FREE PROBABILITY & THE FREE CENTRAL LIMIT THEOREM

Matthew Esmaili Mallory
Department of Statistics, Harvard University
matthewmallory@fas.harvard.edu

Abstract. Free probability theory extends the classical ideas of probability theory to the non-commutative case. The fundamental idea is to abstract away the notion of a sample space and σ-algebra of events to work with more general algebras of random variables. Many classical results in probability have a free analogue, such as the free law of large numbers of free central limit theorem. Such results in free probability, such as the free CLT, generalize known results such as Wigner's semi-circle law for random matrices. More recent work has been in extending the free CLT to dependent and high-dimensional settings, and in applying combinatorial methods in free probability toward concentration inequalities. We begin with an introduction to free probability and then delve into such extensions and applications of the theory.

1. Introduction

Free probability, the study of non-commutative random variables, has become a rapidly expanding field in the last few decades. It sits at the intersection of several topics, such as probability, random matrices, operator algebras, and functional & complex analysis. The notion of free independence was first considered by Voiculescu in 1986 [10], in which the idea was equated to freeness of subgroups $(G_i)_{i \in \mathcal{I}}$ of a group algebra $\mathbb{C}G$. In this manner, it could be viewed as a non-commutative analogue to the classical notion of independence of random variables.

With the idea of free independence in place, it is only natural to then ask if one can extend classical universality results such as the Law of Large Numbers (LLN) or Central Limit Theorem (CLT) to the non-commutative setting. It turns out that this is indeed possible [3, 6, 7], and such a Free CLT and its recent extensions [1] and applications of free probability [2] will be the subject of this paper. We begin with an introduction to free probability.

2. Free Probability

2.1. **Non-Commutative Probability Spaces.** In order to understand free probability, we must first begin with the original foundations of classical probability theory laid out by Kolmogorov in 1933:

(1) We select a **sample space** Ω, with **states** $\omega \in \Omega$.

(2) We then construct a σ-algebra $\mathcal{B} \subseteq 2^{\Omega}$ of **events**, such that $P(\Omega) = 1$.
(3) We build a commutative algebra of random variables $X : \Omega \to \mathbb{C}$, with an expectation function $\mathbb{E}(X)$.

The 3-tuple $(\Omega, \mathcal{B}, P)$ is then called our probability space. In free probability, we wish to "abstract away" both (1) and (2), as this will help us both compute limits and work with more general, non-commutative objects. To accomplish this, a non-commutative probability space is built solely upon the algebra of random variables and an associated linear functional:

Definition 2.1.1. A **C*-algebra** $\mathcal{A}$ over $\mathbb{C}$ is a vector space over $\mathbb{C}$ equipped with a bi-linear operation, an anti-automorphic involution $* : \mathcal{A} \to \mathcal{A}$, and a norm $\|\cdot\| : \mathcal{A} \to [0, \infty)$ under which it is complete and

$$\|ab\| \leq \|a\|\|b\|, \quad \|a^*\| = \|a\|^2.$$

In the same vein, a **von Neumann algebra** is a special type of C*-algebra: it is a $*$-algebra of bounded operators on a Hilbert space, which is also closed in the weak operator topology. Although these are very important, for the rest of the paper we will assume we are working with just an ordinary C*-algebra.

Definition 2.1.2. A (**non-commutative**) **C*-probability space** $(\mathcal{A}, \varphi)$ consists of a unital C*-algebra $\mathcal{A}$ over $\mathbb{C}$, and a unital, positive linear functional $\varphi : \mathcal{A} \to \mathbb{C}$.

One usually also assumes that τ also has the following two properties:

(1) $\varphi(ab) = \varphi(ba)$ for all $a, b \in \mathcal{A}$ (**tracial**)
(2) $\varphi(a^*a) = 0 \implies a = 0_{\mathcal{A}}$ (**faithful**)

Example 2.1.3. Since our course deals with random matrices, it is only natural to consider $\mathcal{A} = M_d\left(L^{\infty-}(\Omega, P)\right)$, the space of $d \times d$ random matrices whose entries have bounded p^{th}-moments of all orders $p \geq 1$. In this case, the obvious choice for our linear functional is

$$\varphi(a) := \frac{1}{d} \int_{\Omega} \operatorname{tr}(a(\omega)) dP(\omega), \quad a \in \mathcal{A},$$

which is the integral of the normalized trace. This makes $(\mathcal{A}, \varphi)$ a proper C*-probability space.

