N}\), the embedding \(\rho_\omega : M \rightarrow \prod_{k \rightarrow \omega} L^\infty(\Omega, M_{n(k)}(\mathbb{C}))\) described above satisfies

\[
\rho_\omega(P) = \rho_\omega(M) \cap \iota_\omega \left( \prod_{k \rightarrow \omega} M_{n(k)}(\mathbb{C}) \right),
\]

and we have commutativity of the diagram

\[
\begin{CD}
M @>{\rho_\omega}>> \prod_{k \rightarrow \omega} L^\infty(\Omega, M_{n(k)}(\mathbb{C})) \\
@VV{E_P}V @VV{E_\omega}V \\
\bar{P} @>{\rho_\omega}>> \prod_{k \rightarrow \omega} M_{n(k)}(\mathbb{C}),
\end{CD}
\]

where \(E\) denotes the trace-preserving conditional expectation.

**Proof.** (1) \(\implies\) (2). Fix \(z \in M\) and write \(z = f(x)\) for some \(f \in \mathcal{F}_{R, \infty}\). By Lemma 3.1 we have \(f(z) \in \mathcal{P}\) if and only if

\[
\lim_{k \rightarrow \infty} \|f(X^{(k)}) - E[f(X^{(k)})]\|_{L^2(\Omega, L^2(M_{n(k)}(\mathbb{C}), \tau_{n(k)})}) = 0.
\]

But since ultraproducts commute with trace-preserving conditional expectation, this occurs if and only if

\[
\|f(y) - E_\omega[f(y)]\|_2 = 0.
\]

This proves that \(\rho_\omega(P) = \rho_\omega(M) \cap \iota_\omega \left( \prod_{k \rightarrow \omega} M_{n(k)}(\mathbb{C}) \right)\).

To check the commutativity condition, it suffices to consider \(z \in P\) (which we already did) and \(z \perp \mathcal{P}\). By Lemma 3.1 \(z \perp \mathcal{P}\) if and only if

\[
\lim_{k \rightarrow \omega} \|E[f(X^{(k)})]\|_2 = 0.
\]

But this is equivalent to \(\|E_\omega[f(y)]\|_2 = 0\).

(2) \(\implies\) (1). Fix \(f \in \mathcal{F}_{R, \infty}\). Then for every ultrafilter \(\omega\), we have

\[
\lim_{k \rightarrow \omega} \|E[f(X^{(k)})]\|_2 = \|E_\omega[f(y)]\|_2 = \|E_\omega[f(x)]\|_2,
\]

where the last equality follows from the commutativity of the diagram and the fact that \(\rho_\omega(x) = y\). Since the ultrafilter \(\omega\) was arbitrary, we have \(\lim_{k \rightarrow \infty} \|E[f(X^{(k)})]\|_2 = \|E_\omega[f(x)]\|_2\). \(\square\)
6.2. Change of Generators and Passing to Subalgebras. Now we show that the hypotheses of Theorem [B] are independent of the choice of generators for \( M \). More precisely, if there exist measures satisfying these hypotheses for one choice of generators \( x \) for \( M \), and if \( y \) is another choice of generators for \( M \), then there also exist measures satisfying these hypotheses for \( y \). This is the special case of Proposition 6.2 below where we take \( Q = M \).

Another special case worth pointing out is if \( P \leq Q \leq M \); in this case, the proposition says that if such measures \( \mu^{(k)} \) for the inclusion \( P \leq M \), then they also exist for the inclusion \( P \leq Q \) when \( Q \) is an intermediate subalgebra between \( P \) and \( M \).

Proposition 6.2. Let \((M, \tau)\) be a tracial W*-algebra, let \( P \) be a W*-subalgebra. Suppose that \( x \in M_{sa}^I \) generates \( M \), let \( y \in M_{sa}^{I'} \), and let \( P = W^*(y) \). Suppose that \( Q \) satisfies the commuting square condition that \( E_P \circ E_Q = E_{P \cap Q} \).

Let \( R \in (0, +\infty) \) and \( R' \in (0, +\infty) \) be bounds for the operator norms of \( x_i \) and \( y_i \), respectively, and \( f \in \mathcal{F}_{R, R'} \) be a function such that \( y \equiv f(x) \).

