UNIVERSALITY OF FREE RANDOM VARIABLES: ATOMS FOR NON-COMMUTATIVE RATIONAL FUNCTIONS

OCTAVIO ARIZMENDI, GUILLAUME CÉBRON, ROLAND SPEICHER, AND SHENG YIN

ABSTRACT. Consider a tuple $(Y_1, \ldots, Y_d)$ of normal operators in a tracial operator algebra setting with prescribed sizes of the eigenspaces for each $Y_i$. We address the question what one can say about the sizes of the eigenspaces for any non-commutative polynomial $P(Y_1, \ldots, Y_d)$ in those operators? We show that for each polynomial $P$ there are unavoidable eigenspaces, which occur in $P(Y_1, \ldots, Y_d)$ for any $(Y_1, \ldots, Y_d)$ with the prescribed eigenspaces for the marginals. We will describe this minimal situation both in algebraic terms - where it is given by realizations via matrices over the free skew field and via rank calculations - and in analytic terms - where it is given by freely independent random variables with prescribed atoms in their distributions. The fact that the latter situation corresponds to this minimal situation allows to draw many new conclusions about atoms in polynomials of free variables. In particular, we give a complete description of atoms in the free commutator and the free anti-commutator. Furthermore, our results do not only apply to polynomials, but much more general also to non-commutative rational functions. Since many random matrix models become asymptotically free in the large $N$ limit, our results allow us to calculate the location and size of atoms in the asymptotic eigenvalue distribution of polynomials and rational functions in randomly rotated matrices.

1. INTRODUCTION

What can we say about the distribution of a non-commutative selfadjoint polynomial $P$ applied to selfadjoint operators $Y_1, \ldots, Y_d$, if we fix the marginal distributions for all the $Y_i$? Of course, there can be many possibilities for those distributions of $P(Y_1, \ldots, Y_d)$, depending on the actual choice of $(Y_1, \ldots, Y_d)$; but it seems that there are some generic features (depending on $P$ and on the marginal distributions) which are common to all realizations. Progress in free probability theory in the last years supports the idea that actually the specific case where we choose all the variables to be freely independent realizes those generic features.

At the moment, questions addressing densities and their smoothness properties of the distributions are out of reach; however, for dealing with atoms there has been quite some

O.A. was supported by Conacyt grant A1- S-9764.
G.C. and S.Y. were partly supported by the Project MESA (ANR-18-CE40-006) and by the Project STARS (ANR-20-CE40-0008) of the French National Research Agency (ANR).
O.A. and R.S. gratefully acknowledge financial support by the SFB-TRR 195 “Symbolic Tools in Mathematics and their Application” of the German Research Foundation (DFG).
S.Y acknowledges with appreciation the support by the National Natural Science Foundation of China (Grant No. 12031004) for his stay at Harbin.

We thank Serban Belinschi, Hari Bercovici, Marek Bożejko, Benoît Collins, Jorge Garza-Vargas, and Dan-Virgil Voiculescu for constructive feedback on an earlier version of this work.

1
progress in the last couple of years. In particular, it was shown that in the case of free variables $X_1, \ldots, X_d$ atoms cannot be created via polynomials; i.e., if each $X_i$ has no atom in its distribution, then for non-constant polynomials $P$, $P(X_1, \ldots, X_d)$ also has no atoms. We will now push all this much further to the general situation where we allow arbitrary atoms for the individual variables. In such a situation the distribution of $P(Y_1, \ldots, Y_d)$ will in general also have atoms. Depending on $P$ and on the atoms for the marginal distributions, there will actually be unavoidable locations and sizes for atoms in the evaluations: for any operator tuple $(Y_1, \ldots, Y_d)$ with the prescribed atoms for the marginals, the distribution of $P(Y_1, \ldots, Y_d)$ will have atoms at those unavoidable locations with the mass of such an atom being at least of the unavoidable size. One of our main results is that in the case of free variables $X_1, \ldots, X_d$, the atoms of $P(X_1, \ldots, X_d)$ are exactly at those unavoidable locations with masses given exactly by the unavoidable sizes. We emphasize that the unavoidable locations and sizes do not depend on any information about the non-atomic parts in the marginal distributions.

One should note that many random matrix ensembles become asymptotically free in the large $N$ limit (for example, independent ensembles whose eigenspaces are rotated randomly against each other). Thus our results say that, in the asymptotic regime, polynomials in randomly rotated matrices have eigenspaces of minimal size; and for many situations we are actually able to calculate the size of those eigenspaces explicitly.

1.1. Main results. Let us now present the precise statements of our main results. For undefined notions the reader is referred to Section 2. In particular, $\mu_X$ denotes the analytic distribution of the selfadjoint (or more general, normal) operator $X$ (see Def. 2.2). The atomic part of this is denoted by $\mu^p_X$ (see Def. 2.4). Our results do not only hold for non-commutative polynomials, but much more general also for non-commutative rational expressions in the variables. The basics about non-commutative rational expressions and functions will be recalled in Subsection 2.4. Note that our input operators are usually normal and bounded, but we do not require this for the outcome after evaluating non-commutative polynomials or rational functions in them. In the general case, the size of an “atom” is replaced by the general notion of the rank of an operator; this is defined in Def. 2.26. For a normal operator $Y$ and a complex number $\lambda$, those concepts are related by $\mu^p_Y(\{\lambda\}) = \mu^p_Y(\{\lambda\}) = 1 - \text{rank}(\lambda - Y)$. We will use the notation $\mu^p_Y$ also in the general, not necessarily normal, case.

**Theorem 1.1.** Let $X_1, \ldots, X_d$ and $Y_1, \ldots, Y_d$ be normal variables in tracial $W^*$-probability spaces with $X_1, \ldots, X_d$ being $*$-free and such that, for each $1 \leq i \leq d$, we have

$$\mu^p_{X_i} \leq \mu^p_{Y_i}.$$  

Then, for each rational expression $R$ in $d$ non-commuting variables, $R(X_1, \ldots, X_d)$ is well-defined whenever $R(Y_1, \ldots, Y_d)$ is well-defined, and we have

$$\text{rank}(R(X_1, \ldots, X_d)) \geq \text{rank}(R(Y_1, \ldots, Y_d));$$

in other words, for each such rational expression $R$ in $d$ non-commuting variables,

$$\mu^p_{R(X_1, \ldots, X_d)} \leq \mu^p_{R(Y_1, \ldots, Y_d)}.$$
Our theorem includes in particular the statement that the location and the size of the unavoidable atoms do not depend on the non-atomic parts in the marginal distributions of the free tuple. We put this down as a separate theorem, as this will be a main step in the proof of the general statement.

**Theorem 1.2.** Let \((X_1, \ldots, X_d)\) and \((Y_1, \ldots, Y_d)\) be two \(d\)-tuples of \(\ast\)-free normal variables in tracial \(W^\ast\)-probability spaces such that, for each \(1 \leq i \leq d\), we have

\[
\mu_{X_i}^p = \mu_{Y_i}^p.
\]

Then, for each rational expression \(R\) in \(d\) non-commuting variables, \(R(X_1, \ldots, X_d)\) is well-defined if and only if \(R(Y_1, \ldots, Y_d)\) is well-defined, and we have

\[
\text{rank}(R(X_1, \ldots, X_d)) = \text{rank}(R(Y_1, \ldots, Y_d));
\]

in other words, for each such rational expression \(R\) in \(d\) non-commuting variables,

\[
\mu_{R(X_1, \ldots, X_d)}^p = \mu_{R(Y_1, \ldots, Y_d)}^p.
\]

Note that this means that for determining the atoms in \(R(X_1, \ldots, X_d)\), for \(X_1, \ldots, X_d\) being \(\ast\)-free, we can model the non-discrete parts in \(\mu_{X_i}\) by any non-atomic distribution; choosing semicircular distributions for this seems to be a good option.

Another consequence of Theorem 1.2 is that it allows us to define, for each non-commutative rational function \(R\), a corresponding rational convolution \(R^\square\) on discrete sub-probability measures on \(\mathbb{C}\), according to

\[
R^\square(\mu_{X_1}^p, \ldots, \mu_{X_d}^p) = \mu_{R(X_1, \ldots, X_d)}^p.
\]

We want to point out that our rational expressions are only allowed to contain the operators, but not their adjoints. The following example shows that an extension of Theorem 1.1 to the \(\ast\)-case cannot hold without additional assumptions. Consider, for \(d = 1\), the \(\ast\)-polynomial \(R(X) = XX^\ast\) and take for \(X\) a Haar unitary variable \(U\). Then the distribution of the normal variable \(U\) is the uniform distribution on the unit circle in \(\mathbb{C}\) and thus has no atoms; hence the condition \(\mu_U^p \leq \mu_Y^p\) is void and we can take for \(Y\) any normal operator; in particular, a normal operator for which \(R(Y) = YY^\ast\) has no atom. On the other hand the distribution of \(R(U) = UU^\ast = 1\) has an atom of weight 1 at \(\lambda = 1\), thus violating \(\mu_U^p \leq \mu_Y^p\). We expect that under stronger assumptions a version of Theorem 1.1 still holds for \(\ast\)-rational functions. We conjecture this to be the case at least when we assume that, for each \(i = 1, \ldots, d\), the \(\ast\)-distribution of \(X_i\) is the same as the \(\ast\)-distribution of \(Y_i\). This will be addressed in forthcoming investigations.

The proof of our general results will proceed via modeling the situation of given marginal distributions, in the case of rational weights for the atoms, by a purely algebraic object, namely by matrices over the free field. The free field is the universal field of fractions of non-commutative polynomials. Those matrices over the free field contain thus all algebraic information common to all realizations by operator tuples \((Y_1, \ldots, Y_d)\) with the given atoms in their marginals. An unavoidable atom at 0 in the distribution of \(P(Y_1, \ldots, Y_d)\) corresponds in this model to the non-invertibility of the matrices. By results of [MSY19], the special operator tuples \((X_1, \ldots, X_d)\), where \(X_1, \ldots, X_d\) are free, yield an analytic model for this algebraic situation. The inner rank on the algebraic side is then the same as the von Neumann operator algebraic rank on the analytic side, which in turn corresponds to the size of the atom in \(P(X_1, \ldots, X_d)\) at 0. (By shifting \(P\) by a constant, we can of course move the location of any atom to 0.) This shows then that, in the case of
rational weights for the input atoms, the unavoidable atoms are realized, both in locations and sizes, exactly in the free situation. General approximation arguments allow to extend this to the general situation of arbitrary weights for the atoms.

Moreover, according to [MSY19 Theorem 5.24], we know that the von Neumann rank is equal to the inner rank over some skew field whenever an operator tuple \((X_1, \ldots, X_d)\) satisfies the strong Atiyah property. So a natural question is that to what extend the von Neumann rank over an operator tuple \((X_1, \ldots, X_d)\) can be identified with an algebraic rank such as inner rank. In the case of rational weights, our result (Theorem 4.1) indeed confirms that von Neumann rank over such tuples is equal to some normalized inner rank. This rank equality explains the rationality of the von Neumann rank for the free tuples of operators with atoms of rational weights.

We want to emphasize that the restriction to a tracial setting for our operator tuples \((Y_1, \ldots, Y_d)\) is crucial for the above arguments. Since the evaluation map from the free field to algebras is only well-defined for stably finite algebras – i.e., in our setting for finite von Neumann algebras – the above modeling with the free field captures only such situations. It is of course an interesting question whether our results extend also to non-tracial situations. In order to address this, new insights are necessary and we will leave this for future work.

The insight that the atoms in the free situation correspond exactly to the unavoidable atoms leads to some very precise statements on atoms for selfadjoint polynomials in free selfadjoint variables \(X_1, \ldots, X_d\). If we have one example of an operator tuple \((Y_1, \ldots, Y_d)\) for which we can calculate the size of an atom for \(P(Y_1, \ldots, Y_n)\), then this size is a bound from above for the size of the atom for \(P(X_1, \ldots, X_d)\). In particular, if we find one realization \((Y_1, \ldots, Y_d)\) of the marginal distributions for which we don’t have an atom at \(\lambda\) for \(P(Y_1, \ldots, Y_d)\) then also \(P(X_1, \ldots, X_n)\) cannot have an atom at \(\lambda\). This has as direct consequence that the only possible locations for atoms of \(P(X_1, \ldots, X_d)\) are at the positions \(P(\lambda_1, \ldots, \lambda_d)\) where, for \(i = 1, \ldots, d\), \(\lambda_i\) is the location of an atom for \(X_i\).

For getting lower bounds on the size of atoms for \(P(X_1, \ldots, X_d)\), we need apriori a bound from below for the size of an atom for \(P(Y_1, \ldots, Y_d)\) when considering all possibilities for \(Y_1, \ldots, Y_d\) with the prescribed atoms for the marginals. This usually has to rely on general arguments about the size of the intersection of the involved eigenspaces for the atoms. It is a bit surprising that in many cases we can match those general lower bounds with the upper bounds from special realizations and thus get the exact size of the atoms in the free situation.

The following theorem gives a general condition for a polynomial \(P\) in two variables ensuring that we can calculate exactly the size of any atom of \(P\).

**Theorem 1.3.** Let \(X\) and \(Y\) be free selfadjoint random variables and let \(a \in \mathbb{R}\). Let \(P\) be a selfadjoint polynomial which satisfies the following: for all \(\lambda, \rho\) such that \(P(\lambda, \rho) = a\) then \(P(\lambda, \rho) \neq a\) and \(P(\lambda, \rho) \neq a\) whenever \(\lambda \neq \tilde{\lambda}\) and \(\rho \neq \tilde{\rho}\).

Then \(P(X, Y)\) has an atom at \(a\) if and only there are \(\lambda\) and \(\rho\) such that \(X\) has an atom at \(\lambda\) of size \(s\) and \(Y\) has an atom at \(\rho\) of size \(t\), such that \(r = t + s - 1 > 0\) and \(P(\lambda, \rho) = a\). Furthermore, if \(r > 0\), and \(s(a)\) and \(t(a)\) denote the mass of these (unique) atoms, then the mass at \(a\) is given by \(s(a) + t(a) - 1\).
In particular, this theorem allows a description of the atoms of the free additive convolution and, when \( a \neq 0 \), also for the free multiplicative convolution, the free commutator and the free anti-commutator. With some additional work we can achieve a full description of the atoms for all these cases.

### 1.2. How to model free variables.

Here we want to give the main ideas – mostly in terms of a concrete example – involved in the proof of our results. To be definite, we will in the following talk about atoms of distributions of (normal) operators. Of course, all of them have to live in a tracial von Neumann algebra setting.

A universal algebraic model for an operator \( X \) without atoms in its distribution is given by a formal variable \( x \). Rational functions of \( x \) model then the corresponding rational function of any such operator. In the case of several operators the free (skew) field, generated by \( d \) formal non-commutative variables \( x_1, \ldots, x_d \), is the universal algebraic model for any \( d \)-tuple of, not necessarily commuting, operators without atoms in their distributions.

If we want models for operators which have atoms in their distributions, then we can model this via matrices over the free field. For example, if we want to model an operator \( X \) with distribution \( \frac{1}{3} \delta_1 + \frac{1}{3} \delta_2 + \frac{1}{3} \tilde{\mu} \), where \( \tilde{\mu} \) is a probability measure without atoms, then the following \( 3 \times 3 \)-matrix over the free field (generated by the formal variable \( x_1 \)) will do the job:

\[
A = \begin{pmatrix}
1 & 0 & 0 \\
0 & 2 & 0 \\
0 & 0 & x_1
\end{pmatrix}.
\]

The weight for an atom at \( \lambda \) corresponds in this algebraic setting to the inner rank (a purely algebraic notion); \( \text{rank}(A - \lambda) = k \) corresponds to an atom at \( \lambda \) of weight \( 1 - k/3 \). In this way, it is clear that with \( n \times n \)-matrices over the free field we can thus model only atoms whose weights are multiples of \( 1/n \).

Now assume that we want to model two operators \( X \) and \( Y \) with prescribed distributions of their atoms. For example, in addition to having an operator \( X \) with distribution \( \frac{1}{3} \delta_1 + \frac{1}{3} \delta_2 + \frac{1}{3} \tilde{\mu} \) as before we would also like to have another operator \( Y \) with only one atom, of size \( 1/3 \) at 0. This latter operator could be modeled by

\[
\tilde{B} = \begin{pmatrix}
0 & 0 & 0 \\
0 & x_2 & 0 \\
0 & 0 & x_3
\end{pmatrix}.
\]

Even though we have used different formal variables to model this one, it is clear that \( A \) and \( \tilde{B} \), living in \( 3 \times 3 \)-matrices over the free field generated by the variables \( x_1, x_2, x_3 \) are not universal models for pairs of operators with the prescribed atoms in their distributions. What we have modeled here are operators where there exist definite relations between their eigenspaces. In order to get rid of those specific relations we should rotate the two matrices against each other, again in a generic way. For this we replace \( \tilde{B} \) by the matrix

\[
B = U \tilde{B} U^{-1} = \begin{pmatrix}
u_{11} & u_{12} & u_{13} \\
u_{21} & u_{22} & u_{23} \\
u_{31} & u_{32} & u_{33}
\end{pmatrix} \begin{pmatrix}0 & 0 & 0 \\
0 & x_2 & 0 \\
0 & 0 & x_3
\end{pmatrix} \begin{pmatrix}
u_{11} & u_{12} & u_{13} \\
u_{21} & u_{22} & u_{23} \\
u_{31} & u_{32} & u_{33}
\end{pmatrix}^{-1}.
\]
The matrix $U$ has here 9 other formal variables as entries, and the matrix $U^{-1}$ is calculated over the free field generated by $u_{11}, \ldots, u_{33}$, thus its entries are rational functions in the $u_{11}, \ldots, u_{33}$; note that the matrix $U$ is invertible over the free field. The matrices $A$ and $B$, living in matrices over the free field generated by the 12 variables $x_1, x_2, x_3, u_{11}, \ldots, u_{33}$, are our wanted universal algebraic model for two operators with the prescribed atomic distributions. Of course, we have to show that any concrete pair of such operators is a specialization of $A$ and $B$. This is essentially the diagonalization of the given operators with respect to the projections onto their eigenspaces, but since we have to stay within our tracial von Neumann algebra setting and also to deal with not necessarily rational weights of the atoms the technical details are a bit more involved.