Example 2.1.4. Let G be a group, and let $\mathbb{C}G$ be its associated **group algebra**:

$$\mathbb{C}G := \left\{ \sum_{g \in G} \alpha_g g \;\middle|\; a_g \in \mathbb{C},\ \text{only finitely many } a_g \neq 0 \right\}.$$

The **canonical trace** on the group algebra is given by

$$\tau_G\left(\sum_{g\in G}\alpha_g g\right) = \alpha_e,$$

which makes $(\mathbb{C}G, \tau_G)$ a faithful, tracial $*$-probability space. This example will turn out to be important later on, as it was the inspiration for free probability in the first place, which we will see soon.

2.2. **$*$-Distributions.** Now that we understand what a $*$-probability space is, we can begin to describe what the distribution of an arbitrary $a \in \mathcal{A}$ would look like. Recall that if X is a classic random variable, then it (often) suffices to understand the moments $\mathbb{E}(X^n)$ for $n \geq 1$. For example, the $\mathcal{N}(0,1)$ distribution is uniquely defined by its sequence of moments $\big(\mathbb{E}(X^n)\big)_{n\geq 1} = (0,1,0,3,0,15,\ldots)$. However, this will not suffice for a general non-commutative random variable, as we need to consider all possible words formed by a and its adjoint a^*:

Definition 2.2.1. A **$*$-moment** of a is given by $\varphi(a^{\varepsilon_1}\cdots a^{\varepsilon_k})$, where $k \geq 0$ and each $\varepsilon_i \in \{1,*\}$. If we let $\mathbb{C}\langle X, X^*\rangle$ be the unital algebra freely generated by two non-commuting indeterminates X, X^*, then the **$*$-distribution** of a is the linear functional $\mu : \mathbb{C}\langle X, X^*\rangle \to \mathbb{C}$ such that

$$\mu(X^{\varepsilon_1}\cdots X^{\varepsilon_k}) = \varphi(a^{\varepsilon_1}\cdots a^{\varepsilon_k}).$$

Definition 2.2.2. The **spectrum** of $a \in \mathcal{A}$ is given by

$$\operatorname{spec}(a) := \{z \in \mathbb{C} : z1_{\mathcal{A}} - a \text{ is not invertible}\}.$$

Of course, this should look familiar: when $\mathcal{A}$ consists of matrices, then the spectrum is simply the set of eigenvalues. When x is self-adjoint, we have $\operatorname{spec}(x) \subseteq \mathbb{R}$, and if u is unitary, then $\operatorname{spec}(u) \subseteq \mathbb{T}$. In general though, for an arbitrary $a \in \mathcal{A}$ the spectrum will lie within the disk $D_{\|a\|}(0)$. It turns out that the spectrum of a random variable is intimately connected to its $*$-distribution, in the following way:

Proposition 2.2.3. Let $a \in \mathcal{A}$ be normal. Then a has a $*$-distribution μ whose support is contained in $\operatorname{spec}(a)$. If φ is faithful, then the support is exactly $\operatorname{spec}(a)$.

In the context of random matrices, this is (trivially) stating that the empirical spectral distribution $\frac{1}{d}\sum_{i=1}^{d}\delta_{\lambda_i}$ of a random matrix A is exactly supported on its spectrum $\{\lambda_1,\ldots,\lambda_d\}$.

2.3. **Semi-Circular Elements & Catalan Numbers.** Recall that in classic probability, the most important distribution is the Normal distribution, as it is the limiting distribution in the CLT. In free probability, the distribution that plays this role is that of the semi-circle:

Definition 2.3.1. Let $x \in \mathcal{A}$ be self-adjoint. If the $*$-distribution of x is given by $\frac{2}{\pi r^2}\sqrt{r^2 - t^2}$ for $t \in [-r, r]$, then x is a **semi-circular element** of radius r.