Suppose that \( \{\mu^{(k)}\} \) is a sequence of probability measures on \( M_{n(k)}(\mathbb{C})^{I_{sa}} \) satisfying the hypotheses (1) - (4) with respect to \( P \subseteq M \) and \( x \) and \( R \). Then \( \{f, \mu^{(k)}\} \) satisfies hypotheses (1) - (4) with respect to \( P \cap Q \subseteq Q \) and \( y \) and \( R' \).

Proof. (1) We have \( \|f_i(a)\|_{}\leq R_i \) for all \( i \in I' \) and any self-adjoint tuple \( a \) with \( \|a_i\| \leq R_i \) for \( i \in I \). In particular, this holds when we evaluate \( f_i \) on the random matrix tuple \( X^{(k)} \) associated to \( \mu^{(k)} \).

(2) Since \( \lambda_{X^{(k)}} \to \lambda_x \) in probability, we have \( \lambda_{f(X^{(k)})} \to \lambda_{f(x)} \) in probability by Proposition 2.6.

(3) Exponential concentration is inherited by \( f, \mu^{(k)} \) by Corollary 2.14.

(4) Let \( q \in C(t_i : i \in I') \). Then \( q \circ f \in \mathcal{F}_{R, \infty} \) since \( \mathcal{F}_{R, \infty} \) is a \( * \)-algebra. Therefore,

\[
\lim_{k \to \infty} \left\| \int q d(\mu^{(k)}) \right\|_2 = \lim_{k \to \infty} \left\| \int q \circ f d\mu^{(k)} \right\|_2 = \|E_P[q(f(x))]\|_2 = \|E_P[q(y)]\|_2 = \|E_{P \cap Q}[q(y)]\|_2,
\]

where we applied the external averaging property for \( \mu^{(k)} \) and the commuting square condition for \( P \cap Q \). \( \square \)

6.3. Separability Issues. Note that the proof of Theorem [B] did not require any separability assumptions on \( P \) and \( M \). However, in the non-separable case, we should not expect there to exist a sequence of random matrix models in general, since indeed \( M \) might not embed into the ultraproduct \( \prod_{n \to \omega} M_{n}(\mathbb{C}) \) when \( \omega \) is a free ultrafilter on \( N \).

From an abstract point of view, the natural adaptation to the non-separable case would be to use a net of random matrix models. That is, we take \( k \to K \) to be an index from some directed set \( K \), and then consider random matrix tuples \( X^{(k)} \) indexed by \( k \in K \). We replace the limits as \( k \to \infty \) in \( N \) by limits as over \( k \in K \), and then the proof of Theorem [B] will go through without further changes. The same is true for the statements in [6.1] and [6.2] we leave the details to the reader.

The proof of Theorem [A] could also be modified to construct a net of random matrix models for \( \mathcal{M}_1 \ast_{\mathcal{P}} \mathcal{M}_2 \) directly without going through the reduction to the separable case. This only requires modification of the construction of microstates for \( \mathcal{M}_1 \) and \( \mathcal{M}_2 \) in [6.2] to obtain a net of *-homomorphisms and microstates indexed by \( k \) in some directed set \( K \) rather than \( k \in \mathbb{N} \). The results about concentration and asymptotic freeness still apply if we arrange that \( n(k) \to \infty \) over \( k \in K \).

However, for the typical situations that arise in random matrix theory, one would only need to consider separable algebras and sequences of random matrix models. Moreover, as we saw in the proof of Theorem [A] the von Neumann algebraic applications can often be reduced to the separable case because 1-bounded entropy is something defined through finitary approximations to begin with.

We remark that the existence of a net of random matrix models satisfying the hypotheses of Theorem [B] for a given generating set \( x \in M_{sa}^I \) is equivalent to following statement: For every finite \( F \subseteq I \), there exists a sequence \( X^{(k)} \) of random F-tuples of \( n(k) \times n(k) \) matrices satisfying (1) operator norm bounds, (2) convergence of \( \lambda_{X^{(k)}} \) to \( \lambda_{X|F} \) in probability, (3) exponential concentration, and (4) for every \( f \in C(t_i : i \in F) \),

\[
\lim_{k \to \infty} \|E[f(X^{(k)})]\|_2 = \|E_{P}[f(x)]\|_2.
\]
6.4. A Free Decomposability Conjecture. So far, we have not found any examples other than amalgamated free products where Theorem B applies. In fact, it is conceivable, based on an analogy with ergodic theory, that the only way the hypotheses of Theorem B can be satisfied is if there exists such an amalgamated free product decomposition. Therefore, we propose the following conjecture.