If we are now interested in the possible atoms appearing in rational functions $R(X, Y)$ for arbitrary operators $X$ and $Y$ with the prescribed weights for their atoms, then doing this calculation for $R(A, B)$ over the free field will yield exactly the unavoidable atoms and weights which are common to all $R(X, Y)$. In particular, in order that $R(X, Y)$ is well-defined (as an unbounded operator), $R(A, B)$ must be well-defined over the free field. Clearly, not all rational functions make sense when we have atoms; for example, $A^{-1}$ makes sense in the free field, but $B^{-1}$ does not. Of course, there are usually choices for $X$ and $Y$ for which $R(X, Y)$ is not defined, even though $R(A, B)$ makes sense.

From the above representation as $3 \times 3$-matrices over the free field, it is also clear that in this case, where the marginal atoms are of weight $1/3$, the atoms of any meaningful rational function in $A$ and $B$ can also have only weights which are multiples of $1/3$. This will in general make a relation to the so-called Atiyah property.

Up to this point we have given universal algebraic models for operator tuples with prescribed atoms in the marginals, relying on the fact that the free field yields via specialization any possible realisation for such situations. This reduction to the free field shows in particular that the non-atomic parts in the prescribed distributions do not come into the calculations. All non-atomic distributions are in the free field just represented by formal variables and thus do not influence the search for the unavoidable atoms in rational functions. We can, and actually will (see Proposition 3.2), even put part of the prescribed atoms into those formal variables which allows us to approximate general weights for atoms via rational weights in the matrices over the free field.

Doing concrete rank calculations over the free field, in order to find the unavoidable atoms, is however not easy – in particular, as we also have to deal with terms like $U^{-1}$. Note that the occurrence of $U^{-1}$ means that even if we are only interested in polynomials in our operators, in our calculations in the free field we have nevertheless to deal with rational functions in the $u_{ij}$.

We will now use another, much more surprising, connection between algebraic and analytic models. Namely, there are actually specific analytic situations which are not just a specialization, but which are isomorphic to the free field. This means that the rank calculations for the weights of unavoidable atoms in the free field can also be done in those analytic settings, using analytic tools. As those analytic settings correspond to free random variables, we have actually quite some tools, provided by free probability theory, for analytic calculations.
The starting point for this correspondence between the algebraic and the analytic side is the characterization from \( \text{[MSY19]} \) that the free field generated by formal variables \( x_1, \ldots, x_d \) is isomorphic to the division closure (in the affiliated unbounded operators) of operators \( X_1, \ldots, X_d \) if and only if those operators maximize a specific quantity \( \Delta(X_1, \ldots, X_d) \). Such a maximal \( \Delta \) is for example achieved if the \( X_1, \ldots, X_d \) are free and each \( X_i \) has no atoms in its distribution. Whereas \( x_1, x_2, x_3 \) in the above example can be chosen as such free variables (with some arbitrary distribution without atoms), we also have to deal with the matrix \( U \). We could just replace all entries of \( U \) by free operators with, say, semicircular distribution and then try to calculate the distribution of those matrices, using operator-valued free probability tools. However, this will get out of hand soon, in particular, for big sizes of the considered matrices and it is not clear whether this brings a clear advantage over rank calculations in the free field. What we do instead is to stay in the scalar-valued setting and make all relevant operators there free. For this we need that the matrix \( U \) itself has to be a Haar unitary matrix which also should be free from \( A \) and \( \tilde{B} \). This can be achieved by taking the entries \( u_{ij} \) as given by a free compression of a Haar unitary by matrix units. It takes then some effort to show that in this case \( \Delta(u_{11}, \ldots, u_{33}) \) is also maximal and that furthermore the \( u_{ij} \) are also free from \( x_1, x_2, x_3 \). Having proved all this gives then that \( \Delta(x_1, x_2, x_3, u_{11}, \ldots, u_{33}) \) is also maximal, and thus we can use this analytic setting to calculate the unavoidable atoms. As it is set up now, we have that \( U \) is a Haar unitary which is \( \ast \)-free from \( \tilde{B} \) and from \( A \). But this means that \( A \) and \( B \) themselves are free and each of them has the prescribed distribution for the atoms (and where we have the freedom to fill in for the non-atomic parts any non-atomic distribution, for example, semicirculars). So in the end we see that if we choose our operators to be free and with the prescribed distribution, then calculating atoms for any rational function in them will give the unavoidable atoms for the prescribed atoms of the marginals.

Note that we can now also turn this connection around and use it to get some information about atoms of polynomials or rational functions of free variables. In particular, we see from the fact that in the free field the non-atomic parts don’t play a role this is now also true for taking rational functions in free variables \( X_1, \ldots, X_d \). The atoms for such rational functions \( R(X_1, \ldots, R_d) \) depend only on the atoms of the operators \( X_i \).

Another advantage of knowing that the situation of free random variables corresponds to the “minimal” situation concerning atoms is that this allows to get results on atoms for polynomial free convolutions. By comparing the free situation with any other suitably chosen example, we can exclude atoms for the free case if they do not show up in the considered example.

1.3. **Organization of the article.** There are six more sections after the introduction and they are organised as follows. In Section 2 we introduce the necessary notions with some basic relevant facts used in the later sections.

In Section 3 we establish several important lemmas and propositions concerning noncommutative random variables. More precisely, it contains two subsections: in the first one we bring non-commutative random variables into some matrix forms that will be used for our main results in Section 4 and 5. In the second one, we prove that the free
compressions of free Haar unitaries have maximal $\Delta$ and thus we can identify them with formal variables.

In Section 4, we first establish algebraic models for $*$-free normal random variables with atoms of rational weights. Some consequences of our first main result are given in the last two subsections of this section. Subsection 4.2 concerns the link between our result and the strong Atiyah conjecture for groups and Atiyah properties for non-commutative variables. In Subsection 4.3 we give the proof of Theorem 1.2 with some applications.

In Section 5 we compare the rank and atoms for matrices in polynomials as well as rational functions between free variables and general random variables. The first subsection is devoted to the proof of Theorem 1.1. As a first consequence, the universality of the rational closure of free variables is stated in Subsection 5.2. Then we show, in Subsection 5.3, that (restricting to polynomials) the minimal situation realized by free variables actually can be obtained via matrices.

Section 6 collects applications and examples that can be studied with the help of our main results. Comparing to specific examples we obtain general criteria to bound the size of the atoms of various general polynomials. By using these methods, in Section 7, we give a complete description of atoms for the free commutator and the free anti-commutator and provide an application on the spectrum of universal covers.

2. Preliminaries

2.1. Tracial $W^*$-probability spaces, affiliated unbounded operators, analytic and point spectrum probabilistic notions. In this section we will define the basic non-commutative probabilistic notions and recall some relevant facts around them.

Definition 2.1. A tracial $W^*$-probability space $(\mathcal{M}, \tau)$ consists of a von Neumann algebra $\mathcal{M}$ and a faithful and normal trace $\tau : \mathcal{M} \to \mathbb{C}$. Elements in $\mathcal{M}$ are often addressed as (non-commutative) random variables.

In the setting of a tracial $W^*$-probability space $(\mathcal{M}, \tau)$, we will usually be interested in the closed and densely defined linear operators affiliated with $\mathcal{M}$. The set of all such affiliated unbounded operators is known to be a $*$-algebra containing $\mathcal{M}$. In a sense it is the non-commutative analogue of the set of all measurable functions of the elements from $\mathcal{M}$; correspondingly, we will denote it by $L^0(\mathcal{M})$. Moreover, the algebra $L^0(\mathcal{M})$ is stably finite (see, for example, [MSY19, Lemma 5.3]), which will allow us to consider the evaluations of rational functions in elements from $L^0(\mathcal{M})$.

Recall that an operator $X$ is called normal if we have $XX^* = X^*X$. Normal operators can be identified with classical (complex-valued) random variables, and thus we associate to each normal operator $X$ in a $W^*$-probability space $(\mathcal{M}, \tau)$ its distribution, which is a probability measure on $\mathbb{C}$ with compact support, and which is determined by the $*$-moments of $X$. In the tracial setting we can also allow affiliated unbounded operators; for those the appearing probability measures do not need to have compact support or even finite moments. We record these facts here as the following definition.

Definition 2.2. Let $(\mathcal{M}, \tau)$ be a tracial $W^*$-probability space.
For a normal operator $X \in \mathcal{M}$, the analytic distribution of $X$ with respect to $\tau$ is defined as the unique Borel probability measure $\mu_X$ on $\mathbb{C}$ that satisfies
\[
\int_{\mathbb{C}} z^k \overline{z}^l \, d\mu_X(z) = \tau(X^k (X^*)^l) \quad \text{for all } k, l \in \mathbb{N}_0.
\]

For a normal unbounded operator $Y \in L^0(\mathcal{M})$ its analytic distribution $\mu_Y$ is defined as the Borel probability measure on $\mathbb{C}$ that is obtained by $\mu_Y = \tau \circ E_Y$, where $E_Y$ denotes the resolution of identity associated to $Y$; this is well-defined, since $E_Y$ takes its values in $\mathcal{M}$ as $Y$ is a normal operator affiliated to $\mathcal{M}$; this extends the concept of analytic distributions for bounded normal operators from (1).

Whereas the analytic distribution only makes sense for normal operators, the atomic part of such a distribution has a much more algebraic meaning and can also be extended to general operators in our setting. This is made precise in the following definitions.

**Definition 2.3.** A discrete sub-probability measure on $\mathbb{C}$ is of the form
\[
\sum_{\lambda \in \mathbb{C}} \alpha_{\lambda} \delta_{\lambda},
\]
where we have that $0 \leq \alpha_{\lambda} \leq 1$ for all $\lambda$ and that $\sum_{\lambda \in \mathbb{C}} \alpha_{\lambda} \leq 1$. Hence in particular, only countably many $\alpha_{\lambda}$ are different from zero. The case that all $\alpha_{\lambda}$ are equal to zero is also allowed. The corresponding discrete sub-probability measure will then be denoted by $0$.

**Definition 2.4.** Let $(\mathcal{M}, \tau)$ be a tracial $W^*$-probability space. Let $X \in L^0(\mathcal{M})$ be given. Its point spectrum is defined as
\[
\sigma_p(X) := \{ \lambda \in \mathbb{C} : \ker(\lambda - X) \neq \{0\}\}.
\]
Furthermore, we define the point spectrum distribution of $X$ as the discrete sub-probability measure
\[
\mu^p_X := \sum_{\lambda \in \sigma_p(X)} \tau(p_{\ker(\lambda - X)}) \delta_{\lambda},
\]
where $p_{\ker(Y)}$ stands for the orthogonal projection onto the kernel of an operator $Y \in L^0(\mathcal{M})$.

Note that each $\lambda \in \sigma_p(X)$ is an eigenvalue of $X$, in the sense that there exists a nonzero eigenvector $\xi \in L^2(\mathcal{M})$ such that $X\xi = \lambda\xi$. Hence we will also call the subspace $\ker(\lambda - X)$ of $L^2(\mathcal{M})$ an eigenspace of $X$. For a given operator $X \in \mathcal{M}$ its point spectrum is the same as its spectrum in $L^0(\mathcal{M})$ since an operator is invertible in $L^0(\mathcal{M})$ if and only if it has trivial kernel; of course, the spectrum with respect to $\mathcal{M}$ is in general larger. Therefore, if two operators $X$ and $Y$ generate isomorphic von Neumann algebras, then $\mu^p_X = \mu^p_Y$. In particular, this is true for operators $X$ and $Y$ with the same $*$-distribution.

The weight $\tau(p_{\ker(\lambda - X)})$ for the eigenvalue $\lambda$ in the definition of $\mu^p_X$ is a measure, as a normalized kind of dimension function, for the size of the corresponding eigenspace; note that in the setting of type $II$ von Neumann algebras those eigenspaces are infinite-dimensional and their dimensions are measured with respect to the size of the whole Hilbert space. It is the basic ingredient in the setting of type $II$ von Neumann algebras.
that we have a continuous dimension, taking on all values in the interval $[0,1]$, for the size of infinite-dimensional subspaces of the ambient Hilbert space.

Moreover, if $X$ in $L^2(M)$ is normal, then $\lambda \in \sigma_p(X)$ if and only if $\lambda$ is an atom of the analytic distribution $\mu_X$, i.e., if $\mu_X(\{\lambda\}) > 0$; in which case we have $\mu_X(\{\lambda\}) = \tau(p_{\ker(\lambda-X)})$. Therefore, for a normal operator $X$, we will also say that $\lambda \in \sigma_p(X)$ is an atom of $X$. Clearly, for a normal $X$ we can decompose the measure $\mu_X$ as the sum of an atomic part (which is a discrete measure supported on $\sigma_p(X)$ and which coincides with the point spectrum distribution $\mu^p_X$ of $X$) and a non-atomic part (which is a measure which has no atoms).

If $X \in M$ is an operator which is not normal, a canonical probability measure $\mu_X$ called the Brown measure can still be constructed on the spectrum of $X$, generalizing the definition of the analytic distribution of normal variables. Thanks to [HS09], we always have the inequality $\mu^p_X \leq \mu_X$. In particular, each $\lambda \in \sigma_p(X)$ is an atom of the Brown measure $\mu_X$. However, this inequality can be strict, and the weight of the atom for the Brown measure can be larger than the weight in the point spectrum distribution.

2.2. Free random variables. The situation corresponding to unavoidable atoms will be given by free random variables. Free probability theory and its fundamental notion of freeness or free independence are by now well-established non-commutative analogues of classical probability theory and its notion of independence. We refer to the standard references [MS17, NS06, VDN92] for the basic notions and results from free probability theory. For our constructions only a few properties of free random variables will be important. We will not explicitly address random matrices in the following, but at least we want to point out that often random matrix models become freely independent in the large size limit, thus our results have quite some implications for atoms in the asymptotic eigenvalue distribution of polynomials or rational functions in random matrices.

The notion of freeness refers to a setting of a non-commutative probability space $(M, \tau)$ and given subalgebras $A_i$.

**Definition 2.5.** Let $(M, \tau)$ be a $W^*$-probability space and let $I$ be an index set.

1) Let, for each $i \in I$, $A_i \subset M$, be a unital subalgebra. The subalgebras $(A_i)_{i \in I}$ are called free or freely independent, if $\tau(a_1 \cdots a_k) = 0$ whenever we have: $k$ is a positive integer; $a_j \in A_{(i)}$ (with $i(j) \in I$) for all $j = 1, \ldots, k$; $\tau(a_j) = 0$ for all $j = 1, \ldots, k$; and neighboring elements are from different subalgebras, i.e., $i(1) \neq i(2), i(2) \neq i(3), \ldots, i(k-1) \neq i(k)$.

2) Let, for each $i \in I$, $X_i \in M$. The random variables $(X_i)_{i \in I}$ are called free or freely independent, if their generated unital subalgebras are free, i.e., if $(A_i)_{i \in I}$ are free, where, for each $i \in I$, $A_i$ is the unital subalgebra of $M$ which is generated by $X_i$.

3) Let, for each $i \in I$, $X_i \in M$. The random variables $(X_i)_{i \in I}$ are called *-free or *-freely independent, if their generated unital *-subalgebras are free, i.e., if $(A_i)_{i \in I}$ are free, where, for each $i \in I$, $A_i$ is the unital *-subalgebra of $M$ which is generated by $X_i$.

For selfadjoint $X_i$, there is no difference between freeness and *-freeness. In the case of general normal $X_i$, we will usually consider *-free variables.
Let us point out that the notion of freeness, which looks quite algebraic, goes well with
von Neumann algebras. Namely, if unital subalgebras $A_i$ are free in $(M, \tau)$, then also
their generated von Neumann subalgebras are free.

A crucial fact on freeness is that if we have a tuple $X = (X_1, \ldots, X_d)$ of free selfadjoint
operators, then for any selfadjoint non-commutative polynomial $P$, the analytic distribu-
tion of the evaluation $P(X) := P(X_1, \ldots, X_d)$ (which is again a selfadjoint operator) is
uniquely determined by the marginal analytic distributions $\mu_{X_i}$. Hence, for each selfad-
joint polynomial $P$ in $d$ non-commuting variables we have a corresponding polynomial free
convolution $P$, mapping $d$-tuples of probability measures on $\mathbb{R}$ to another probability
measure on $\mathbb{R}$, corresponding to $P(\mu_{X_1}, \ldots, \mu_{X_d}) = \mu_{P(X_1, \ldots, X_d)}$. However, there are no
explicit descriptions of this functional relationship and getting qualitative and quantitative
properties of $\mu_{P(X)}$ out of corresponding information for the marginal distributions
is one of the big problems. Up to now such concrete descriptions were only known for
the case of the sum and the product of free random variables (corresponding to the opera-
tions of additive and multiplicative, respectively, free convolution). On the conceptual
side, there was quite some progress by relating polynomials in free independent variables
via linearization to the operator-valued version of additive convolution. This general ap-
proach goes back to the work [BMS17], and was brought to fruition for the problem of
dealing with atoms in the recent work [BBL21]. However, it is hard to get explicit results
on atoms from those approaches.