We often work with the semi-circle of radius 2, since that is the one with zero mean and unit variance. Thus, if the radius is not specified, assume it is $r = 2$. In general, the variance $\mathrm{var}(x) := \varphi(x^2) - \varphi(x)^2$ of a semi-circular element is given by $r^2/4$. It turns out that the moments of a semi-circular element have a very nice form:

$$\varphi(x^k) = \begin{cases} 0, & k \text{ is odd} \\ C_{k/2}, & k \text{ is even,} \end{cases}$$

where $C_p := \frac{1}{p+1}\binom{2p}{p}$ is the p^{th} **Catalan number**. This can be proven by calculating the integral given by the $*$-distribution μ of x:

$$\varphi(x^k) = \frac{1}{2\pi}\int_{-2}^{2} t^k \sqrt{4 - t^2}\,\mathrm{d}t,$$

and we can notice that if the variance is σ^2 instead of 1, the even moments will be $\sigma^k C_{k/2}$. It turns out that the semi-circular distribution is completely determined by its moments, and this fact will be vital for proving the Free CLT later, as we will show that these are exactly the limiting moments of our $\sqrt{N}$-scaled sample average.

3. Free Independence

Independence is an important assumption in statistics; in many cases, it allows us to greatly simplify expressions that arise in the calculation of important statistics or limiting distributions. Recall that for two classical random variables X and Y, the standard notion of independence states that

$$\mathbb{E}[f(X)] = \mathbb{E}[g(Y)] = 0 \implies \mathbb{E}[f(X)g(Y)] = 0.$$

It turns out that a similar notion holds for non-commutative random variables, so long as we account for the fact that $f(X)g(Y)$ and $g(Y)f(X)$ might not be equal:

Definition 3.1. Let (A, φ) be a C*-probability space. Then $(A_i)_{i\in\mathcal{I}}$ are **freely independent** (**free**) if $\varphi(a_1 \cdots a_k) = 0$ whenever $a_j \in \mathcal{A}_{i_j}$ with $\varphi(a_j) = 0$ for all $1 \le j \le k$, and no neighboring indices are from the same subalgebra: $i_1 \neq i_2$, $i_2 \neq i_3$, …, $i_{k-1} \neq i_k$.

We can also talk about free independence of a collection of random variables: $(a_i)_{i\in\mathcal{I}}$ are free if the generated sub-algebras $\big(\sigma(1,a_i)\big)_{i\in\mathcal{I}}$ are free. We also say that $(a_i)_{i\in\mathcal{I}}$ are $*$-free if $\big(\sigma(1,a_i,a_i^*)\big)_{i\in\mathcal{I}}$ are free. We will sometimes write $a \perp\!\!\!\perp^f b$ to say a and b are freely independent. Now, we are able to relate this to freeness of subgroups:

Definition 3.2. Let G be a group. Then subgroups $(G_i)_{i\in\mathcal{I}}$ are (**torsion**) **free** if for every $k \geq 1$ and $i_1,\dots,i_k \in \mathcal{I}$ all distinct, we have

$$g_1 \in G_{i_1} \setminus \{e\}, \dots, g_k \in G_{i_k} \setminus \{e\} \implies g_1 \cdots g_k \neq e.$$

Essentially, it means that there is no non-trivial relationship among the subgroups. Recall that in Example 2.1.4, we considered the non-commutative probability space $(\mathbb{C}G, \tau_G)$. It turns out that this was exactly Voiculescu's initial reasoning for calling it "free probability" in the first place, as can be understood through the following connection:

Proposition 3.3. Let G be a group. Then the subgroups $(G_i)_{i\in\mathcal{I}}$ are free in G if and only if the subalgebras $(\mathbb{C}G_i)_{i\in\mathcal{I}}$ are freely independent in $(\mathbb{C}G, \tau_G)$.