**Conjecture.** Let $\mathcal{P} \leq \mathcal{M}$ be an inclusion of tracial von Neumann algebras. Suppose there exist generators $(x_i)_{i \in I}$, a sequence $n(k) \to \infty$ of natural numbers, and self-adjoint $n(k) \times n(k)$ random matrices $(X^{(k)}_i)_{i \in I}$ satisfying (1)-(4) of Theorem B. Then there is von Neumann subalgebra $\mathcal{N} \leq \mathcal{M}$, and an atomic $\mathcal{D} \leq \mathcal{M}$ so that $\mathcal{N}, \mathcal{P}$ are freely independent with amalgamation over $\mathcal{D}$, and so that $\mathcal{M} = \mathcal{N} \vee \mathcal{P}$.

This conjecture is motivated by the following analogy with ergodic theory. Suppose $\mathbb{Z} \curvearrowright^T (X, \mu)$ is a probability measure-preserving action, and suppose we are given a measurable map $\alpha : X \to A$ where $A$ is a finite set. Assume further that $\alpha$ is a *generator*, i.e. $\bigvee_n T^n(\{f \circ \alpha : f \in \mathbb{C}^A\}) = L^\infty(X, \mu)$. In this case, the partition $\{1_{\alpha^{-1}(a)} : a \in A\}$ is analogous to the $(x_i)_{i \in I}$ in the free probabilistic context. Moreover, there is a natural interpretation of *microstates* for such actions (see [Bow10]).

We can then form a natural analogue of having a sequence of measures which is asymptotically supported on the microstate spaces, satisfies exponential concentration of measure, and has the external averaging property. In the setting of probability measure-preserving $\mathbb{Z}$-actions, the existence of a sequence of measures which satisfies the analogue of these three properties is equivalent to the *almost blowing up property* (first defined in [MS94, Definition 4] under the name “blowing up property”). This property is one of the conditions from Ornstein theory (see [Orn70a, Orn70b, OWS7, Orn74, Shi73]) which guarantees that the $\mathbb{Z}$-action is isomorphic to a Bernoulli shift. (This is shown in [MS94, 3.2.2].) This relation to concentration of measure and the action being Bernoulli was recently used to great effect in [Aus18] to solve the weak Pinsker conjecture.

So in the context of free probability, it should be the case that these three conditions (asymptotic concentration on microstates, exponential concentration of measure, and the external averaging property) imply that $\mathcal{M}$ is “freely Bernoulli over $\mathcal{P}$.” Note that free products may be viewed as a free analogue of Bernoulli shifts. For example if we consider the infinite free product $P^\infty \times \mathbb{Z}$ and if $\mathbb{Z} \curvearrowright P^\infty \times \mathbb{Z}$ by free shifts, then $P^\infty \times \mathbb{Z} \cong P \ast L(\mathbb{Z})$. So one interpretation of being “freely Bernoulli over $\mathcal{P}$” is being free complemented with amalgamation over an atomic subalgebra. So Conjecture 6.4 may be viewed as the natural free probability analogue of the fact from Ornstein theory that the almost blowing up property for a $\mathbb{Z}$-action implies the action is Bernoulli.

**References**

[AGZ09] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni, *An introduction to random matrices*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2009.

[AK19] Scott Atkinson and Srivatsav Kunnawalkam Elayavalli, *On ultraproduct embeddings and amenability for tracial von Neumann algebras*, arXiv e-prints (2019), arXiv:1907.03359

[Aus18] Tim Austin, *Measure concentration and the weak Pinsker property*, Publ. Math. Inst. Hautes Études Sci. **128** (2018), 1–119. MR 3905465

[BB03] Serban T. Belinschi and Hari Bercovici, *A property of free entropy*, Pacific J. Math. **211** (2003), no. 1, 35–40.

[BC15] Rémi Boutonnet and Alessandro Careri, *Maximal amenable von Neumann subalgebras arising from maximal amenable subgroups*, Geometric and Functional Analysis **25** (2015), no. 6, 1688–1705.

[BDJ08] Nathanial P. Brown, Kenneth J. Dykema, and Kenley Jung, *Free entropy dimension in amalgamated free products*, Proceedings of the London Mathematical Society **97** (2008), no. 2, 339–367.