Our starting point here to address general polynomials (and also rational functions)
in free variables is the crucial recent insight [MSY19] that in the case of marginal distri-
butions without atoms, rational functions in such free variables can be modelled by an
abstract algebraic object, the so-called free field. In the next subsection we will recall the
basic properties of this object. By realizing marginal distributions with atoms (of rational
weights) via matrices over the free field we will then be able to get our results on atoms
of functions of free variables, not only for polynomials, but also for non-commutative
rational functions.

Our results show that the polynomial convolutions $P$, as well as their generalizations
to rational functions, can actually also be considered as mappings within the class of
discrete sub-probability measures; and in this setting they also make sense not just for
selfadjoint polynomials and measures on $\mathbb{R}$, but actually for arbitrary rational functions
and discrete sub-probability measures on $\mathbb{C}$. For each non-commutative rational function
$R$, we have a corresponding rational convolution $R$ on discrete sub-probability measures
on $\mathbb{C}$, according to $R(\mu_{X_1}, \ldots, \mu_{X_d}) = \mu_{R(X_1, \ldots, X_d)}$.

Note that it is crucial for our arguments that the operators $X_1, \ldots, X_d$ are normal, even
though the evaluation $R(X_1, \ldots, X_d)$ does not have to be so. Indeed, if we start from
non-normal operators it will be hard to find polynomials which preserve normality.

2.3. The category of $\mathbb{C}[x_1, \ldots, x_d]$-algebras, rational and division closures. We
denote by $\mathbb{C}[x_1, \ldots, x_d]$ the algebra of non-commutative polynomials in the indeterminates
$x_1, \ldots, x_d$.

In this article, we are interested in $d$-tuples of elements in various algebras. It is natural
to represent such a tuple as a $\mathbb{C}[x_1, \ldots, x_d]$-algebra in the category of $\mathbb{C}[x_1, \ldots, x_d]$-
algebras. The following definition is based on the so-called $\mathcal{R}$-ring, where $\mathcal{R}$ is another
ring with some homomorphism (see, for example, [Coh06 Section 7.2]). We change the name accordingly since we are mainly interested in algebras instead of rings.

**Definition 2.6.** A $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebra is a unital complex algebra $A$ together with a homomorphism $\phi : \mathbb{C}\langle x_1, \ldots, x_d \rangle \to A$. We denote $\Sigma_\phi := \prod_{n=1}^{\infty} \{ A \in M_n(\mathbb{C}\langle x_1, \ldots, x_d \rangle) : \phi(A) \text{ is invertible in } M_n(A) \}$, where we also denote by $\varphi$ the matricial amplification $\varphi((a_{ij})_{i,j=1}^{n}) := ((\varphi(a_{ij}))_{i,j=1}^{n})$ for any homomorphism $\varphi$.

Let $A$ and $B$ be $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebras with homomorphisms $\phi_A : \mathbb{C}\langle x_1, \ldots, x_d \rangle \to A$ and $\phi_B : \mathbb{C}\langle x_1, \ldots, x_d \rangle \to B$.

1. An $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebra homomorphism $f : A \to B$ is a homomorphism such that $f \circ \phi_A = \phi_B$.

2. An $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebra subhomomorphism is a $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebra homomorphism $f : A_f \to B$, where $A_f$ is a $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-subalgebra of $A$ such that
   - its homomorphism $\phi_f : \mathbb{C}\langle x_1, \ldots, x_d \rangle \to A_f$ agrees with $\phi_A$
   - and $\Sigma_{\phi_f} \subset \Sigma_{\phi_A}$.

3. Two subhomomorphisms from $A$ to $B$ are called equivalent if they agree on an $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-subalgebra $A_0$ of $A$ such that their common restriction to $A_0$ is again a subhomomorphism.

4. A specialization from $A$ to $B$ is an equivalence class of subhomomorphisms from $A$ to $B$.

Note that the definition of subhomomorphism (Item (2) of Definition 2.6) is slightly different from the one in [Coh06 Section 7.2] since subhomomorphisms therein are considered for epic skew fields instead of algebras (see Section 2.4). However, the two definitions coincide if one restrict oneself to epic skew fields: if $K$ and $L$ are two epic skew fields, a subhomomorphism $f$ from $K$ to $L$ with $f : K_f \to L$ satisfies that an element $x \in K_f$ such that $x \notin \ker f$ is invertible in $K_f$.

In the sequel, for a unital complex algebra $A$, the $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebra structure usually comes from the evaluation of some tuple $X = (X_1, \ldots, X_d)$ over $A$. Namely, we usually take the homomorphism $\phi : \mathbb{C}\langle x_1, \ldots, x_d \rangle \to A$ as $\phi(P) = P(X)$, $\forall P \in \mathbb{C}\langle x_1, \ldots, x_d \rangle$. In particular, we also write $\Sigma_X := \Sigma_\phi$ in such a case.

The following construction is an important example of $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebra that will be considered later.

**Definition 2.7.** Let $X = (X_1, \ldots, X_d)$ be a tuple of elements in a unital complex algebra $A$. The rational closure of $X$ in $A$ is the set, denoted by $A_X$, of all entries of inverses of evaluations of matrices in $\Sigma_X$ at $X$.

For a tuple $X = (X_1, \ldots, X_d)$ over a unital complex algebra $A$, we denote by $\mathbb{C}\langle X \rangle$ the subalgebra of $A$ which is generated by $X_1, \ldots, X_d$. A rational closure $A_X$ of $X$ in $A$ is always a subalgebra of $A$ and contains $\mathbb{C}\langle X \rangle$ (see [Coh06 Section 7.1] for more details).

Let $X$ and $Y$ be two tuples over unital complex algebras $A$ and $B$ respectively. If $f : A_f \to B_Y$ is a subhomomorphism from a $A_X$ to $B_Y$ with the homomorphism $\phi_f : \mathbb{C}\langle x_1, \ldots, x_d \rangle \to A_f$, then $f$ is actually surjective due to the property $\Sigma_Y \subset \Sigma_{\phi_f}$. Hence $f$
naturally induce an isomorphism from $\mathcal{A}_f/\ker(f)$ onto $\mathcal{B}_Y$. Moreover, if $\mathcal{B}_Y = \mathcal{A}_X$, then $\mathcal{A}_f = \mathcal{A}_X$ and $f$ reduces to the identity map, as follows.

**Proposition 2.8.** Let $X = (X_1, \ldots, X_d)$ be a tuple over a unital complex algebra $\mathcal{A}$. The identity map is the only specialization from the rational closure $\mathcal{A}_X$ to itself.

**Proof.** Let $\phi : \mathbb{C}\langle x_1, \ldots, x_d \rangle \to \mathcal{A}_X$ be given by the evaluation map at $X$ inside $\mathcal{A}_X$. Let $f : \mathcal{A}_f \to \mathcal{A}_X$ be a subhomomorphism from $\mathcal{A}_X$ to itself, with $\mathcal{A}_f \subset \mathcal{A}_X$ and $\phi_f : \mathbb{C}\langle x_1, \ldots, x_d \rangle \to \mathcal{A}_X$ given by the evaluation map at $X$ inside $\mathcal{A}_X$.

We have $\Sigma_\phi \subset \Sigma_{\phi_f}$ because of the subhomomorphism $f : \mathcal{A}_f \to \mathcal{A}_X$, and as a consequence, all entries of evaluations of inverses of matrices in $\Sigma_\phi$ at $X$ are in fact entries of evaluations of inverses of matrices in $\Sigma_{\phi_f}$, and by consequence belong to $\mathcal{A}_f$. As a consequence, $\mathcal{A}_X \subset \mathcal{A}_f$ which implies that $\mathcal{A}_X = \mathcal{A}_f$.

The identity map and the homomorphism $f : \mathcal{A}_X \to \mathcal{A}_X$ coincides on $\mathbb{C}\langle X \rangle$ so they also coincide on the rational closure $\mathcal{A}_X$ of $X$. Indeed, if $c \in \mathcal{A}_X$, there exists some $m \times m$ matrix $A \in \Sigma_X$ and some integers $1 \leq i, j \leq m$ such that $c$ is the $(i, j)$-entry of $A(X)^{-1}$.

Now, we have

$$I_m = f(I_m) = f(A(X)A(X)^{-1}) = f(A(X))f(A(X)^{-1}) = A(X)f(A(X)^{-1}),$$

which means that $f(A(X)^{-1}) = A(X)^{-1}$, and $f(c) = c$. \hfill \square

Moreover, if $f : \mathcal{A} \to \mathcal{B}$ is an $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-isomorphism for two $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebras $\mathcal{A}$ and $\mathcal{B}$, then for any tuple $X$ over $\mathcal{A}$ and $Y$ over $\mathcal{B}$, $f$ naturally induce an isomorphism from $\mathcal{A}_X$ onto $\mathcal{B}_Y$. This allows us to give a criterion for the isomorphism of rational closures that are given by tuples over tracial $W^*$-probability spaces. That is, if two tuples have the same $*$-distribution, then they generate isomorphic von Neumann algebras and thus have isomorphic rational closures. We record this observation as the following remark.

**Remark 2.9.** Let $(\mathcal{M}_1, \tau_1)$ and $(\mathcal{M}_2, \tau_2)$ be two tracial $W^*$-probability spaces. Suppose that $X = (X_1, \ldots, X_d) \in \mathcal{M}_1^d$ and $Y = (Y_1, \ldots, Y_d) \in \mathcal{M}_2^d$ are two tuples that have the same $*$-distribution. Then the rational closures $L^0(\mathcal{M}_1)_X$ and $L^0(\mathcal{M}_2)_Y$ are isomorphic.

A rational closure $\mathcal{A}_X$ has a property called division closed, i.e., if $a \in \mathcal{A}_X$ is invertible in $\mathcal{A}$, then $a^{-1} \in \mathcal{A}_X$. The smallest subalgebra with such a property is another important example of $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebras.

**Definition 2.10.** Let $X = (X_1, \ldots, X_d)$ be a tuple over a unital complex algebra $\mathcal{A}$. The division closure of $X$ in $\mathcal{A}$ is the smallest division closed subalgebra of $\mathcal{A}$ containing $\mathbb{C}\langle X \rangle$. We denote it by $\mathbb{C}\langle X \rangle_\div$.

A priori the division closure $\mathbb{C}\langle X \rangle_\div$ is contained in the rational closure $\mathcal{A}_X$ for a given tuple $X$. However, they may agree under certain circumstance, for example, when one of them is a skew field (see, for example, [MSY19, Proposition 4.9]).

As $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-subalgebras inside rational closures, the isomorphism of division closures follows from that of rational closures.

**Remark 2.11.** Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two unital complex algebras with the division closure $D_i := \mathbb{C}\langle X_i \rangle_\div$ for some tuple $X_i$ over $\mathcal{A}_i$, $i = 1, 2$. If $f : \mathcal{A}_1 \to \mathcal{A}_2$ is an $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-isomorphism, then $f(D_1) = D_2$. 


Proof. It’s not difficult to see that $f$ preserves the division closedness. So $f(D_1)$ is division closed, and thus $D_2 \subseteq f(D_1)$. Hence $f^{-1}(D_2) \subseteq D_1$. Note that $f^{-1}(D_2)$ is division closed, we have $D_1 \subseteq f^{-1}(D_2)$. Therefore, $D_1 \subseteq f^{-1}(D_2) \subseteq D_1$, from which the conclusion follows. \hfill \Box

2.4. Free field. In this subsection, we want to introduce the free field that is constituted by non-commutative rational functions. We will also introduce some closely related notions such as rational expressions and linear representations.

Let us point out that the free field is actually a non-commutative algebra, hence “free skew field” might be a more precise name (at least for people coming from a commutative setting).

It is well-known that commutative polynomials can be embedded uniquely into the field of fractions, which is a field consisting of the quotients of polynomials. However, the non-commutative analogue of the field of fractions is highly non-trivial. In order to introduce this object, let us first give some necessary notations.

Definition 2.12. A $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebra $K$ is called a $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-field if $K$ is a skew field.

Let $K$ be a $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-field with $\phi: \mathbb{C}\langle x_1, \ldots, x_d \rangle \to K$.

1. We call $K$ epic if $K$ is generated by the image $\phi(\mathbb{C}\langle x_1, \ldots, x_d \rangle)$, i.e., there is no proper subfield of $K$ containing $\phi(\mathbb{C}\langle x_1, \ldots, x_d \rangle)$.

2. An epic $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-field $K$ is called field of fractions of $\mathbb{C}\langle x_1, \ldots, x_d \rangle$ if the homomorphism $\phi$ is injective.

One may expect that, like in the commutative case, a (skew) field of fractions of $\mathbb{C}\langle x_1, \ldots, x_d \rangle$ (if it exists) is unique. However, it turns out that there exist fields of fractions of $\mathbb{C}\langle x_1, \ldots, x_d \rangle$ which are not isomorphic (see, for example, [Coh06, Exercise 7.2.13]). To understand the non-commutative counterpart of commutative rational functions, we need some universal property for $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-fields with the help of subhomomorphisms.

Definition 2.13. An epic $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-field $U$ is called a universal $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-field if for any epic $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-field $K$ there is a unique specialization from $U$ to $K$. If $U$ is in addition a field of fractions of $\mathbb{C}\langle x_1, \ldots, x_d \rangle$, then we call $U$ the universal field of fractions of $\mathbb{C}\langle x_1, \ldots, x_d \rangle$.

There indeed exists a universal field of fractions of $\mathbb{C}\langle x_1, \ldots, x_d \rangle$, which we call the free field and denote it by $\mathbb{C}\langle x_1, \ldots, x_d \rangle$. We won’t go further into the details of the constructions of the free field and refer the interested reader to [Coh06, Chapter 7] for more details. An element in the free field is called a non-commutative rational function.

A non-commutative rational function can be represented by non-commutative rational expressions. They are syntactically valid combinations of $\mathbb{C}$ and symbols $x_1, \ldots, x_d$ with $+,$, $\cdot$, $^{-1}$, and $()$, which are respectively corresponding to addition, multiplication, taking inverse, and ordering these operations. Non-commutative rational expressions are formal objects or expressions rather than functions. For example, $x_1 + (-1) \cdot x_1$ and $0$ are two distinct non-commutative rational expressions though they represent the same non-commutative rational function. We refer here to [HW15] for a rigorous definition of
non-commutative rational expressions that is based on the graph theory and by which arithmetic operations for rational expressions can be interpreted as operations on graphs. Moreover, in [Ami66], Amitsur constructed the free field \( \mathbb{C}\langle x_1, \ldots, x_d \rangle \) as equivalence classes of non-commutative rational expressions (see also [KV, Section 2]). For simplicity, we usually omit "non-commutative" and call them rational expressions unless otherwise indicated. The same convention will also be used for non-commutative polynomials and non-commutative rational functions.

Now we can recursively define the domain of rational expressions.

**Definition 2.14.** Let \( A \) be a unital complex algebra. For each rational expression \( R \) in indeterminates \( x_1, \ldots, x_d \), we define its \( A \)-domain \( \text{dom}_A(R) \subseteq A^d \) together with its evaluation \( R(X) \) for any tuple \( X = (X_1, \ldots, X_d) \in \text{dom}_A(r) \) by the following rules:

1. If \( R = \lambda \) for any \( \lambda \in \mathbb{C} \), we define \( \text{dom}_A(R) = A^d \) and \( R(X) = \lambda \).
2. If \( R = x_i \) for some \( 1 \leq i \leq d \), we define \( \text{dom}_A(R) = A^d \) and \( R(X) = X_i \).
3. For two rational expressions \( R_1, R_2 \), we define
   \[
   \text{dom}_A(R_1 \cdot R_2) = \text{dom}_A(R_1 + R_2) = \text{dom}_A(R_1) \cap \text{dom}_A(R_2)
   \]
   and
   \[
   (R_1 \cdot R_2)(X) = R_1(X) \cdot R_2(X),
   R_1 + R_2)(X) = R_1(X) + R_2(X).
   \]
4. For a rational expression \( R \), we define
   \[
   \text{dom}_A(R^{-1}) = \{X \in \text{dom}_A(R) : R(X) \text{ is invertible in } A\}
   \]
   and
   \[
   R^{-1}(X) = (R(X))^{-1}.
   \]

A very useful tool for studying rational expressions or functions is the so-called linearization or linear representation. It is well-known and used in many different theories of mathematics (see, for example, [CR99, HMS18]). In order to introduce this tool, let us first define linear matrices.