The above proposition should make intuitive sense: if no non-trivial relations exist among the elements of these sets G_i, then if each $\sum_{g\in G_i} \alpha_g g$ has $\alpha_e = 0$, the same will be true of their product.

4. The Free Central Limit Theorem

Once again, we can recall that in the classical CLT, we are given a sequence of i.i.d. centered random variables $X_1, X_2, \dots$, and we wish to understand the distribution of

$$\frac{X_1 + \ldots + X_N}{\sqrt{N}} \text{ as } N \to \infty.$$

Well, we can do the same thing for free random variables! First we need to understand convergence in distribution in the non-commutative case though:

Definition 4.1. Let $(\mathcal{A}_N, \varphi_n)$ with $a_N \in \mathcal{A}_N$, $N \geq 1$, and $(\mathcal{A}, \varphi)$ with $a \in \mathcal{A}$, be non-commutative probability spaces. Then we say that $a_N \in \mathcal{A}_n$ **converges in distribution** to a ($a_N \xrightarrow{d} a$) if for every $n \geq 1$, $\varphi_N(a_N^n) \to \varphi(a^n)$.

We see that this is essentially saying we need every moment to converge. We will understand these terms in the proof of our following main theorem:

Theorem 4.2 (Free Central Limit Theorem)**.** Let $(\mathcal{A}, \varphi)$ be a C*-probability space and $a_1, a_2, \ldots \in \mathcal{A}$ be free, identically distributed, and self-adjoint with common mean $\varphi(a_i) = 0$

and variance $\varphi(a_i^2) = \sigma^2$. Then

$$\frac{a_1 + \ldots + a_N}{\sqrt{N}} \xrightarrow{d} \mu_s,$$

where μ_s is the distribution of a semi-circular element of radius 2σ.

Proof. In the course lecture notes [5], a proof using R-transforms is given. Instead, we will work from first principles, so let us now consider moments of the form $\varphi\left((a_1 + \ldots + a_N)^n\right)$. The first thing we can notice is that by our assumptions, the trace of certain products depends only on whether or not certain indices are the same or different; for example,

$$\varphi(a_1^2 a_2) = \varphi(a_3^2 a_7) = \varphi(a_{10}^2 a_4).$$

Thus, our moment can be expanded as

$$\varphi\left((a_1 + \ldots + a_N)^n\right) = \sum_{\pi \in P_n} k_\pi A_\pi^N,$$

where

(1) $\pi \in \mathcal{P}[n]$ is a partition of $[n] = \{1, 2, \ldots, n\}$.
(2) k_π is the common value of $\varphi(a_{r_1} \cdots a_{r_n})$, in which $r_i = r_j$ if and only if $i \sim_\pi j$ (they are in the same block within the partition π).
(3) A_π^N is the number of ways to choose all of our $1 \leq r_i \leq N$ for $i \leq n$ such that $\varphi(a_{r_1} \cdots a_{r_n}) = k_\pi$.

In terms of the example given above, if we take $n = N = 3$, then we have

$$k_{\{(1,2),(3)\}} = \varphi(a_1^2 a_2), \quad A_\pi^3 = 6 = \left|\{a_1^2 a_2, a_1^2 a_3, a_2^2 a_1, a_2^2 a_3, a_3^2 a_1, a_3^2 a_2\}\right|.$$

Now, we can make the following realization: if the partition π has a singleton block, then $k_\pi = 0$, since free independence tells us that

$$\varphi(a_1 b a_2) = \varphi(a_1 a_2)\varphi(b) \text{ whenever } \{a_1, a_2\} \perp\!\!\!\perp^f b,$$

and thus when we pull out this singleton term which is centered, we get zero. Thus, only partitions π having pairs or larger blocks will contribute. However, it turns out we will need exactly *all* pairs! To see this, note that by a simple combinatorial argument,

$$A_\pi^N = \frac{N!}{(N - |\pi|)!} \approx N^{|\pi|},$$

where $|\pi|$ is the number of blocks. Thus, when we throw in the square root of N to our moment, we obtain

$$\lim_{N\to\infty} \varphi\left(\left(\frac{a_1 + \ldots + a_N}{\sqrt{N}}\right)^n\right) = \lim_{N\to\infty} \sum_{\pi \in \mathcal{P}[n]} k_\pi A_\pi^N N^{-n/2} = \lim_{N\to\infty} \sum_{\pi \in \mathcal{P}[n]} N^{|\pi| - n/2} k_\pi.$$