[BH18] Rémi Boutonnet and Cyril Houdayer, *Amenable absorption in amalgamated free product of von Neumann algebras*, Kyoto J. Math. **58** (2018), no. 3, 583–593.

[Blo09] Gordon Blower, *Random matrices: High dimensional phenomena*, London Mathematical Society Lecture Note Series, Cambridge University Press, 2009.

[Bow10] Lewis Bowen, *Measure conjugacy invariants for actions of countable sofic groups*, J. Amer. Math. Soc. **23** (2010), no. 1, 217–245. MR 2552252

[Bro14] Arnaud Brothier, *The cup subalgebra of a II₁ factor given by a subfactor planar algebra is maximal amenable*, Pacific J. Math. **269** (2014), no. 1, 19–29. MR 3233908

[BW16] Arnaud Brothier and Chenxi Wen, *The cup subalgebra has the absorbing amenability property*, Internat. J. Math. **27** (2016), no. 2, 1650013. 6. MR 3464393

[Ceb13] G Cébron, *Free convolution operators and free Hall transform*, Journal of Functional Analysis **265** (2013), 2645–2708.

[CFRW10] Jan Cameron, Junsheng Fang, Mohan Ravichandran, and Stuart White, *The radial masa in a free group factor is maximal injective*, J. Lond. Math. Soc. (2) **82** (2010), no. 3, 787–809. MR 2739068
The large-

B. K. Driver, B. C. Hall, and T. Kemp,

Model theory of operator algebras i: stability

J. Dixmier,

Some unique group-measure space decomposition results

Ionut Chifan and Jesse Peterson,

Two applications of free entropy, Math. Ann. 308 (1997), no. 3, 547–558. MR 1457745

Junsheng Fang,

On maximal injective subalgebras of factors of type II1, Ann. of Math. (2) 147 (1998), no. 1, 143–157. MR 1609522

Alice Guionnet and Edward Maurel-Segala,

Combinatorial aspects of random matrix models, Alea 1 (2006), 241–279.

Liming Ge and Sorin Popa,

On some decomposition properties for factors of type II1, Duke Math. J. 94 (1998), no. 1, 79–101. MR 1635904

Alin Galatan and Sorin Popa,

Smooth bimodules and cohomology of II1 factors, J. Inst. Math. Jussieu 16 (2019), no. 1, 57–137. MR 3635672

Alice Guionnet and Dimitri Shlyakhtenko,

Free diffusions and matrix models with strictly convex interaction, Geom. and Functional Analysis 18 (2009), no. 6, 1875–1916.

Ben Hayes,

Relative entropy and the Pinsker product formula for sofic groups, Groups Geom. Dyn., to appear.

Ben Hayes,

Mixing and spectral gap relative to Pinsker factors for sofic groups, Proceedings of the 2014 Maui and 2015 Qinhuangdao conferences in honour of Vaughan F. R. Jones’ 60th birthday, Proc. Centre Math. Appl. Austral. Nat. Univ., vol. 46, Austral. Nat. Univ., Canberra, 2017, pp. 193–221. MR 3635672

Cyril Houdayer,

On some free products of von Neumann algebras which are free Araki-Woods factors, Int. Math. Res. Not. IMRN (2018), no. 1, 57–137. MR 3801429

Cyril Houdayer and Dimitri Shlyakhtenko,

Strongly solid II1 factors with an exotic MASA, Int. Math. Res. Not. IMRN (2007), no. 23, Art. ID rnm008, 21. MR 2377217

A class of II1 factors with an exotic abelian maximal amenable subalgebra, Trans. Amer. Math. Soc. 366 (2014), no. 7, 3693–3707. MR 3192613

Structure of II1 factors arising from free Bogoljubov actions of arbitrary groups, Adv. Math. 260 (2014), 414–457. MR 3209358

Gamma stability in free product von Neumann algebras, Comm. Math. Phys. 336 (2015), no. 2, 831–851. MR 3322388

Cyril Houdayer and Dimitri Shlyakhtenko,

Strongly solid II1 factors with an exotic MASA, Int. Math. Res. Not. IMRN (2011), no. 6, 1352–1380. MR 2806507

Masaki Izumi, Roberto Longo, and Sorin Popa,

A Galois correspondence for compact groups of automorphisms of von Neumann algebras with a generalization to Kac algebras, J. Funct. Anal. 155 (1998), no. 1, 25–63. MR 1622812

Adrian Ioana,

Rigidity of von neumann algebras, Delivered at the 2018 Proceedings of the International Congress of Mathematicians. Preprint arXiv:1712.00151.