**Definition 2.15.** A matrix \( A \) in \( M_m(\mathbb{C}\langle x_1, \ldots, x_d \rangle) \) (which we also identify with \( M_m(\mathbb{C}) \otimes \mathbb{C}\langle x_1, \ldots, x_d \rangle \)) is called linear if it is of the form

\[
A = A_0 \otimes 1 + A_1 \otimes x_1 + \cdots + A_d \otimes x_d,
\]

where \( A_0, A_1, \ldots, A_d \) are \( m \times m \) matrices over \( \mathbb{C} \). Note that many authors also call such a matrix affine linear since it allows a constant term.

The following notion of linear representation of rational expressions is taken from [HMS18] and will be important in the sequel.

**Definition 2.16.** Let \( R \) be a rational expression. A formal linear representation of \( R \) of dimension \( m \) is a tuple \( \rho = (u, Q, v) \) consisting of an \( m \times m \) linear matrix \( Q \) over \( \mathbb{C}\langle x_1, \ldots, x_d \rangle \), a row \( u \in M_{1,m}(\mathbb{C}) \) and a column \( v \in M_{m,1}(\mathbb{C}) \) for some integer \( m \) such that for any complex algebra \( A \)

- \( \text{dom}_A(R) \subseteq \text{dom}_A(Q^{-1}) \), where
  \[
  \text{dom}_A(Q^{-1}) := \{X \in A^d : Q(X) \text{ is invertible in } M_m(A)\}\]
for any tuple $X = (X_1, \ldots, X_d) \in \text{dom}_A(R)$, we have $R(X) = -uQ(X)^{-1}v$.

We denote its display by
$$L := \begin{pmatrix} 0 & u \\ v & Q \end{pmatrix}.$$

For each rational expression, we can construct a formal linear representation, see [HMS18, Section 5.2]. We remark that rational functions also have linear representations which are more or less the same as the ones of rational expressions. But we won’t go into details of their linear representations since they are not needed in this article. We refer the interested reader to [CR99, Mai15] for more details.

2.5. Sylvester rank function. We will need two rank functions in the sequel. One is defined algebraically for general rings and the other one is defined via von Neumann algebra theory. However, they share some common characteristics, which are captured by the notion of algebraic rank function ([Mal80]) or Sylvester matrix rank function ([Sch85, Section 7]). Let us introduce this notion as Sylvester rank function. Since we are mainly interested in objects like von Neumann algebras and free fields, we will state all definitions for unital complex algebras.

**Definition 2.17.** For a unital complex algebra $A$, a Sylvester rank function on $A$ is an $\mathbb{R}^+\text{-valued function } \text{rk}$ on the set of rectangular matrices over $A$ satisfying:

1. $\text{rk}(0) = 0$ and $\text{rk}(1) = 1$;
2. $\text{rk}(AB) \leq \min(\text{rk}(A), \text{rk}(B))$;
3. $\text{rk}\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) = \text{rk}(A) + \text{rk}(B)$;
4. $\text{rk}\left(\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}\right) \geq \text{rk}(A) + \text{rk}(B)$.

Moreover, if $\text{rk}(A) > 0$ for any non-zero matrix $A$, then we say $\text{rk}$ is faithful.

Item 4 is not really needed in our study but we keep it for the sake of completeness. We record here some basic properties of a Sylvester rank function.

**Lemma 2.18.** Let $A$ be a unital complex algebra with a Sylvester rank function $\text{rk}$.

1. If $A, B$ are matrices over $A$ such that $B$ is invertible, then $\text{rk}(AB) = \text{rk}(A)$.
2. If $A \in M_n(A)$ is invertible, then $\text{rk}(A) = n$. In particular, $\text{rk}(I_n) = n$.
3. $\text{rk}$ is subadditive, i.e., for any rectangular matrices $A, B$,
   $$\text{rk}(A + B) \leq \text{rk}(A) + \text{rk}(B).$$

**Proof.** Let $A, B$ be matrices over $A$ such that $B$ is invertible. According to Item 2 of Definition 2.17, $\text{rk}(AB) \leq \text{rk}(A)$. Since $A = (AB)B^{-1}$, we apply Item 2 again to see that $\text{rk}(A) \leq \text{rk}(AB)$. Hence we have $\text{rk}(AB) = \text{rk}(A)$.

Combining the fact $\text{rk}(1) = 1$ and Item 3 of Definition 2.17 we see that $\text{rk}(I_n) = n$. Now, let $A \in M_n(A)$ be an invertible matrix, then by the above paragraph we have $\text{rk}(A) = \text{rk}(I_n) = n$ since $AA^{-1} = I_n$. 
Finally, we consider the matrix factorization
\[
\begin{pmatrix}
A + B & 0 \\
0 & 0
\end{pmatrix} = \begin{pmatrix} 1 & 1 \\
0 & 0
\end{pmatrix} \begin{pmatrix} A & 0 \\
0 & B
\end{pmatrix} \begin{pmatrix} 1 & 0 \\
0 & 1
\end{pmatrix}.
\]
Then the rank of the left hand side is \(\text{rk}(A + B)\) thanks to Item 3 of Definition 2.17 while the rank of the right hand side is smaller than or equal to
\[
\text{rk}\left( \begin{pmatrix} A & 0 \\
0 & B
\end{pmatrix} \right) = \text{rk}(A) + \text{rk}(B)
\]
thanks to Item 2 and 3. This completes the proof of the lemma. □

In general a square matrix \(A\) with maximal \(\text{rk}(A)\) may not be invertible. Let us introduce the following notion of regularity for Sylvester rank function which connects the maximality of a Sylvester rank function with the invertibility of matrices.

**Definition 2.19.** Let \(A\) be a unital complex algebra with a Sylvester rank function \(\text{rk}\). If for any integer \(n\) and any matrix \(A \in M_n(A)\) with \(\text{rk}(A) = n\), \(A\) is invertible, then we say \(\text{rk}\) is regular.

We know that for each rational expression we can associate a formal linear representation. The explicit link between the rank of a polynomial expression and the rank of its formal linearization has been established in [BBL21, Lemma 3.4]. Let us reformulate this rank equality in our more general situation, for rational expressions and an arbitrary tuple of random variables.

**Lemma 2.20.** Let \(A\) be a complex algebra with a Sylvester rank function \(\text{rk}\) and let \(X = (X_1, \ldots, X_d)\) be a given tuple over \(M_n(A)\) for some integer \(n\). Let \(R\) be a rational expression and \(X \in \text{dom}_{M_n(A)}(R)\). For any formal linear representation \(\rho = (u, Q, v)\) of \(R\) of dimension \(m\) (see Definition 2.16), we have
\[
\text{rk}(R(X)) + nm = \text{rk}(L(X)),
\]
where \(L\) is the display of \(\rho\).

**Proof.** Since \(X \in \text{dom}_{M_n(A)}(R)\), we have \(Q(X)\) is invertible and \(R(X) = -uQ(X)^{-1}v\). Thus the factorization
\[
\begin{pmatrix}
R(X) & 0 \\
0 & I_{mn}
\end{pmatrix} = \begin{pmatrix} -uQ(X)^{-1}v & 0 \\
0 & I_{mn}
\end{pmatrix} = \begin{pmatrix} I_n & -uQ(X)^{-1} \\
0 & Q(X)^{-1}
\end{pmatrix} \begin{pmatrix} 0 & u \\
v & Q(X)
\end{pmatrix} \begin{pmatrix} I_n & 0 \\
-Q(X)^{-1}v & I_{mn}
\end{pmatrix}
\]
follows. Hence
\[
\text{rk}\left( \begin{pmatrix} R(X) & 0 \\
0 & I_{mn}
\end{pmatrix} \right) = \text{rk}\left( \begin{pmatrix} 0 & u \\
v & Q(X)
\end{pmatrix} \right)
\]
as the rank function \(\text{rk}\) is preserved by invertible matrices. So we see \(\text{rk}(L(X)) = \text{rk}(R(X)) + \text{rk}(I_{nm}) = \text{rk}(R(X)) + nm\), which completes the proof. □
Proposition 2.21. Let $\mathcal{A}$ and $\mathcal{B}$ be two unital complex algebras with faithful and regular Sylvester rank functions $\text{rk}_{\mathcal{A}}$ and $\text{rk}_{\mathcal{B}}$ respectively. Let $X = (X_1, \ldots, X_d) \in \mathcal{A}^d$, $Y = (Y_1, \ldots, Y_d) \in \mathcal{B}^d$ be two tuples satisfying that for any matrix $A$ over $\mathbb{C}(x_1, \ldots, x_d)$,
\[
\text{rk}_{\mathcal{B}}(A(Y)) \leq \text{rk}_{\mathcal{A}}(A(X)).
\]
Then we have the following.

(1) There is a unique specialization from the rational closure $\mathcal{A}_X$ to $\mathcal{B}_Y$.

(2) For any rational expression $R$ in variables $x_1, \ldots, x_d$ such that $R(Y)$ is well-defined, $R(X)$ is also well-defined and
\[
\text{rk}_{\mathcal{B}}(R(Y)) \leq \text{rk}_{\mathcal{A}}(R(X)).
\]

Proof. Let us first give a proof of the first item which is inspired by the proof of [Z19 Theorem 6.3]. Let $Z := (X_1 \oplus Y_1, \ldots, X_d \oplus Y_d) \in \mathcal{C} := \mathcal{A} \oplus \mathcal{B}$. For all $A \in M_n(\mathbb{C}(x_1, \ldots, x_d))$,
\[
A(Z) = A(X) \oplus A(Y),
\]
which means that $\Sigma_Z = \Sigma_X \cap \Sigma_Y$. So we see that the rational closure $\mathcal{C}_Z \subset \mathcal{A}_X \oplus \mathcal{B}_Y$ and thus we can consider the projection maps $\pi_1 : \mathcal{C}_Z \to \mathcal{A}_X$ and $\pi_2 : \mathcal{C}_Z \to \mathcal{B}_Y$.

Let $c$ be an element in $\mathcal{C}_Z$. There exists some $m \times m$ matrix $A \in \Sigma_Z$ and some integers $1 \leq i, j \leq m$ such that $c$ is the $(i, j)$-entry of $A(Z)^{-1}$. That is, we can write $c = e_i A(Z)^{-1} e_j^T$, where $(e_1, \ldots, e_m)$ stands for the canonical basis of $\mathbb{C}^m$. Suppose that $\pi_1(c) = e_i A(X)^{-1} e_j^T = 0$. According to the proof of Lemma 2.20
\[
\text{rk}_{\mathcal{A}}(\pi_1(c)) + m = \text{rk}_{\mathcal{A}}\left(\begin{pmatrix} 0 & e_i \\ e_j^T & A(X) \end{pmatrix}\right) \quad \text{and} \quad \text{rk}_{\mathcal{B}}(\pi_2(c)) + m = \text{rk}_{\mathcal{B}}\left(\begin{pmatrix} 0 & e_i \\ e_j^T & A(Y) \end{pmatrix}\right).
\]
(Lemma 2.20 addresses linear matrices but its proof actually works for matrices that are not necessarily linear.) So by our assumption,
\[
\text{rk}_{\mathcal{B}}(\pi_2(c)) \leq \text{rk}_{\mathcal{A}}(\pi_1(c)) = 0.
\]
Because of the faithfulness of $\text{rk}$, it allows us to say that $\pi_2(c) = 0$. Consequently, $c = (0, 0)$. Hence $\pi_1$ is injective and it is possible to define an $\mathbb{C}(x_1, \ldots, x_d)$-algebra isomorphism $g : \text{im}(\pi_1) \to \mathcal{C}_Z$.

Now, we consider $f := \pi_2 \circ g : \text{im}(\pi_1) \to \mathcal{B}_Y$ and thus $\Sigma_{\phi_f} = \Sigma_Z$. Then we want to see that it is a subhomomorphism from $\mathcal{A}_X$ to $\mathcal{B}_Y$, i.e., $\Sigma_Y \subset \Sigma_Z$. If $A \in \Sigma_Y$, we know that, by the regularity of $\text{rk}_{\mathcal{B}}$,
\[
n = \text{rk}_{\mathcal{B}}(A(Y)) \leq \text{rk}_{\mathcal{A}}(A(X)).
\]
Thanks to the regularity of $\text{rk}_{\mathcal{A}}$, $A \in \Sigma_X$ and thus $\Sigma_Y \subset \Sigma_X$. It follows that $\Sigma_Y \subset \Sigma_Z$, as desired.

Finally, in order to see that the above subhomomorphism induces a unique specialization, let $f' : \mathcal{A}_{f'} \to \mathcal{B}_Y$ be another subhomomorphism with the homomorphism $\phi_{f'} : \mathbb{C}(x_1, \ldots, x_d) \to \mathcal{A}_{f'}$. Then $f'$ is equivalent to $f$ since $\mathcal{A}_{f'} \subset \text{im}(\pi)$ and $f'$ agrees with $f$ on $\mathcal{A}_{f'}$. The inclusion $\mathcal{A}_{f'} \subset \text{im}(\pi)$ is due to the fact that $\Sigma_Y \subset \Sigma_{\phi_{f'}} \subset \Sigma_X$ and thus $\Sigma_{\phi_{f'}} \subset \Sigma_Z$. Moreover, $f'$ and $f$ agree since $f'(A(X)^{-1}) = f(A(X)^{-1}) = A(Y)^{-1}$ when $A \in \Sigma_{\phi_{f'}}$.

Now, let $\mathcal{R}$ be the set of rational expressions such that $R(Y)$ and $R(X)$ are well-defined, and $\text{rk}_{\mathcal{B}}(R(Y)) \leq \text{rk}_{\mathcal{A}}(R(X))$. This set contains the constant and the variables $x_1, \ldots, x_d$. 
Moreover, it is stable by sums and products: it is clear for the well-definedness, and the inequality $\text{rk}_B(R(Y)) \leq \text{rk}_A(R(X))$ is a consequence of Lemma 2.20 as there exist formal linear representations for sums and products.

In order to conclude that $\mathcal{R}$ contains all rational expressions such that $R(Y)$ is well-defined, it suffices to prove that, for all $R \in \mathcal{R}$, whenever $R(Y)$ is invertible, then $R(X)$ is invertible and

$$\text{rk}_B(R(Y)^{-1}) \leq \text{rk}_A(R(X)^{-1}).$$

So let $R \in \mathcal{R}$ such that $R(Y)$ is invertible. We have

$$1 = \text{rk}_B(R(Y)) \leq \text{rk}_A(R(X))$$

which implies that $R(X)$ is invertible, because of the regularity of $\text{rk}$. The inequality

$$\text{rk}_B(R(Y)^{-1}) \leq \text{rk}_A(R(X)^{-1})$$

is again a consequence of Lemma 2.20 as there exist formal linear representations for the inverse of any rational expression. \hfill $\square$

**Corollary 2.22.** Let $\mathcal{A}$ and $\mathcal{B}$ be two unital complex algebras with faithful and regular Sylvester rank functions $\text{rk}_A$ and $\text{rk}_B$ respectively. Let $X = (X_1, \ldots, X_d) \in \mathcal{A}^d$, $Y = (Y_1, \ldots, Y_d) \in \mathcal{B}^d$ be two tuples satisfying that for any matrix $A$ over $\mathbb{C}\langle x_1, \ldots, x_d \rangle$, $\text{rk}_A(A(X)) = \text{rk}_B(A(Y))$.

Then the rational closures $\mathcal{A}_X$ and $\mathcal{B}_Y$ are isomorphic. Moreover, for any rational expression $R$ in variables $x_1, \ldots, x_d$, $R(Y)$ is well-defined if and only if $R(X)$ is also well-defined and

$$\text{rk}_B(R(Y)) = \text{rk}_A(R(X)).$$

Not that if each tuple is closed under taking the adjoint (for example if they are self-adjoint variables), it is also possible thanks to [JZ19, Theorem 6.3] to go beyond and prove an isomorphism between the regular closures (which include the rational closures).

**Proof.** Thanks to Lemma 2.21 we know that there is a specialization from $\mathcal{A}_X$ to $\mathcal{B}_Y$ and another one from $\mathcal{B}_Y$ to $\mathcal{A}_X$. Thanks to Proposition 2.8, those two specialisations are inverses of each other, and by consequence, are isomorphisms of algebras. The last part of the corollary is a consequence of the last part Proposition 2.8. \hfill $\square$

**Lemma 2.23.** Let $\mathcal{A}$ be a unital complex algebra with a Sylvester rank function $\text{rk}$, and let $X = (X_1, \ldots, X_d)$ and $Y = (Y_1, \ldots, Y_d)$ be given tuples over $\mathcal{A}$.

1. For a given matrix $A \in M_n(\mathbb{C}\langle x_1, \ldots, x_d \rangle)$, there exists a constant $c := c(A)$ such that

$$\text{rk}(A(X) - A(Y)) \leq c \max_{i=1,\ldots,d} \{\text{rk}(X_i - Y_i)\}$$

and $c \leq n(d + 2d^2 + \cdots + kd^k)$, where $k$ is the degree of $A$, i.e., the maximal degree of all entries of $A$. Moreover, if $A$ is linear, i.e., $k = 1$, we have

$$\text{rk}(A(X) - A(Y)) \leq n \cdot (\text{rk}(X_1 - Y_1) + \ldots + \text{rk}(X_d - Y_d)).$$

2. Let $R$ be a rational expression such that $X, Y \in \text{dom}_A(R)$. For any formal linear representation $(u, Q, v)$ of $R$ of dimension $n$, we have

$$\text{rk}(R(X) - R(Y)) \leq n \cdot (\text{rk}(X_1 - Y_1) + \ldots + \text{rk}(X_d - Y_d)).$$
Proof. First we want to prove Item (1). We denote $L := \max_{i=1,\ldots,d}\{\rk(X_i - Y_i)\}$ for the sake of simplicity. Suppose that $A = M \otimes x_{i_1} \cdots x_{i_k}$ is a matrix-valued monomial of degree $k$ for indices $i_1, \ldots, i_k \in \{1, \ldots, d\}$. Then we want to show that $\rk(A(X) - A(Y)) \leq n k L$.