We notice that this limit is non-zero if and only if $|\pi| \geq n/2$, since otherwise we have a limit of $1/N^a$ for $a > 0$, which is zero. However, since $|\pi|$ is bounded above by $n/2$ (since it has no singletons), we conclude that π must exactly be a pairing (every block is of size two), thus

$$\lim_{N\to\infty} \varphi\left(\left(\frac{a_1 + \ldots + a_N}{\sqrt{N}}\right)^n\right) = \sum_{\pi \in \mathcal{P}_2[n]} k_\pi,$$

where $\mathcal{P}_2[n]$ is all pairings of $[n]$. This immediately tells us that this limit is zero for n odd, since there is no way to pair an odd number of things. So, assume $n = 2k$ for some k. We need to determine which pairings still do not contribute. Recall that by the definition of free independence, if our pairing does not pair any neighbors (i.e., there is no pair of the form $\{r, r+1\} \in \pi$), then each consecutive index must be distinct, meaning k_π is also zero. Thus, we must have a pair of neighbors, giving us that

$$k_\pi = \varphi(a_{r_1} \cdots a_{r_j}^2 \cdots a_{r_n}) = \sigma^2 \varphi(a_{r_1} \cdots a_{r_{j-1}} a_{r_{j+1}} \cdots a_{r_n}).$$

Clearly, we can continue this with the right-hand term, and at each stage we either obtain another σ^2 out front, or the entire thing becomes zero if there is no neighbor pair. It turns out that such a pairing is called a **non-crossing pairing**, i.e. there are no $a < b < c < d$ such that (a, c) and (b, d) are pairs. We conclude that when n is even,

$$\lim_{N\to\infty} \varphi\left(\left(\frac{a_1 + \ldots + a_N}{\sqrt{N}}\right)^n\right) = S_n \sigma^n,$$

where $S_n = \left|\mathrm{NC}_2[n]\right|$ is the number of non-crossing pairings of $[n]$. From what we know in class, S_{2k} is exactly the k^{th} Catalan number! Since our limiting moments are exactly those of the semi-circular distribution with radius 2σ, and this distribution is uniquely determined by its moments, we are done. $\square$

One main takeaway from this proof is the following connection: going from classical probability to free probability means going from all pairings to non-crossing pairings. If we wanted to calculate the moments above for classically independent random variables, we would have simply gotten $\sigma^n(n-1)!!$ for the even moments, which coincides with the Normal distribution. In the non-commutative case, we were forced to consider only pairings which are non-crossing. There is also a Free Law of Large Numbers (whose proof will not be written down here), but essentially boils down to the same idea from the classical case, this time with a condition on the distribution:

Theorem 4.3 (Free Law of Large Numbers)**.** Let $(a_n)_{n\geq 1}$ be freely independent with the same distribution μ. Then the following are equivalent:

(1) There exist constants $(M_n)_{n\in\mathbb{N}}$ such that $\bar{X}_n - M_n \xrightarrow{d} \delta_0$.
(2) $\lim_{t\to\infty} t\mu\left(\{x : |x| > t\}\right) = 0$.

Of course, in the case that (2) holds, we can simply choose $M_n = \int_{-n}^{n} t d\mu(t)$. This second condition is essentially stating that our distribution puts measure outside the ball $B_t(0)$ at a decreasing rate faster than $1/t$, which is of course true for common distributions such as sub-Gaussians or sub-exponentials. One also usually requires in the above Theorem 4.3 that φ be **normal**, i.e. $a_n \uparrow a \implies \varphi(a) = \sup_n \varphi(a_n)$ (I know, probabilists have a total overloading of the word normal). One can check [6] for more similar results on limit theorems in free probability, with regards to both sums and products of free random variables.