Classification and rigidity for von Neumann algebras, European Congress of Mathematics, Eur. Math. Soc., Zürich, 2014, pp. 601–625.

Cartan subalgebras of amalgamated free product II1 factors, Ann. Sci. Éc. Norm. Supér. (4) 48 (2015), no. 1, 71–130, With an appendix by Ioana and Stefaan Vaes. MR 335839

Adrian Ioana, Jesse Peterson, and Sorin Popa,

Amalgamated free products of weakly rigid factors and calculation of their symmetry groups, Acta Math. 200 (2008), no. 1, 85–153. MR 2386109

David Jekel,

An elementary approach to free entropy theory for convex potentials, To appear in Analysis and PDE Journal.

Conditional expectation, entropy, and transport for convex gibbs laws in free probability, arXiv:1906.10051.
Kenley Jung, *Amenability, tubularity, and embeddings into $\mathcal{R}^\omega$*, Math. Ann. **338** (2007), no. 1, 241–248. MR 2295511

Kenley Jung, *Strongly 1-bounded von Neumann algebras*, Geom. Funct. Anal. **17** (2007), no. 4, 1180–1200. MR 2373014

T. Kemp, *The large-n limits of Brownian motions on $gl(n)$*, Int. Math. Res. Not. IMRN (2016), no. 13, 4012–4057.

T. Kemp, *Heat kernel empirical laws on $u(n)$ and $gl(n)$*, J. Theoret. Probab. **30** (2017), no. 2, 397–451.

David Kerr, *Sofic measure entropy via finite partitions*, Groups Geom. Dyn. **7** (2013), no. 3, 617–632. MR 3095712

Brian Leary, *Maximal Amenability with Asymptotic Orthogonality in Amalgamated Free Products*, arXiv e-prints (2019).

Michel Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, vol. 89, American Mathematical Society, Providence, RI, 2001. MR 1849347

Elizabeth S. Meckes and Mark W. Meckes, *Gaussian von Neumann algebras*, Proc. Amer. Math. Soc. **135** (2007), no. 7, 2331–2340. MR 2316379

Donald Ornstein, *Bernoulli shifts with the same entropy are isomorphic*, Advances in Math. **4** (1970), 337–352. MR 0257322

Donald Ornstein, *Two Bernoulli shifts with infinite entropy are isomorphic*, Advances in Math. **5** (1970), 339–348 (1970). MR 0274716

Donald Ornstein and Benjamin Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, J. Analyse Math. **48** (1987), 1–141. MR 910005

Narutaka Ozawa, *On a class of II$_1$ factors with at most one Cartan subalgebra*, Ann. of Math. (2) **172** (2010), no. 1, 713–749. MR 2680430

Narutaka Ozawa and Sorin Popa, *On a class of II$_1$ factors with at most one Cartan subalgebra, II*, Amer. J. Math. **132** (2010), no. 3, 841–866. MR 2669099

Donald Ornstein and Benjamin Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, J. Analyse Math. **48** (1987), 1–141. MR 910005

Narutaka Ozawa, *A remark on amenable von Neumann subalgebras in a tracial free product*, Proc. Japan Acad. Ser. A Math. Sci. **91** (2015), no. 7, 104.

M. S. Pinsker, *Dynamical systems with completely positive or zero entropy*, Soviet Math. Dokl. **1** (1960), 937–938. MR 0152628

Sandeepan Parekh, Koichi Shimada, and Chenzhu Wen, *Maximal amenability of the generator subalgebra in q-Gaussian von Neumann algebras*, J. Operator Theory **80** (2018), no. 1, 125–152. MR 3835452

Jesse Peterson and Andreas Thom, *Group cocycles and the ring of affiliated operators*, Invent. Math. **185** (2011), no. 3, 561–592. MR 2827955

Sandeepan Parekh and Stefaan Vaes, *Unique Cartan decomposition for II$_1$ factors arising from arbitrary actions of free groups*, Acta Math. **212** (2014), no. 1, 141–198. MR 3179609

Sandeepan Parekh and Stefaan Vaes, *Unique Cartan decomposition for II$_1$ factors arising from arbitrary actions of hyperbolic groups*, J. Reine Angew. Math. **694** (2014), 215–239. MR 3259044