Note that we can write

$$A(X) - A(Y) = (M \otimes 1) \cdot \left( I_n \otimes \left( \prod_{j=1}^{k} X_{i_j} - \prod_{j=1}^{k} Y_{i_j} \right) \right),$$

as a consequence of Item (2) and (3) of Definition 2.17, we have

$$\rk(A(X) - A(Y)) \leq n \cdot \rk(\prod_{j=1}^{k} X_{i_j} - \prod_{j=1}^{k} Y_{i_j}).$$

Since

$$\sum_{j=1}^{k} X_{i_j} - \sum_{j=1}^{k} Y_{i_j} = \sum_{j=1}^{k} X_{i_1} \cdots X_{i_{j-1}} (X_{i_j} - Y_{i_j}) Y_{i_{j+1}} \cdots Y_{i_k},$$

by Item (2) of Definition 2.17 and Item (3) Lemma 2.18,

$$\rk(\prod_{j=1}^{k} X_{i_j} - \prod_{j=1}^{k} Y_{i_j}) \leq k \cdot \rk(X_{i_j} - Y_{i_j}) \leq k L,$$

as desired. Therefore, if $A$ is a general matrix in $M_n(C(x_1, \ldots, x_d))$, the desired inequality follows by counting the monomials in $A$ with degree smaller or equal than $k$. Moreover, the particular inequality for the linear case with respect to $\sum_{i=1}^{d} \rk(X_i - Y_i)$ can be seen by a similar argument.

Next, we want to prove Item (2). Since $X, Y \in \text{dom}_A(R)$, $Q(X)$ and $Q(Y)$ are invertible, and we have $R(X) = -u Q(X)^{-1} v$ and $R(X) = -u Q(Y)^{-1} v$. We write

$$\rk(R(X) - R(Y)) = \rk(u Q(X)^{-1} - Q(Y)^{-1}) v) \leq \rk(Q(X)^{-1} - Q(Y)^{-1})$$

due to the matrix factorization

$$\begin{pmatrix} u(Q(X)^{-1} - Q(Y)^{-1}) v & 0 \\ 0 & 0_{n-1} \end{pmatrix} = \begin{pmatrix} u \\ 0 \end{pmatrix} (Q(X)^{-1} - Q(Y)^{-1}) \begin{pmatrix} v \\ 0 \end{pmatrix}.$$

We compute

$$\rk(Q(X)^{-1} - Q(Y)^{-1}) = \rk(Q(Y)^{-1} (Q(Y) - Q(X)) Q(X)^{-1}) \leq \rk(Q(Y) - Q(X))$$

since the rank is preserved by invertible matrices. We conclude thanks to the first item for linear matrices. \qed

2.6. Inner rank and von Neumann rank. The first example of a Sylvester rank function is based on the inner rank, which is defined as follows.

**Definition 2.24.** Let $A$ be a unital complex algebra. For $A \in M_{m,n}(A)$, $m, n \in \mathbb{N}$, we define the **inner rank** $\rho_A(A)$ by

$$\rho_A(A) := \min\{r \geq 1 : A = PQ, P \in M_{m \times r}(A), Q \in M_{r \times n}(A)\}$$

and $\rho_A(0) = 0$ for any rectangular matrix $0$. In addition, we call $A$ full if $\rho_A(A) = \min\{m, n\}$.
The first two items in Definition 2.17 follow from the definition of inner rank. But an inner rank may not be a Sylvester rank function in general. In the sequel, inner rank will majorly be considered on the unital complex algebra \( \mathbb{C} \langle x_1, \ldots, x_d \rangle \) of non-commutative polynomials. This algebra is known as a Sylvester domain (see, for example, its definition and properties \[Coh06, \text{Section 5.5}\]). On a Sylvester domain, Item 3 is satisfied and it’s not difficult to see that Item 4 also holds. So \( \rho_{\mathbb{C} \langle x_1, \ldots, x_d \rangle} \) is a faithful Sylvester function.

Moreover, the inner rank on \( \mathbb{C} \langle x_1, \ldots, x_d \rangle \) extends to the free field \( \mathbb{C}(x_1, \ldots, x_d) \) (which is actually a property of a Sylvester domain). So we will not distinguish these two inner ranks and usually denote them by \( \rho \). On the free field, the inner rank \( \rho \) is actually regular, i.e., full matrices are invertible over the free field. Let us record here this important fact as a remark.

**Remark 2.25.** The inner rank \( \rho \) on the free field \( \mathbb{C}(x_1, \ldots, x_d) \) is a faithful and regular Sylvester rank function.

Another important example of Sylvester rank functions is an analytic rank on von Neumann algebras, which is defined as follows.

**Definition 2.26.** Let \( (\mathcal{M}, \tau) \) be a tracial \( \text{W}^* \)-probability space. For \( A \in M_n(L^0(\mathcal{M})) \), we define its von Neumann rank by

\[
\text{rank}(A) := \text{Tr}_n \otimes \tau(p_{\text{im}(A)}),
\]

where \( p_{\text{im}(A)} \) is the orthogonal projection onto the closure of the image of \( A \) and \( \text{Tr}_n \) is the unnormalized trace on \( M_n(\mathbb{C}) \).

Recall that in Definition 2.4 the weight of each \( \lambda \) in the point spectrum \( \sigma_p(X) \) of \( X \in L^0(\mathcal{M}) \) is defined with the help of kernel. So we also have \( \mu^p_X = \sum_{\lambda \in \sigma_p(X)} (1 - \text{rank}(\lambda - X)) \delta_\lambda \), thanks to the well-known equality \( \text{rank}(Y) = 1 - \tau(p_{\text{ker}(Y)}) \), that holds for any operator \( Y \in L^0(\mathcal{M}) \) (see, for example, [MSY19, Lemma 5.2]). Therefore, for two operator \( X, Y \in L^0(\mathcal{M}) \) we have \( \mu^p_X \leq \mu^p_Y \) if and only if \( \text{rank}(\lambda - Y) \leq \text{rank}(\lambda - X) \) for each \( \lambda \in \mathbb{C} \).

It is probably well-known to experts that the von Neumann rank is indeed a Sylvester rank function on \( L^0(\mathcal{M}) \). We state this here as a lemma and include a proof for reader’s convenience.

**Lemma 2.27.** Let \( (\mathcal{M}, \tau) \) be a tracial \( \text{W}^* \)-probability space. Then rank is a faithful and regular Sylvester rank function on \( L^0(\mathcal{M}) \).

**Proof.** Clearly, Item 1, 3 and 4 of Definition 2.17 follows directly from the definition of the von Neumann rank. Suppose that \( A, B \in M_n(L^0(\mathcal{M})) \) and let us prove Item 2. First, we see that \( \text{rank}(AB) \leq \text{rank}(A) \) since \( \text{im}(AB) \subset \text{im}(A) \). Similarly, we have \( \text{rank}((AB)^*) = \text{rank}(B^* A^*) \leq \text{rank}(B^*) \). Recall that \( \text{rank}(X) = \text{rank}(X^*) \) since \( \tau \) is tracial, we infer that \( \text{rank}(AB) \leq \text{rank}(B) \). So desired inequality \( \text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B)) \) follows.

Moreover, rank is clearly faithful. And the regularity amounts to say that \( A \in M_n(L^0(\mathcal{M})) \) is invertible if and only if \( \text{ker}(A) \neq \{0\} \), which is a well-known fact (see, for example, [MSY19, Lemma 5.5]). \( \square \)
Lemma 2.28. Let $X = (X_1, \ldots, X_d)$ be a tuple of $*$-free normal variables in some tracial $W^*$-probability space $(\mathcal{M}, \tau)$. Then, there exists a constant $C_X > 0$ such that, for all $N \geq 2$, (enlarging the $W^*$-probability space if necessary), there exists a tuple of $*$-free normal variables $X' = (X'_1, \ldots, X'_d)$ such that, for all $1 \leq i \leq d$:

- the weights of the atoms of $\mu_{X'_i}$ are multiples of $1/N$ which are smaller than the corresponding weights of the atoms of $\mu_{X_i}$;
- $\text{rank}(X_i - X'_i) \leq C_X/N$.

**Proof.** Let us prove the result for one variable (the general case can be deduced by taking the free product of $d$ $*$-algebras).

By the spectral theorem, we write $X = \int_{\sigma(X)} \lambda dE_{\lambda}$, where $E_{\lambda}$ is the spectral resolution of $X$. Suppose that $\lambda_1, \ldots, \lambda_l$ are the distinct atoms of $X$ and $q_i := E_{\lambda_i}$ is the orthogonal projection onto the eigenspace $\ker(\lambda_i - X)$ for each $i = 1, \ldots, l$. Moreover, we set $q_0 := 1 - \sum_{i=1}^l q_i$, then $\tau(q_0) = 1 - \sum_{i=1}^l \tau(q_i)$ since $q_i$’s are orthogonal with each other. Therefore, we could consider $X$ in the following form

$$X = q_0 X q_0 + \sum_{i=1}^l \lambda_i q_i.$$ 

We consider a variable with no atoms $S \in \mathcal{M}$ (let say a semicircular variable), and $l$ Bernoulli variables $\beta_i$ of respective sizes $|N\tau(q_i)|/N\tau(q_i)$ in some probability space $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. Then we define

$$X' := q_0 X q_0 \otimes 1 + \sum_{i=1}^l (\lambda_i q_i \otimes \beta_i + S \otimes (1 - \beta_i)) \in \mathcal{M} \otimes L^\infty(\Omega, \mathcal{F}, \mathbb{P}),$$

whose atoms are $\lambda_1, \ldots, \lambda_l$ with respective weights $|N\tau(q_i)|/N$, where $\lfloor c \rfloor$ stands for the largest integer which is smaller than or equal to the number $c \in \mathbb{R}$. Identifying $\mathcal{M}$ as a subalgebra of the finite $W^*$-probability space $\mathcal{M} \otimes L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, we can compute the rank

$$\text{rank}(X - X') \leq \sum_{i=1}^l \text{rank}(1 - \beta_i) \leq \sum_{i=1}^l \left(1 - \frac{|N\tau(q_i)|}{N\tau(q_i)}\right) \leq \frac{1}{N} \left(\sum_{i=1}^l \frac{1}{\tau(q_i)}\right).$$

\[ \square \]

Let $\mu_1, \mu_2$ be two probability measure on $\mathbb{R}$. Their **Kolmogorov distance** is defined as

$$d_{Kol}(\mu_1, \mu_2) := \sup_{t \in \mathbb{R}} |\mathcal{F}_{\mu_1}(t) - \mathcal{F}_{\mu_2}(t)|,$$

where $\mathcal{F}_\mu$ is the cumulative distribution function for a probability measure $\mu$. Let $X, Y \in L^0(\mathcal{M})$ be selfadjoint random variables for some $W^*$-probability space $(\mathcal{M}, \tau)$. Then the Kolmogorov distance of their analytic distributions $\mu_X, \mu_Y$ is controlled by the rank of their difference, as follows.

**Lemma 2.29.** [CMM+21] Lemma 25] Let $X, Y \in L^0(\mathcal{M})$ be selfadjoint random variables, where $(\mathcal{M}, \tau)$ is some $W^*$-probability space. Then

$$d_{Kol}(\mu_X, \mu_Y) \leq \text{rank}(X - Y).$$
Therefore, if $R$ is a rational expression such that $X, Y \in \text{dom}_{A}(R)$, and such that $R(X)$ and $R(Y)$ are selfadjoint, then

$$d_{Kol}(\mu_{R(X)}, \mu_{R(Y)}) \leq \text{rank}(R(X) - R(Y)),$$

and it allows to state the following approximation theorem.

**Theorem 2.30.** Let $X = (X_1, \ldots, X_d)$ and $Y = (Y_1, \ldots, Y_d)$ be two tuples of free selfadjoint random variables in some $W^{*}$-probability space.

- Let $A \in M_n(\mathbb{C} \langle x_1, \ldots, x_d \rangle)$. There exists a constant $c := c(A)$ such that

  $$|\text{rank}(A(X) - A(Y))| \leq c \max_{i=1, \ldots, d} \{d_{Kol}(\mu_{X_i}, \mu_{Y_i})\}.$$

- Let $R$ be a self-adjoint rational expression in $d$ noncommuting variables such that $X, Y \in \text{dom}_{A}(R)$. There exists a constant $c := c(R)$ such that

  $$d_{Kol}(\mu_{R(X)}, \mu_{R(Y)}) \leq c \max_{i=1, \ldots, d} \{d_{Kol}(\mu_{X_i}, \mu_{Y_i})\}.$$

**Proof.** Let $\delta = \max\{d_{Kol}(\mu_{X_i}, \mu_{Y_i}) : i = 1, \ldots, d\}$. As in the proof of [BV93, Theorem 4.11] we may suppose $\{X_i, Y_i\}$ live in a $W^{*}$-probability space $(\mathcal{M}, \tau)$ so that there exist projections $\{p_i\}_{i=1}^{n} \in \mathcal{M}$ such that $\tau(p_i) \geq 1 - \delta$, $p_i X_i p_i = p_i Y_i p_i$ and $\{X_i, Y_i, p_i\}$ is a free family. Decomposing $X_i$ as $p_i X_i p_i + (1 - p_i) X_i p_i + p_i X_i (1 - p_i) + (1 - p_i) X_i (1 - p_i)$, we can compute

$$\text{rank}(X_i - Y_i) \leq \text{rank}((1 - p_i)(X_i - Y_i)p_i)$$

$$+ \text{rank}(p_i(X_i - Y_i)(1 - p_i)) + \text{rank}((1 - p_i)(X_i - Y_i)(1 - p_i))$$

$$\leq 3 \cdot \text{rank}(1 - p_i) \leq 3\delta.$$

As a consequence of Lemma 2.23 we have

$$|\text{rank}(A(X) - A(Y))| \leq K \max_{i=1, \ldots, d} \{\text{rank}(X_i - Y_i)\} \leq 3K\delta$$

for a certain constant $K$. In particular, for any formal linear representation $(u, Q, v)$ of $R$ of dimension $n$, we have

$$\text{rank}(R(X) - R(Y)) \leq n \cdot (\text{rank}(X_1 - Y_1) + \ldots + \text{rank}(X_d - Y_d)) \leq 3n\delta,$$

thanks to Lemma 2.23 again. Then the desired the result follows thanks to Lemma 2.29.

\[\square\]

### 2.7. The quantity $\Delta$

The free field $\mathbb{C} \langle x_1, \ldots, x_d \rangle$ can seem quite abstract. However, it is possible to represent it as a concrete algebra of operators. In [MSY19], a link was discovered between the free field and the rational closure of tuples of operators maximizing a certain quantity $\Delta$. Let us introduce this quantity and recall some results of [MSY19].

Let us denote by $\mathcal{F}(L^2(\mathcal{M}, \tau))$ the $\mathcal{M} - \mathcal{M}$ Hilbert bimodule of finite-rank operators on $L^2(\mathcal{M}, \tau)$, and by $J$ Tomita’s conjugation operator, i.e., the conjugate-linear map $J : L^2(\mathcal{M}, \tau) \to L^2(\mathcal{M}, \tau)$ such that $J(x) = x^*$ for $x \in \mathcal{M}$. Considering a tuple $X = (X_1, \ldots, X_d)$ of noncommutative random variables in $\mathcal{M}$, the quantity $\Delta(X)$ was introduced in [CS05] as

$$\Delta(X) = d - \dim \left\{ (T_1, \ldots, T_d) \in \mathcal{F}(L^2(\mathcal{M}, \tau))^d : \sum_{i=1}^{d} [T_i, JX_iJ] = 0 \right\}_{HS}^H.$$
where the closure is taken with respect to the Hilbert-Schmidt norm. The quantity $\Delta(X)$ is maximal whenever $\Delta(X) = d$, or equivalently whenever there do not exist a $d$-tuple $(T_1, \ldots, T_d) \in \mathcal{F}(L^2(M, \tau))^d$ such that $\sum_{i=1}^d [T_i, X_i] = 0$.

**Theorem 2.31.** [MSY19, Theorem 1.1] Let $X = (X_1, \ldots, X_d)$ be a $d$-tuple of operators in a tracial $W^*$-probability space $(M, \tau)$. The following are equivalent:

1. $\Delta(X) = d$;
2. the evaluation map sending $P \in \mathbb{C}\langle x_1, \ldots, x_d \rangle$ to $P(X)$ extends to an isomorphism from $\mathbb{C}\langle x_1, \ldots, x_d \rangle$ to the rational closure $L^0(M)_X$ of $X$;
3. for any $n \in \mathbb{N}$ and any matrix $A \in M_n(\mathbb{C}\langle x_1, \ldots, x_d \rangle)$, we have
   $$\text{rank}(A(X)) = \rho(A).$$

Note that the rank equality in Item (3) of the above theorem actually holds over the free field provided that one of the equivalent properties holds for a given tuple $X$. That is, Item (3) can be replaced by Item (3’): for any $M_n(\mathbb{C}\langle x_1, \ldots, x_d \rangle)$, we have $\text{rank}(A(X)) = \rho(A)$. This is not explicitly stated in [MSY19, Theorem 1.1] but it can be easily read out from the proof of [MSY19, Theorem 5.6]. We summarize it as a remark that will needed in Section 4.