5. Free CLT in Dependent and Other Cases

Of course, one of the crucial assumptions in the Free CLT is that the random variables are, well, free. This is not always desirable in practice, since many real-word examples might not meet this criterion of independence, rendering us unable to use the proof technique above to study such a limiting distribution. One example is if each a_n is a random matrix whose entries are dependent on those of a_{n-1}, such as a matrix of growing dimension $n \times n$, in which at each stage we simply append a new row and column. In this case, $a_{n+\varepsilon} \mid a_n$ only consists of $\varepsilon^2 + 2\varepsilon n$ new entries.

Furthermore, the Free CLT only considers empirical averages, of which for example the popular U-statistics

$$\frac{1}{n^2} \sum_{i,j=1}^{n} f(a_i, a_j),$$

and many other statistics, are not. This is not a problem for the classical CLT, in which the assumptions of independence, taking empirical averages, and even identical distributions have all been relaxed to varying degrees of generality.

In [1], Austern considers under what conditions a free CLT can be extended to both empirical averages and U-statistics of dynamical systems, such as stationary or quantum exchangeable sequences. Essentially, in the same way that the classical CLT is extended to the dependent case by imposing size/growth restraints on certain mixing conditions which quantify how far a sequence is from independence, we can do the same with "free mixing" coefficients in the non-commutative case. Recall that for a sequence of classical random variables, the strong-mixing coefficients are defined as

$$\alpha_i := \sup_{\substack{A \in \sigma(X_{-\infty:0}) \\ B \in \sigma(X_i:\infty)}} \left| P(A \cap B) - P(A)P(B) \right|,$$

where $\sigma(X_{-\infty:0})$ and $\sigma(X_i : \infty)$ denote the σ-algebra of events generated by $\ldots, X_{-1}, X_0$ or $X_i, X_{i+1}, \ldots$, respectively. If these coefficients decrease rapidly to 0 as $i \to \infty$, then the sequence is very weakly dependent. In re-defining such a quantity for non-commutative random variables, the paper is quite elaborate in its definition of such coefficients, so I will omit them for now. The important takeaway is that their main theorem (Theorem 2 of [1]) bounds the normed difference between a dynamical system and a semi-circular family using these mixing coefficients.

Similar results exist for extending the Free CLT to free convolutions of unbounded operators and a multivariate Free CLT, which can be found in [4] and [8], respectively.

6. Free Probability for Concentration Inequalities

Techniques from free probability can also be applied toward obtaining tighter concentration inequalities than those that are classically known. One common such concentration inequality is the **non-commutative Khintchine inequality**: if we have

$$X := \sum_{i=1}^{n} g_i A_i,$$

where each $g_i \sim \mathcal{N}(0,1)$ are independent and $A_i \in M_d(\mathbb{C})$ are fixed (and often self-adjoint), then the inequality states that

$$\sigma(X) \lesssim \mathbb{E}\|X\| \lesssim \sigma(X)\sqrt{\log(d)},$$

where $\sigma(X)^2 = \|\mathbb{E}(X^2)\| = \|A_1^2 + \ldots + A_n^2\|$. Thus, we can essentially compute the expected spectral norm up to this $\log(d)^{1/2}$ factor, solely based on the norm of our fixed matrices A_i. However, in high-dimensional regimes where $d \gg n$, the gap between the lower and upper bounds can become severe, rendering this inequality useless.