David Kerr, *Sofic measure entropy via finite partitions*, Groups Geom. Dyn. **7** (2013), no. 3, 617–632. MR 3095712

Narutaka Ozawa and Sorin Popa, *On a class of II$_1$ factors with at most one Cartan subalgebra*, Ann. of Math. (2) **172** (2010), no. 1, 713–749. MR 2680430

Sandeepan Parekh, Koichi Shimada, and Chenzhu Wen, *Maximal amenability of the generator subalgebra in q-Gaussian von Neumann algebras*, J. Operator Theory **80** (2018), no. 1, 125–152. MR 3835452

Jesse Peterson and Andreas Thom, *Group cocycles and the ring of affiliated operators*, Invent. Math. **185** (2011), no. 3, 561–592. MR 2827955

Sandeepan Parekh and Stefaan Vaes, *Unique Cartan decomposition for II$_1$ factors arising from arbitrary actions of free groups*, Acta Math. **212** (2014), no. 1, 141–198. MR 3179609

Sandeepan Parekh and Stefaan Vaes, *Unique Cartan decomposition for II$_1$ factors arising from arbitrary actions of hyperbolic groups*, J. Reine Angew. Math. **694** (2014), 215–239. MR 3259044

E.M. Rains, *Combinatorial properties of brownian motion on the compact classical groups*, Journal of Theoretical Probability **10** (1997), 659–679.

B. Seward, *The Koopman representation and positive Rokhlin entropy*, arXiv:1602.06680.
A RANDOM MATRIX APPROACH TO ABSORPTION IN FREE PRODUCTS

[B. Hayes] DEPARTMENT OF MATHEMATICS, UNIVERSITY OF VIRGINIA, 141 CABELL DRIVE, KERCHOF HALL, CHARLOTTESVILLE, VA, 22904
E-mail address: brh5c@virginia.edu

(D. Jekel) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, BOX 951555, LOS ANGELES, CA 90095-1555
E-mail address: davidjekel@math.ucla.edu

(B. Nelson) DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, 619 RED CEDAR ROAD, C212 WELLS HALL, EAST LANSING, MI 48824-3402
E-mail address: brent@math.msu.edu

(T. Sinclair) DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, 150 N. UNIVERSITY ST., WEST LAFAYETTE, IN 47907-2067
E-mail address: tsincla@purdue.edu

[Sewb] ———, Weak containment and Rokhlin entropy, [arXiv:1602.06680]

[She06] Junhao Shen, Maximal injective subalgebras of tensor products of free group factors, J. Funct. Anal. 240 (2006), no. 2, 334–348. MR 2261686

[Shi73] Paul C. Shields, The theory of Bernoulli shifts, The University of Chicago Press, Chicago, Ill.-London, 1973, Chicago Lectures in Mathematics. MR 0442198

[Shi96] ———. The ergodic theory of discrete sample paths, Graduate Studies in Mathematics, vol. 13, American Mathematical Society, Providence, RI, 1996. MR 1400225

[Vae07] Stefaan Vaes, Rigidity results for Bernoulli actions and their von Neumann algebras (after Sorin Popa), Astérisque (2007), no. 311, Exp. No. 961, viii, 237–294, Séminaire Bourbaki. Vol. 2005/2006. MR 2359046

[Vae10] ———. Rigidity for von Neumann algebras and their invariants, Proceedings of the International Congress of Mathematicians. Volume III, Hindustan Book Agency, New Delhi, 2010, pp. 1624–1650. MR 2827858 (2012g:46006)

[Voi91] Dan-Virgil Voiculescu, Limit laws for random matrices and free products, Invent. Math. 104 (1991), no. 1, 201–220. MR 1094052

[Voi94] ———, The analogues of entropy and of Fisher’s information measure in free probability theory II, Invent. Math. 118 (1994), no. 3, 411–440. MR 1296352

[Voi96] ———, The analogues of entropy and of Fisher’s information measure in free probability theory. III. The absence of Cartan subalgebras, Geom. Funct. Anal. 6 (1996), no. 1, 172–199. MR 1371236

[Voi98] ———, A strengthened asymptotic freeness result for random matrices with applications to free entropy, Internat. Math. Res. Not. IMRN (1998), no. 1, 41–63. MR 1601878

[Wen16] Chenxu Wen, Maximal amenability and disjointness for the radial masa, J. Funct. Anal. 270 (2016), no. 2, 787–801. MR 3423903