**Remark 2.32.** For a $d$-tuple $X$ over some $W^*$-probability space with $\Delta(X) = d$, the evaluation map is a rank-preserving homomorphism from $\mathbb{C}\langle x_1, \ldots, x_d \rangle$ to $L^0(M)$.

We also have the following additivity of the quantity $\Delta$.

**Proposition 2.33** (Theorem 3.3 of [CS05]). Let $X_1, \ldots, X_d, X_{d+1}, \ldots X_d'$ be in a tracial $W^*$-probability space $(M, \tau)$. Whenever $X_1, \ldots, X_d$ is $*$-free from $X_{d+1}, \ldots X_d'$, we have

$$\Delta(X_1, \ldots, X_d') = \Delta(X_1, \ldots, X_d) + \Delta(X_{d+1}, \ldots, X_d').$$

**Proof.** As remarked in [CMM+21, Theorem 16], the proof of [CS05, Theorem 3.3], which is written for self-adjoint variables, is valid for not necessarily self-adjoint $*$-free variables. $\square$

## 3. Matrix models of random variables

### 3.1. Matrix model of normal random variables

The goal of this subsection is to bring normal random variables to a matrix form of which the information of atoms can be easily read out. This matrix form will play an important role in Section 4 and 5.

Clearly, if $Y_1, \ldots, Y_d$ are random variables in the $W^*$-probability space $(M_n(\mathbb{C}), \text{tr}_n)$ for some $n \in \mathbb{N}$, then they are already in a matrix form (with atoms of rational weights). So in the following we focus on the random variables that have non-trivial non-atomic parts. We will present two constructions that give us a matrix model for a given random variable.

In the first one, we consider normal random variable whose atoms are of rational weights. For such a variable, we will construct a diagonal matrix such that they have the same distribution. If we have multiple such random variables, the diagonal matrices produced by this procedure will lose the information on the relation between the given random variables. However, with the help of freely independent Haar unitaries, we can further...
build a freely independent tuple of matrices which copy the distribution of each of the
given random variables.

In the second one, we will show that for a normal variable \( Y \), its matrix amplification
\( I_n \otimes Y \) can be decomposed into a form of \( U D U^* \), where \( U \) is unitary and \( D \) is a diagonal
matrix containing the information of atoms. Comparing to the first construction, we can
keep the information of the relation between the given random variables.

The first construction is given as the following lemma.

**Lemma 3.1.** Let \( Y \) be a normal random variable in some \( W^* \)-probability space \((\mathcal{M}, \tau)\)
such that there exists \( n \in \mathbb{N} \) with \( \mu_Y(\{ \lambda \}) = \frac{1}{n} \mathbb{N} \) for each \( \lambda \in \mathbb{C} \). Then there is some
\( W^* \)-probability space \((\mathcal{N}, \varphi)\) such that the distribution of \( Y \) is the one of

\[
D := \begin{pmatrix}
\alpha_1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \alpha_m \\
\gamma_1 & \cdots & 0 \\
0 & \vdots & \ddots \\
0 & \cdots & \gamma_m\end{pmatrix} \in (M_n(\mathcal{N}), \text{tr}_n \circ \varphi),
\]

where \( \gamma_1, \ldots, \gamma_{m_0} \) are *-free normal random variables without atoms, and \( \{\alpha_1, \ldots, \alpha_m\} \in \mathbb{C} \)
is the (multi)-set of atoms of \( Y \), in such a way that the point spectrum distribution \( \mu^\varphi \)
of \( Y \) is \( \frac{1}{n} (\delta_{\alpha_1} + \cdots + \delta_{\alpha_m}) \).

**Proof.** First, recall that we can write \( Y = q_0 Y q_0 + \sum_{i=1}^l \lambda_i q_i \), where \( \lambda_1, \ldots, \lambda_l \in \sigma_p(Y) \),
\( q_1, \ldots, q_l \) are orthogonal projections onto the eigenspaces of \( Y \) and \( q_0 := 1 - \sum_{i=1}^l q_i \) (see
the proof of Lemma 2.25). We denote \( \mathcal{M}_0 := q_0 \mathcal{M} q_0 \) and \( \tau_0(\cdot) := \frac{1}{\tau(q_0)} \tau(\cdot) \), which form a
compressed \( W^* \)-probability space of \((\mathcal{M}, \tau)\). We consider \( \gamma := q_0 Y q_0 \) as a random variable
in \( \mathcal{M}_0 \). Since \( Y \) commutes with \( q_0 \), we see that \( \gamma \) is normal as \( Y \) is normal. Moreover,
\( \lambda - \gamma \) is injective on \( \text{im}(q_0) \) for any \( \lambda \in \mathbb{C} \) otherwise there is some subspace of \( \text{im}(q_0) \)
which is an eigenspace of \( Y \). So the analytic distribution of \( \gamma \) in \((\mathcal{M}_0, \tau_0)\) has no atom.

Next, let \( n \) be an integer such that \( \tau(q_1), \ldots, \tau(q_l) \) can be written as fractions with their
denominators to be \( n \). We denote by \( m_i := m \tau(q_i) \) for each \( i = 1, \ldots, l, m := \sum_{i=1}^l m_i \) and
\( m_0 := n - m \). Now we define \( (\mathcal{N}, \varphi) := (\mathcal{M}_0, \tau_0)^{\otimes m_0} \) to be the free product of \( m_0 \) copies of
\((\mathcal{M}_0, \tau_0)\). Moreover, let \( \gamma_1, \ldots, \gamma_{m_0} \in \mathcal{N} \) to be \( m_0 \) copies of \( \gamma \in \mathcal{M}_0 \) such that \( \gamma_1, \ldots, \gamma_{m_0} \)
are freely independent (and of course, they are normal random variables with non-atomic
analytic distributions). Let us set \( D \in M_n(\mathcal{N}) \) to be a diagonal matrix whose diagonal
entries consists of

- \( m \) constants \( \alpha_1, \ldots, \alpha_m \in \mathbb{C} \), which are given by \( m_i \) copies of the constant \( \lambda_i \) for
each \( 1 \leq i \leq l \);
- \( \gamma_1, \ldots, \gamma_{m_0} \).

Then we see that \( D \) and \( Y \) have the same *-distribution since for any polynomial \( p \in \mathbb{C}[x, x^*] \),

\[
\varphi(p(D, D^*)) = \sum_{i=1}^m \frac{1}{n} p(\alpha_i, \overline{\alpha_i}) + \frac{1}{n} \sum_{j=1}^{m_0} \varphi(p(\gamma_j, \gamma_j^*))
\]
\[
\begin{align*}
= \sum_{i=1}^{l} \frac{m_i}{n} p(\lambda_i, \overline{\lambda_i}) + \frac{1}{n} \sum_{j=1}^{m_0} \varphi(p(\gamma_j, \gamma_j^*)) \\
= \sum_{i=1}^{l} \frac{m_i}{n} (p(\lambda_i, \overline{\lambda_i})) + \frac{m_0}{n} \tau_0(p(\gamma, \gamma^*)) \\
= \sum_{i=1}^{l} \tau(q_i)p(\lambda_i, \overline{\lambda_i}) + \tau(q_0)p(p(\gamma, \gamma^*)) \\
= \sum_{i=1}^{l} \tau(q_i)p(\lambda_i, \overline{\lambda_i}) + \tau(q_0p(Y, Y^*)\cdot q_0) \\
= \tau(p(Y, Y^*/)).
\end{align*}
\]

Since we consider random variables with non-trivial non-atomic part, we will assume the von Neumann algebra to be of type II\(_1\). Then we can embed it, with also preserving the trace, into a type II\(_1\) factor. Thanks to the factoriality, we can construct a unitary matrix by partial isometries that come from equivalent projections. Then we use this unitary matrix to diagonalize our random variables. The details of these statements are given in the following proposition and its proof.

**Proposition 3.2.** Let \(Y_1, \ldots, Y_d\) be normal random variables in some tracial \(W^*\)-probability space. Then they can be embedded into a tracial \(W^*\)-probability space \((\mathcal{M}, \tau)\) such that \(\mathcal{M}\) is a factor. Let \(n \in \mathbb{N}\) be given. Then for each \(Y_k\), there exist a unitary matrix \(U_k \in M_n(\mathcal{M})\) and a diagonal matrix \(D_k \in M_n(\mathcal{M})\) such that

- \(U_kD_kU_k^* = I_n \otimes Y_k\) for each \(k = 1, \ldots, d\);
- for every atom \(\lambda\) of \(Y\) with weight \(w\), there are \([nw]\) diagonal elements of \(D_k\) appear as \(\lambda\).

**Proof.** Thanks to [HW00, Proof of Theorem 2.6] or [Ole12, Theorem A.1.5], a type II\(_1\) von Neumann algebra with a tracial state can be embedded into a type II\(_1\) factor such that the embedding preserves also the trace. (Note that the normal trace on a factor is unique.) Therefore, let us assume that \((\mathcal{M}, \tau)\) is a tracial \(W^*\)-probability space such that \(\mathcal{M}\) is a factor. It’s enough to prove the statement for one normal random variable \(Y \in \mathcal{M}\). First, we write

\[
Y = q_0Yq_0 + \sum_{i=1}^{l} \lambda_iq_i.
\]

as in the proof of Lemma 2.28 where \(\lambda_1, \ldots, \lambda_i \in \sigma_p(Y)\), \(q_1, \ldots, q_l\) are orthogonal projections onto the eigenspaces of \(Y\) and \(q_0 := 1 - \sum_{j=1}^{l} q_j\).

Next, for each \(i = 0, 1, \ldots, l\), let \(m_i := [n\tau(q_i)]\). We expect to find \(m_i\) subprojections of \(q_i\) with trace equal to \(\frac{1}{n}\). This can be done as follows. We know that there exists a projection \(p_n \in \mathcal{M}\) with \(\tau(p_n) = \frac{1}{n}\) since \(\mathcal{M}\) is a type II\(_1\) factor (see, for example, [KR86, Proposition 8.5.3]). If \(m_i \geq 1\), i.e., \(\frac{1}{n} \leq \tau(q_i)\), then by the comparison theory of the factor we infer from \(\tau(p_n) \leq \tau(q_i)\) that \(p_n \precsim q_i\) (see, for example, [KR86, Section 8.5]). This implies that there exists a projection \(p_1 \leq q_i\) with \(\tau(p_1) = \tau(p_n) = \frac{1}{n}\). Similarly we can...
find a projection $p_2^i \leq q_i - p_1^i$ with $\tau(p_2^i) = \frac{1}{n}$ if $m_i \geq 2$. We repeat this procedure until there are $m_i$ orthogonal subprojections $p_1^i, \ldots, p_{m_i}^i$ of $q_i$ with equal trace $\frac{1}{n}$. Therefore, we write $q_i = \sum_{j=1}^{m_i} p_j^i + p_0^i$, where $p_0^i := q_i - \sum_{j=1}^{m_i} p_j^i$ stands for the remaining part. Note that $p_1^i, \ldots, p_{m_i}^i, p_0^i$ commute with $Y$, we see that

$$q_i Y q_i = \left( \sum_{j=1}^{m_i} p_j^i + p_0^i \right) Y \left( \sum_{j=1}^{m_i} p_j^i + p_0^i \right) = \sum_{j=1}^{m_0} p_j^i Y p_j^i + p_0^i Y p_0^i.$$ Therefore, we could write

$$Y = \sum_{i=0}^{m_i} \sum_{j=0}^{p_i^j} p_j^i Y p_j^i.$$ Let us set $p^{i+1} := \sum_{i=0}^{m_i} p_0^i$. Then we have

$$\tau(p^{i+1}) = 1 - \sum_{i=0}^{m_i} \frac{m_i}{n} = n - \frac{\sum_{i=0}^{m_i} m_i}{n}.$$ We denote $m_{i+1} := n - \sum_{i=0}^{m_i} m_i$. Furthermore, we divide also $p^{i+1}$ into $m_{i+1}$ subprojections with trace $\frac{1}{n}$ and denote them by $p_1^{i+1}, \ldots, p_{m_{i+1}}^{i+1}$. So we finally write $Y$ as

$$Y = \sum_{i=0}^{m_{i+1}} \sum_{j=1}^{m_{i+1}} p_j^i Y p_j^i.$$ For simplicity’s sake, we rewrite the indices of $p_j^i$ as $1, \ldots, n$ and thus

$$Y = \sum_{j=1}^{n} p_j Y p_j.$$ Now let us consider the $W^*$-probability space $(M_n(M), \text{tr}_n \circ \tau)$. For each pair of projections $(p_i, p_j)$ ($i < j$) we can find an partial isometry $u_{ij} \in M$ such that $u_{ij} u_{ij}^* = p_j$ and $u_{ij} u_{ij} = p_i$ since $p_i$ and $p_j$ are equivalent relative to $M$ (recall that the equivalence of projections is equivalent to the equality of traces of projections, see, for example [KR86, Section 8.5]). Let $u_{ii} := p_i$ and $u_{ij} := u_{ji}^*$ if $i > j$, then we have $u_{ij} u_{ij} = p_j$ for $i, j = 1, \ldots, n$. We define

$$U := (u_{ij})_{i,j=1}^n,$$ which is a selfadjoint unitary operator in $M_n(M)$. Moreover, let us define

$$D := \begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{pmatrix} \in M_n(M),$$ where $\alpha_j := \sum_{i=1}^{n} u_{ji} Y u_{ij}$. Note that if $p_j$ satisfies $Y p_j = \lambda p_j$ for $\lambda \in \sigma(p)(Y)$, then

$$\alpha_j = \lambda \sum_{i=1}^{n} u_{ji} u_{ij} = \lambda \sum_{i=1}^{n} p_i = \lambda.$$ Moreover, if $\lambda$ has weight $w$, then there are $|nw|$ $p_i$’s that are subprojections of the projection $p_{\ker(\lambda-Y)}$. So $\lambda$ appears $|nw|$ many times in $\alpha_j$’s.
Finally, from a direct verification
\[ UDU = \left( \sum_{j=1}^{n} u_{kj}a_{ji}u_{jl} \right)_{k,l=1}^{n} \]
\[ = \left( \sum_{i,j=1}^{n} u_{kj}u_{ji}u_{jl} \right)_{k,l=1}^{n} \]
\[ = \left( \sum_{j=1}^{n} u_{kj}u_{ji}u_{jl} \right)_{k,l=1}^{n} \]
\[ = \left( \delta_{kl} \sum_{j=1}^{n} p_j Y p_j \right)_{k,l=1}^{n} \]
\[ = I_n \otimes Y, \]
we see that \( UDU \in M_n(\mathcal{M}) \) is indeed a matrix model for random variable \( Y \).

3.2. Non-commutative derivations and \( \Delta \). Let \((\mathcal{M}, \tau)\) be a tracial \( W^*\)-probability space. A non-commutative derivation \( \delta \) on \( \mathcal{M} \) with domain \( A \subset \mathcal{M} \) is a (possibly unbounded) linear map
\[ \delta : L^2(\mathcal{M}, \tau) \supseteq A \to L^2(\mathcal{M}, \tau) \otimes L^2(\mathcal{M}, \tau), \]
such that \( A \) is a \( * \)-algebra which is weakly dense in \( \mathcal{M} \), and such that, for all \( X_1, X_2 \in A \), we have the Leibniz rule
\[ \delta(X_1X_2) = \delta(X_1) \cdot X_2 + X_1 \cdot \delta(X_2). \]

For all \( \eta \in L^2(\mathcal{M}, \tau) \otimes L^2(\mathcal{M}, \tau) \), we say that \( \delta^*(\eta) \) exists (or \( \eta \) is in the domain of \( \delta^* \)) if \( \delta^*(\eta) \) is a vector of \( L^2(\mathcal{M}) \) such that, for all \( X \in \mathcal{A} \), we have
\[ \langle \eta, \delta X \rangle = \langle \delta^*(\eta), X \rangle. \]

For all \( \eta \in A \otimes A \), the map \( A \mapsto A \sharp \eta \), given by \((A_1 \otimes A_2)\sharp(x \otimes y) = (A_1 x \otimes y A_2)\) on tensors, extends continuously to the space \( L^2(\mathcal{M}, \tau) \otimes L^2(\mathcal{M}, \tau) \). Because we have \( \langle \delta(X)\sharp \eta, 1 \otimes 1 \rangle = \langle \delta(X), \eta^* \rangle \), the existence of \( \delta^*(\eta^*) \) is equivalent to the existence of \( (\delta^0)^*(1 \otimes 1) \) for the non-commutative derivation \( \delta^0 : X \mapsto \delta(X)\sharp \eta \). We have the following consequence.

**Lemma 3.3.** [Mai15] Lemma 4.3 Let \( \delta \) be a non-commutative derivation on \( \mathcal{M} \) with domain \( A \subset \mathcal{M} \). The following are equivalent:

- \( \delta^*(1 \otimes 1) \) exists;
- \( \delta^*(\eta) \) exists for all \( \eta \in A \otimes A \);
- \( (\delta^0)^*(1 \otimes 1) \) exists for all \( \eta \in A \otimes A \) (where \( \delta^0 \) is given by \( X \mapsto \delta(X)\sharp \eta \)).