In [2], the authors show that free probability can accurately portray the spectral statistics of $\sum_i g_i A_i$ when our fixed matrices A_i are non-commutative. To this end, they define the standard model and its "free" counterpart like so:

$$X := A_0 + \sum_{i=1}^{n} g_i A_i, \qquad X_{\text{free}} := A_0 \otimes \mathbf{1} + \sum_{i=1}^{n} A_i \otimes s_i,$$

where $s_1, \ldots, s_n$ are free semi-circular elements. Then it turns out that the spectrum of X is close to that of X_{free}:

Theorem 6.1 (Theorem 2.1 of [2]). For the above model with $A_0, \ldots, A_n$ all self-adjoint, we have that for every $t \geq 0$,

$$P\left(\text{spec}(X) \subseteq \text{spec}(X_{\text{free}}) \pm c_t\right) \geq 1 - e^{-t^2},$$

where $c_t := C\left(\tilde{v}(X)\log(d)^{3/4} + \sigma_*(X)t\right)$ (for some universal constant $C > 0$) quantifies the non-commutativity of the matrices A_i. In the case of non-self-adjoint coefficient matrices, the authors still obtain a probabilistic lower bound of the form

$$P\left(\|X\| > \|X_{\text{free}}\| + c_t\right) \leq e^{-t^2}$$

and a sure upper bound

$$\mathbb{E}\|X\| \leq \|X_{\text{free}}\| + C\tilde{v}(X)\log(d)^{3/4}.$$

Of course, in order for these bounds to be useful, $\|X_{\mathbf{free}}\|$ must be readily computable in practice, in which case the following lemma can help us:

Lemma 6.2. When the A_i are self-adjoint, we have

$$\|X_{\mathbf{free}}\| = \max_{\varepsilon=\pm 1} \inf_{Z>0} \lambda_{\max}\left(Z^{-1} + \varepsilon A_0 + \sum_{i=1}^{n} A_i Z A_i\right),$$

where this infimum is over all positive-definite, self-adjoint $Z \in M_d(\mathbb{C})$, and $\lambda_{\max}(\cdot)$ is the largest eigenvalue. In the proofs of their results, the authors use techniques described above for the Free CLT proof, such as the difference between summing over all pairings and summing over only non-crossing pairings. Essentially, they are able to show that even though the moments of the matrix X might depend on all pairings, the crossing pairings still come close to vanishing in many cases via the non-commutativity of the A_i. This just goes to show the deep connection between free probability and random matrix theory.

7. Conclusion

Overall, we have seen that free probability is a complex and fascinating field of active research. Many results in the non-commutative case mirror those in the classical setting, with subtle changes such as partitions becoming non-crossing partitions or Normal limits becoming semi-circular. The ideas of free probability have come a long way from Voiculescu's first papers on the topic in the late 1980s, finding applications in areas such as random matrix theory and concentration inequalities as seen above. Current active research in free probability aims to extend results such as the free CLT to dependent cases and high-dimensional regimes, and even attempt to construct new invariants of von Neumann algebras. I hope this was a good introduction to the field and you are excited to continue learning more about it!

References

[1] Austern, M., 2020. A free central-limit theorem for dynamical systems. *arXiv preprint arXiv:2005.10923.*

[2] Bandeira, A.S., Boedihardjo, M.T. and van Handel, R., 2023. Matrix concentration inequalities and free probability. *Inventiones mathematicae*, pp.1-69.

[3] Bercovici, H. and Pata, V., 1996. The law of large numbers for free identically distributed random variables. *The Annals of Probabilit y, 24*(1), pp.453-465.

[4] Bercovici, H. and Voiculescu, D., 1993. Free convolution of measures with unbounded support. *Indiana University Mathematics Journal, 42*(3), pp.733-773.

[5] Edelman, A., 2022. *Random matrix theory.*

[6] Kargin, V., 2008. *Limit theorems in free probability theory.* New York University.

[7] Nica, A. and Speicher, R., 2006. *Lectures on the combinatorics of free probability* (Vol. 13). Cambridge University Press.

[8] Speicher, R., 1990. A new example of 'independence'and 'white noise'. *Probability theory and related fields, 84*(2), pp.141-159.

[9] Tao, T., 2023. *Topics in random matrix theory* (Vol. 132). American Mathematical Society.

[10] Voiculescu, D., 1986. Addition of certain non-commuting random variables. *Journal of functional analysis, 66*(3), pp.323-346.