Let \( P_{\Omega} : L^2(\mathcal{M}, \tau) \to L^2(\mathcal{M}, \tau) \) be the orthogonal projection onto the trace vector 1. The Hilbert space \( L^2(\mathcal{M}, \tau) \otimes L^2(\mathcal{M}, \tau) \) is isometrically isomorphic to the space of Hilbert-Schmidt operators on \( L^2(\mathcal{M}, \tau) \) via the map \( A \sharp P_{\Omega} \), given by \((A_1 \otimes A_2)\sharp P_{\Omega} = A_1 P_{\Omega} A_2\) for all \( A_1 \otimes A_2 \in A \otimes A \). The existence of \( \delta^*(1 \otimes 1) \) allows to define some unbounded operators on \( L^2(\mathcal{M}, \tau) \) with very interesting properties, thanks to the following theorem.
Theorem 3.4. ([Shl06] Theorem 1) Let \( \partial \) be a non-commutative derivation on \( \mathcal{M} \) with domain \( \mathcal{A} \subset \mathcal{M} \). If \( \partial^*(1 \otimes 1) \) exists, there exists a closable unbounded operator \( D : L^2(\mathcal{M}, \tau) \to L^2(\mathcal{M}, \tau) \) such that \( [D, X] = (\partial X) \sharp P_\Omega \) for all \( X \in \mathcal{A} \).

Proof. We recall briefly here the construction (it is written for self-adjoint variables in [Shl06], however the proof is valid also in our context, since \( \mathcal{A} \) is a dense \( * \)-algebra in \( \mathcal{M} \)). Let us consider the densely defined operator

\[
D := (\text{id} \otimes \tau) \circ \partial : L^2(\mathcal{M}, \tau) \supseteq \mathcal{A} \to L^2(\mathcal{M}, \tau)
\]

which is such that, for all \( A \in \mathcal{A} \), we have \([D, A] = \partial A \sharp P_\Omega \). Now, the existence of \( \partial^*(1 \otimes 1) \) allows us to define an adjoint

\[
D^* : L^2(\mathcal{M}, \tau) \supseteq \mathcal{A} \to L^2(\mathcal{M}, \tau)
\]

by the formula

\[
D^*(A) := A \cdot \partial^*(1 \otimes 1) - (\partial(\partial^*) \sharp P_\Omega)^* (1) = A \cdot \partial^*(1 \otimes 1) - \tau \otimes \text{id}[(\partial(\partial^*))^*].
\]

As a consequence, \( D \) is closable (see also [MS17] Theorem 8 and Exercise 3). \( \Box \)

We use the previous theorem in order to give a criterion for the maximality of \( \Delta \).

Proposition 3.5. Let \( X_1, \ldots, X_d \) be in a tracial \( W^* \)-probability space \( \mathcal{M} := W^*(X_1, \ldots, X_d) \) endowed with the state \( \tau \). Assume that there exist non-commutative derivations \( \partial_i \) (1 \( \leq \) \( i \) \( \leq \) \( d \)) on \( \mathcal{M} \) with domain \( \mathbb{C}(X_1, \ldots, X_d, X_1^*, \ldots, X_d^*) \), such that \( \partial_i(X_j) = \delta_{ij} 1 \otimes 1 \) and such that \( \partial_i^*(1 \otimes 1) \) exists. Then \( \Delta(X_1, \ldots, X_d) = d \).

In our setting, \( \partial_i \) is defined on \( \mathbb{C}(X_1, \ldots, X_d, X_1^*, \ldots, X_d^*) \) and not only on \( \mathbb{C}(X_1, \ldots, X_d) \) (as it is usually the case for self-adjoint variables). The condition in the Lemma above is exactly the existence of a conjugate system if \( X_i = X_i^* \).

Proof. Thanks to Theorem 3.4, the existence of the derivation \( \partial_i \) allows to define an (unbounded) dual system \( (D_1, \ldots, D_d) \) in the following sense. For each 1 \( \leq \) \( i \) \( \leq \) \( d \), the densely defined operator \( D_i : L^2(\mathcal{M}, \tau) \to L^2(\mathcal{M}, \tau) \) is such that, for all \( A \in \mathbb{C}(X_1, \ldots, X_d, X_1^*, \ldots, X_d^*) \), we have \([D_i, A] = (\partial_i A) \sharp P_\Omega \). In particular, we have

\[
[D_i, X_j] = \delta_{ij} P_\Omega.
\]

In order to conclude that \( \Delta(X_1, \ldots, X_d) = d \), it suffices now to use [MSY19] Proposition 6.2, which is also valid if the dual system is in the algebra of affiliated operators (the proof works mutatis mutandis). \( \Box \)

Lemma 3.6. If a unitary variable \( U \) has a \( L^2 \)-density with respect to the Lesbesgue measure, then there exists a non-commutative derivation \( \partial \) on \( W^*(U) \) with domain \( \mathbb{C}[U, U^{-1}] \) such that \( \partial(U) = 1 \otimes 1 \) and such that \( \partial^*(1 \otimes 1) \) exists. Furthermore, \( \Delta(U) = 1 \).

Proof. Thanks to [Vol99] Proposition 8.7, \( \partial^*(1 \otimes 1) \) exists for the unique non-commutative derivation \( d \) on \( W^*(U) \) with domain \( \mathbb{C}[U, U^{-1}] \) such that \( d(U) = 1 \otimes U \). Using Lemma 3.5, it implies that \( \partial^*(1 \otimes 1) \) exists for the non-commutative derivation \( \partial : X \mapsto d(X) \sharp (1 \otimes U^{-1}) \), which is such that \( \partial(U) = (1 \otimes U) \sharp (1 \otimes U^{-1}) = 1 \otimes 1 \). Proposition 3.5 yields \( \Delta(U) = 1 \). \( \Box \)

An important realization of distributions in matrix form is via the free compression by matrix units. Precise definitions and more information on this can be found in [NS06 Lecture 14].
Theorem 3.7. Let $X^{(1)}, \ldots, X^{(d)}$ be in a tracial $W^*$-probability space. If $(X^{(i)}_{ij})_{i,j=1}^n$ are given by the free compressions of $X^{(1)}, \ldots, X^{(d)}$ by matrix units, and if there exist non-commutative derivations $\partial^{(\ell)} (1 \leq \ell \leq d)$ on $W^*(X^{(1)}, \ldots, X^{(d)})$ with domain $\mathbb{C}(X^{(1)}, \ldots, X^{(d)})$, $(X^{(1)}), \ldots, (X^{(d)})^*$ such that $\partial^{(\ell)}(X^{(i)}_{ij'}) = \delta_{i,i'} \delta_{j,j'} 1 \otimes 1$ and such that $(\partial^{(\ell)})^*(1 \otimes 1)$ exists, then there exist the non-commutative derivations $\partial^{(\ell)}$ on $M := W^*(X^{(i)}_{ij})$ with domain $\mathbb{C}(X^{(i)}_{ij}, (X^{(i)}_{ij})^*: i,j,\ell)$ such that $\partial^{(\ell)}(X^{(i)}_{kl}) = \delta_{i,k} \delta_{j,l} 1 \otimes 1$ and such that $(\partial^{(\ell)})^*(1 \otimes 1)$ exists. Furthermore, $\Delta(X^{(i)}_{ij}: i,j,\ell) = \Delta_{n^2}$. 

Proof. The general case being a straightforward adjustment of the case $d = 1$, let us prove the result for $d = 1$.

Without loss of generality, we can assume that

$$X_{ij} = E_{11}X E_{11} \in (E_{11}W^*(X) \ast M_n(\mathbb{C})E_{11}, n(\varphi \ast tr_n)) = (\mathcal{M}, \tau),$$

where $W^*(X)$ is a finite $W^*$-probability space with state $\varphi$. We extend $\partial$ to a non-commutative derivative $d$ on $W^*(X) \ast M_n(\mathbb{C})$ with domain $\mathbb{C}(X, X^*) \ast M_n(\mathbb{C})$ by imposing $d|_{\mathbb{C}(X,X^*)} = \partial$ and $d|_{M_n(\mathbb{C})} = 0$. Thanks to [Voi99, Proposition 5.14], we know that

$$(\tau \ast tr_n) \circ d = (\tau \ast tr_n) \circ \partial \circ E_{W^*(X)},$$

where $E_{W^*(X)}$ is the conditional expectation from $W^*(X) \ast M_n(\mathbb{C})$ to $E_{W^*(X)}$. As a consequence, we have $d^*(1 \otimes 1) = \partial^*(1 \otimes 1)$.

Now, let us define the non-commutative derivative $d_{ij}$ given by $d_{ij}(A) = d(A)\sharp(E_{11} \otimes E_{1j})$. Using Lemma 3.3, we know that $d_{ij}(1 \otimes 1)$ exists. They are defined on $W^*(X) \ast M_n(\mathbb{C})$ with domain $\mathbb{C}(X, X^*) \ast M_n(\mathbb{C})$, but restricting them to the domain $E_{11}\mathbb{C}(X, X^*) \ast M_n(\mathbb{C})E_{11}$, we get the non-commutative derivatives $\partial_{ij}$ on $E_{11}W^*(X) \ast M_n(\mathbb{C})E_{11}$ such that $\partial_{ij}(X_{kl}) = \delta_{i,k} \delta_{j,l} 1 \otimes 1$. We can compute

$$\langle \partial_{ij}(A), 1 \otimes 1 \rangle_{L^2(\mathcal{M}, \tau) \otimes L^2(\mathcal{M}, \tau)} = n^2\langle d_{ij}(A), 1 \otimes 1 \rangle_{L^2(W^*(X) \ast M_n(\mathbb{C}), \varphi \ast tr_n) \otimes L^2(W^*(X) \ast M_n(\mathbb{C}), \varphi \ast tr_n)}$$

$$= n^2\langle A, d^*_{ij}(1 \otimes 1) \rangle_{L^2(W^*(X) \ast M_n(\mathbb{C}), \varphi \ast tr_n)}$$

which shows that $\partial^*_{ij}(1 \otimes 1) = n d^*_{ij}(1 \otimes 1)$ exists.

Proposition 3.3 yields $\Delta(X_{ij}: i,j) = n^2$. \hfill \Box

The previous lemma and the previous theorem (for $d = 1$) yield the following result.

Corollary 3.8. If $(U_{ij})_{i,j=1}^n$ are given by the free compressions of a unitary variable $U$ by matrix units, and if $U$ has a $L^3$-density with respect to the Lebesgue measure, then $\Delta(U_{ij}: i,j) = n^2$.

Note that free compressions of unitary variables correspond to quantum random variables on the unitary dual group, as studied in [CU16], and Haar unitaries correspond to the particular case of the Haar trace on the unitary dual group. Another particular case of interest is the unitary free Brownian motion $U(t)$, which has a bounded density whenever $t > 0$ (see [Voi99, Proposition 1.6]): the free compressions $(U_{i,j}(t))_{i,j=1}^n$ of the unitary free Brownian motion have maximal $\Delta$ whenever $t > 0$.

Our main use of the previous corollary will be to model free variables in matrix forms. For this we use Haar unitaries, written in matrix form $(U_{ij})_{i,j=1}^n$ as above, where the
entries are given by free compressions of a Haar unitary \( U \). Note that this means that the matrix \( (U_{ij})_{i,j=1}^{n} \) is free from \( M_n(\mathbb{C}) \), see [NS06, Theorem 14.20]. We will use such Haar unitaries to make other matrices free among themselves by conjugating with the Haar unitaries. For this it will be crucial that such \( (U_{ij})_{i,j=1}^{n} \) are free from all other matrices where the entries are free from the \( U_{ij} \). This fact is provided by the following proposition. This can be derived as a direct consequence of Proposition 3.7 of [NSS02], but we prefer to give a self-contained proof which is more in line with our setting. We will only need the special case where \( \mathcal{A} \) is generated by a Haar unitary; for this case a direct proof can also be found in the recent preprint [MPS21, Theorem 17].

**Proposition 3.9.** Let \( (\mathcal{M}, \tau) \) be a tracial \( W^* \)-probability space and \( \mathcal{A} \) and \( \mathcal{B} \) be two \( * \)-free subalgebras of \( \mathcal{M} \).

If \( M_n(\mathcal{A}) \) is \( * \)-free from \( M_n(\mathcal{C}) \) in \( (M_n(\mathcal{M}), \text{tr}_n \otimes \tau) \), then \( M_n(\mathcal{A}) \) is also \( * \)-free from \( M_n(\mathcal{B}) \) in \( (M_n(\mathcal{M}), \text{tr}_n \otimes \tau) \).

**Proof.** Without loss of generality, we can replace \( \mathcal{B} \) by \( u\mathcal{B}u^* \) with \( u \) \( * \)-free from \( \mathcal{A} \) and try to prove that \( M_n(u\mathcal{B}u^*) = (I_n \otimes u)M_n(\mathcal{B})(I_n \otimes u)^* \) is \( * \)-free from \( M_n(\mathcal{A}) \).

We already know that \( M_n(\mathcal{A}) \) is \( * \)-free with amalgamation from \( (I_n \otimes u) \) with respect to the conditional expectation \( E_{M_n(\mathcal{C})} := \text{id} \otimes \tau \). However, \( E_{M_n(\mathcal{C})} \) is equal to \( \text{tr}_n \otimes \tau \) on \( M_n(\mathcal{A}) \) (because of the \( * \)-freeness of \( M_n(\mathcal{A}) \) from \( M_n(\mathcal{C}) \)) and \( E_{M_n(\mathcal{C})} \) is equal to \( \text{tr}_n \otimes \tau \) on \( W^*(I_n \otimes u) = I_n \otimes W^*(u) \). As a consequence, the \( * \)-freeness with amalgamation of \( M_n(\mathcal{A}) \) from \( (I_n \otimes u) \) with respect to \( E_{M_n(\mathcal{C})} \) implies the \( * \)-freeness with respect to \( \text{tr}_n \otimes \tau \).

Because we just proved that \( (I_n \otimes u) \) is a Haar unitary which is \( * \)-free from \( M_n(\mathcal{A}) \), we get the \( * \)-freeness of \( (I_n \otimes u)M_n(\mathcal{B})(I_n \otimes u)^* \) from \( M_n(\mathcal{A}) \).

\( \square \)

4. **Algebraic model for free variables with atoms**

In the first subsection, we will build an algebraic model for \( * \)-free normal random variables that have atoms of rational weights. The main argument we will use is that we can compare the free situation with matrices over the free skew field, by the use of the compressions from the previous section. This result is based on, but also largely extends, the main result in [NSY19], in the sense that now we can deal with variables with atoms of rational weights. In particular, we will establish a rank equality between the analytic objects and the algebraic objects. This allows us to describe the point spectrum distribution of rational expressions evaluated at normal variables with prescribed atoms. Moreover, this rank equality also explains the strong Atiyah conjecture for the free products of finite groups as the analytic rank is equal to some normalized inner rank.

As our algebraic model contains no information about the non-atomic parts of the given random variables, a natural consequence is that for a rational expression \( R \) and a tuple \( X = (X_1, \ldots, X_d) \) of \( * \)-free normal random variables such that \( R(X) \) is well-defined, the point spectrum distribution of \( R(X) \) does not depend on the non-atomic parts of \( X_1, \ldots, X_d \). In the second subsection we will show that this is even true if the weights of the atoms of \( X_1, \ldots, X_d \) are not necessarily rational; this will then give the full proof of Theorem [4.2]. This phenomenon was known for the polynomial \( P(x_1, x_2) = x_1 + x_2 \) according to [BV98, Theorem 7.4]. Our theorem pushes such a result to general rational functions. Moreover, the approximation to the general case actually allows us to calculate the point spectrum distribution.
4.1. Algebraic model for free variables with atoms of rational weights. Let \( X = (X_1, \ldots, X_d) \) be a tuple of \(*\)-free normal random variables in some \( W^\ast\)-probability space such that there exists \( n \in \mathbb{N} \) with \( \mu_{X_k}(\{\lambda\}) \in \frac{1}{n} \mathbb{N} \) for each \( 1 \leq k \leq d \) and each \( \lambda \in \mathbb{C} \). For each \( 1 \leq k \leq d \), the point spectrum distribution \( \mu_{X_k}^p \) of \( X_k \) can be written as

\[
\sum_{\ell=1}^{\ell_k} \frac{1}{n} \delta_{\alpha_\ell^{(k)}},
\]

where the multiset \( \{\alpha_\ell^{(k)}\}_{\ell \in \mathbb{N}} \subset \mathbb{C} \) is the set of atoms of \( X_k \).

Now, for each \( 1 \leq k \leq d \), we consider indeterminate formal variables \( u_{ij}^{(k)}_1 (1 \leq i, j \leq n) \) and \( y_\ell^{(k)} (1 \leq \ell \leq n - \ell_k) \). We denote by \( \mathcal{P} \) the unital complex algebra of noncommutative polynomials formed by these \( dn^2 + dn - \sum_k \ell_k \) variables and by \( \mathcal{R} \) its universal field of fractions. For each \( k = 1, \ldots, d \) we define the matrix

\[
\bar{U}_k := (u_{ij}^{(k)})_{i,j=1}^n \in M_n(\mathcal{P})
\]

and the diagonal matrix

\[
\bar{D}_k := \begin{pmatrix}
\alpha_1^{(k)} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \alpha_{\ell_k}^{(k)} \\
y_1^{(k)} & \cdots & 0 \\
0 & \ddots & \vdots \\
0 & \cdots & y_{n-\ell_k}^{(k)}
\end{pmatrix} \in M_n(\mathcal{P}).
\]

**Theorem 4.1.** Let \( X = (X_1, \ldots, X_d) \) be a tuple of \(*\)-free normal random variables in some \( W^\ast\)-probability space \( (\mathcal{M}, \tau) \) such that the weights of atoms are multiples of \( \frac{1}{n} \) for a certain \( n \in \mathbb{N} \). Let \( \bar{U}_k \) and \( \bar{D}_k \), \( k = 1, \ldots, d \) be defined as in (1) and (2). We set \( \bar{X}_k := \bar{U}_k \bar{D}_k \bar{U}_k^{-1} \in M_n(\mathcal{R}) \) and denote \( \bar{X} := (\bar{X}_1, \ldots, \bar{X}_d) \). Then, we have the following:

- For each integer \( m \) and \( A \in M_m(\mathbb{C} \langle x_1, \ldots, x_d \rangle) \), we have

\[
\text{rank}(A(X)) = \frac{1}{n} \rho_R(A(\bar{X})) \in [0, m] \cap \frac{1}{n} \mathbb{N}.
\]

- For any rational expression \( R \) in variables \( x_1, \ldots, x_d \), \( R(X) \) exists in \( L^0(\mathcal{M}) \) if and only if \( R(\bar{X}) \) exists in \( M_n(\mathcal{R}) \). Moreover, provided well-definedness, we have for any \( \lambda \in \mathbb{C} \),

\[
\mu^p_{R(X)}(\{\lambda\}) = 1 - \frac{1}{n} \rho_R(\lambda - R(\bar{X})) \in [0, 1] \cap \frac{1}{n} \mathbb{N}.
\]

**Proof.** The second item is a consequence of the first one, thanks to Corollary 2.22. So let us prove the first item.

Let \( U_1, \ldots, U_d \) be \(*\)-free Haar unitary random variables in some \( W^\ast\)-probability space \( (\mathcal{M}_1, \tau_1) \). Let \( U_{ij}^{(k)} \in \mathcal{N} := E_{11} \mathcal{M}_1 \ast M_n(\mathbb{C}) E_{11} \), \( 1 \leq i, j \leq n, 1 \leq k \leq d \) be their compressions by free matrix units as in Theorem 3.7. According to Corollary 3.8, \( U_{ij}^{(k)} \), \( 1 \leq i, j \leq n, 1 \leq k \leq d \) maximize their \( \Delta \).
Moreover, the matrix
\[
\hat{U}_k := (U_{ij}^{(k)})_{1 \leq i, j \leq n} \in (M_n(N), \text{tr}_n \circ \tau)
\]
is a Haar unitary random variable and thus \(\text{rank}(\hat{U}_k) = n\). By Theorem 2.31 we infer that \(\rho(\hat{U}_k) = \text{rank}(\hat{U}_k) = n\). So we see that \(\hat{U}_k\) is a full matrix and thus \(\hat{U}_k^{-1}\) is matrix over \(\mathcal{R}\).
Therefore the evaluation of \(\hat{U}_k^{-1}\) is exactly the operator \(\hat{U}_k^* = (\hat{U}_k)^{-1}\).
Next, thanks to Lemma 3.1 we know that the distribution of \(\hat{U}_k\) is a Haar unitary random variable and thus \(\text{rank}(\hat{U}_k) = n\). By Theorem 2.31 we infer that \(\rho(\hat{U}_k) = \text{rank}(\hat{U}_k) = n\). So we see that \(\hat{U}_k\) is a full matrix and thus \(\hat{U}_k^{-1}\) is matrix over \(\mathcal{R}\).

Therefore the evaluation of \(\hat{U}_k^{-1}\) is exactly the operator \(\hat{U}_k^* = (\hat{U}_k)^{-1}\).

We consider now the \(\ell\)-free tuple \(U^{(k)} = (\gamma_1^{(k)}, \ldots, \gamma_{n-\ell_k}^{(k)})\) of \(\mathcal{R}\)-free normal random variables and then we define the matrix
\[
D_k := \begin{pmatrix}
\alpha_{1}^{(k)} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \alpha_{\ell_k}^{(k)} \\
\gamma_{1}^{(k)} & \cdots & 0 \\
0 & \vdots & \ddots \\
0 & \cdots & \gamma_{n-\ell_k}^{(k)}
\end{pmatrix}
\]
where \(\gamma_{1}^{(k)}, \ldots, \gamma_{n-\ell_k}^{(k)}\) are normal random variables without atoms which are \(\ast\)-free in some tracial \(W^\ast\)-probability space. Because \(\gamma_{1}^{(k)}, \ldots, \gamma_{n-\ell_k}^{(k)}\) are without atoms, we have \(\Delta(\gamma_{1}^{(k)}), \ldots, \gamma_{n-\ell_k}^{(k)}\) have the same distribution as the original tuple \(\gamma_{1}^{(k)}, \ldots, \gamma_{n-\ell_k}^{(k)}\).

We further consider the \(W^\ast\)-probability space \((N * M_2, \tau * \tau_2)\), in which we have
\[
\Delta(U_{ij}^{(k)}, \gamma_{\ell}^{(k)} : 1 \leq k \leq d, 1 \leq i, j \leq n, 1 \leq \ell \leq n - \ell_k) = dn - d \sum_{k=1}^{d} \ell_k.
\]
Let us further consider the \(W^\ast\)-probability space \((N * M_2, \tau * \tau_2)\), in which we have
\[
\Delta(U_{ij}^{(k)}, \gamma_{\ell}^{(k)} : 1 \leq k \leq d, 1 \leq i, j \leq n, 1 \leq \ell \leq n - \ell_k) = dn^2 + dn - \sum_{k=1}^{d} \ell_k.
\]

Thanks to Remark 2.32, this maximality of \(\Delta\) gives us an isomorphism \(u_{ij}^{(k)} \mapsto U_{ij}^{(k)}, y_{\ell}^{(k)} \mapsto \gamma_{\ell}^{(k)}\) from the free field \(\mathcal{R}\) to its image such that the respective ranks are preserved.

Now let us define
\[
\hat{X}_k := \hat{U}_k D_k \hat{U}_k^* \in (M_n(N * M_2), \text{tr}_n \circ (\tau * \tau_2)) \quad \forall k = 1, \ldots, d,
\]
and denote \(\hat{X} := (\hat{X}_1, \ldots, \hat{X}_d)\). Then we have for any \(A \in M_\binom{n}{2}(\mathbb{C})\)
\[
\text{rank}(A(\hat{X})) = \rho_\mathcal{R}(A(\hat{X})).
\]
Hence the first part of the desired result follows if \(\text{rank}(A(\hat{X})) = \frac{1}{n}\text{rank}(A(\hat{X}))\). Such a rank equality naturally holds if \(\hat{X}\) has the same distribution as the tuple \(X\).

Let us verify that \(\hat{X}\) has the same distribution as the original tuple \(X\). Note that for each \(k = 1, \ldots, d\), \(\hat{X}_k\) has the same distribution as \(X_k\) and \(\hat{U}_1, \ldots, \hat{U}_d\) are freely independent Haar unitary random variables. So it suffices to show that \(\hat{U}_1, \ldots, \hat{U}_d\) are \(\ast\)-free from \(D_1, \ldots, D_d\). Since the family \(\{\hat{U}_1, \ldots, \hat{U}_d\}\) is \(\ast\)-free from \(M_n(\mathbb{C})\) (which can be seen with the help of the identification \((M_1, \tau_1) * (M_n(\mathbb{C}), \text{tr}_n) \simeq (M_n(N), \text{tr}_n \circ \tau)\) and
the entries of $\hat{U}_1, \ldots, \hat{U}_d$ are $*$-free from the entries of $D_1, \ldots, D_d$, the desired result is a consequence of the Proposition 3.9.

**Remark 4.2.** Note that the requirement on the normality in the above theorem comes from the requirement on the normality for modelling a tuple of $*$-free random variables by their matrix model (see Lemma 3.7). So actually we can also build an algebraic model mutatis mutandis if

1. we put $Y_1, \ldots, Y_d \in M_n(M_2)$ for some integer $n$ and $W^*$-probability space $(M_2, \tau_2)$ such that $\Delta(\gamma_1, \ldots, \gamma_\ell) = 1$ (but $\gamma_1, \ldots, \gamma_\ell$ and $Y_1, \ldots, Y_d$ may not necessarily be normal);

2. for $k = 1, \ldots, d$, we set $\hat{X}_k := \hat{U}_k Y_k \hat{U}_k^*$, where $\hat{U}_k$ is defined as in Theorem 4.1.

In particular, we can also set $\ell = 0$, i.e.,

1. $Y_1, \ldots, Y_d \in M_n(\mathbb{C})$;

2. for $k = 1, \ldots, d$, $\hat{X}_k := \hat{U}_k Y_k \hat{U}_k^*$, where $\hat{U}_k$ is defined as in Theorem 4.1.

Then an algebraic model can be built mutatis mutandis.

Furthermore, we can actually model the free product of algebras rather than individual variables if these algebras are matrix algebras over some variables that have maximal $\Delta$. We record this observation as the following proposition.

**Proposition 4.3.** Let $A_1, \ldots, A_d$ be subalgebras of $M_n(\mathbb{C}(\gamma_1, \ldots, \gamma_\ell))$, where $\ell$ is an integer and $\gamma_1, \ldots, \gamma_\ell$ are in some $W^*$-probability space $(\mathcal{M}_2, \tau_2)$ such that $\Delta(\gamma_1, \ldots, \gamma_\ell) = 1$. We denote by $\Phi : \mathbb{C}(y_1, \ldots, y_\ell) \to \mathbb{C}(\gamma_1, \ldots, \gamma_\ell)$ the isomorphism determined by $\Phi(y_i) = \gamma_i$, $i = 1, \ldots, \ell$, where $y_i$'s are formal variables. For $k = 1, \ldots, d$, we set $\hat{A}_k := \hat{U}_k \Phi^{-1}(A_k) \hat{U}_k^{-1}$ which are subalgebras of $M_n(\mathbb{R})$, where $\mathbb{R}$ and $\hat{U}_k$ are defined as in Theorem 4.1. Then $*_{k=1}^d \hat{A}_k$ is isomorphic to the subalgebra generated by $\hat{A}_1, \ldots, \hat{A}_d$ with rank preserved.

**Proof.** For $k = 1, \ldots, d$, we set $\hat{A}_k := \hat{U}_k A_k \hat{U}_k^*$ which are subalgebras of $M_n(\mathbb{N} * \mathcal{M}_2)$, where $\mathbb{N}$ and $\hat{U}_k$ are defined as in Theorem 4.1. For each $k = 1, \ldots, d$, let $\Psi_k : A_k \to M_n(\mathbb{N} * \mathcal{M}_2)$ be the homomorphism determined by

$$\Psi_k(A) := \hat{U}_k A \hat{U}_k^*, \quad \forall A \in A_k.$$ 

Clearly the images $\Psi_1(A_1), \ldots, \Psi_d(A_d)$ are $*$-free in $(M_n(\mathbb{N} * \mathcal{M}_2), \text{tr}_n \circ (\tau * \tau_2))$. Thus $*_{k=1}^d A_k$ is isomorphic to the subalgebra generated by $\Psi_1(A_1), \ldots, \Psi_d(A_d)$ with preserved ranks. Finally, the isomorphism with preserved ranks between the subalgebra generated by $\Psi_1(A_1), \ldots, \Psi_d(A_d)$ and the subalgebra generated by $A_1, \ldots, A_d$ can be seen by a similar argument as in Theorem 4.1. □

### 4.2. Atiyah conjecture and division closure

In this subsection, let us briefly discuss the link between our results obtained so far and the Atiyah conjecture.

First, let us explain the consequence of the our results for the Atiyah conjecture of discrete groups. For a group $G$, we denote by $\text{lcm}(G)$ the least common multiple of the orders of finite subgroups of $G$, if it exists (with a convention that $\text{lcm}(G) = 1$ if $G$ is torsion-free). For a group $G$ with $\text{lcm}(G) < \infty$, we say that $G$ satisfies the strong Atiyah...
conjecture (see, for example, [Luc02, Conjecture 10.2]) if for every \( m \in \mathbb{N} \) and every matrix \( A \in M_m(\mathbb{C}G) \) we have
\[
\text{rank} A \in \frac{1}{\text{lcm}(G)} \mathbb{Z}.
\]

Let us consider finite groups \( G_1, \ldots, G_d \) with orders \( n_1, \ldots, n_d \) respectively. For each \( k = 1, \ldots, d \), we regard the group algebra \( \mathbb{C}G_k \) as the subalgebra of \( M_{n_k}(\mathbb{C}) \) via the left-regular representation. In order to apply Proposition 4.3, we need to put \( \mathbb{C}G_1, \ldots, \mathbb{C}G_d \) (considered as matrix subalgebras) into a matrix algebra \( M_n(\mathbb{C}) \) for some integer \( n \). Naturally, the smallest convenient choice for this \( n \) is \( \text{lcm}(n_1, \ldots, n_d) \). Hence we can have the following corollary, thanks to Proposition 4.3.

**Corollary 4.4.** Let \( G_1, \ldots, G_d \) be finite groups with orders \( n_1, \ldots, n_d \). Then their free product \( G := G_1 * G_2 * \cdots * G_d \) satisfies the strong Atiyah conjecture. Moreover, for any matrix \( A \in M_m(\mathbb{C}G) \), \( \text{rank} A \) is equal to the normalized inner rank of some \( nm \times nm \) matrix over some free field \( \mathcal{R} \) (as constructed in Proposition 4.3).

Actually, it is known that the strong Atiyah conjecture is stable with respect to finite free products (see, for example, [Rei98, Theorem 7.7]). Therefore our corollary can be understood as a free probabilistic proof or explanation of the strong Atiyah conjecture for free products of finite groups.

Let us make a further remark on the rational/division closure of \( \mathbb{C}G \). We actually have the isomorphism between the rational/division closure of \( \mathbb{C}G \) in the algebra of affiliated operators with respect to the group von Neumann algebra \( \mathcal{L}(G) \) and the rational/division closure of the subalgebra generated by \( \overline{\mathbb{C}G_k} := \overline{U_k \mathbb{C}G_k U_k^{-1}} \) (as defined in Proposition 4.3) due to the rank equality. This can be seen similarly with the help of proof of Corollary 2.22. So we see that the division closure of \( \mathbb{C}G \) is a subalgebra of a matrix algebra over some free field. This observation seems to be related to the known result that if a group \( G \) satisfies the strong Atiyah conjecture, then its division closure is some direct sum of finite matrix algebras over skew fields (see, for example, [KLS17]).

Finally, let us discuss a bit the general case that may not come from groups. In [SS15], a similar notion was introduced under the name of Atiyah property in order to include random variables that may not come from groups. Let us introduce this property with some slight modification.

**Definition 4.5.** Let \( \mathcal{A} \) be a \( * \)-subalgebra of a tracial \( W^* \)-probability space \((\mathcal{M}, \tau)\) and let \( \Gamma \) be an additive subgroup of \( \mathbb{R} \) containing \( \mathbb{Z} \). We say that \( \mathcal{A} \) satisfies the Atiyah property with \( \Gamma \) if for every \( m \in \mathbb{N} \) and every matrix \( A \in M_m(\mathcal{A}) \) we have that \( \text{rank} A \in \Gamma \).

Then we have the following as a corollary of Theorem 4.1.

**Corollary 4.6.** Let \( X = (X_1, \ldots, X_d) \) be a tuple of \( * \)-free normal random variables with compactly supported probability measures \( \mu_k \) as distribution, respectively. Suppose that for each \( k = 1, \ldots, d \), there exists \( n_k \in \mathbb{N} \) such that all the atoms of \( \mu_k \) have weights contained in \( \frac{1}{n_k} \mathbb{N} \). Then \( \mathbb{C}(X) \) has the Atiyah property with group \( \frac{1}{n} \mathbb{N} \), where \( n := \text{lcm}(n_1, \ldots, n_d) \).

Recall that the result in [SS15, Theorem 3.4] says that \( \mathbb{C}(X) \) has the Atiyah property with group \( \frac{1}{n} \mathbb{N} \), where \( n := n_1 \cdots n_d \). We improve the denominator for the weight of the
atoms from the product $n_1 \cdots n_d$ to their least common multiple $\text{lcm}(n_1, \ldots, n_d)$. Furthermore, we also improve their result in the sense that we can extract precise information on the positions and the weights of atoms via the calculation of the inner rank of matrices in formal variables.

For a pair of self-adjoint free variables $X = (X_1, X_2)$, Corollary 4.6 has already been established in [BBL21], as the following result shows. Interestingly, this result is also valid for free variables with atoms of irrational weights. In this respect, it is much more general than [SS15, Theorem 3.4] or Corollary 4.6. We are not able to attain this level of generality with our approach because we did not succeed to obtain the Atiyah property via an approximation procedure.

**Theorem 4.7** (Corollary 3.3 of [BBL21]). Let $X = (X_1, X_2)$ be a tuple of free self-adjoint random variables with compactly supported probability measures $\mu_1, \mu_2$ as distribution, respectively. Suppose there exists an additive subgroup $\Gamma$ of $\mathbb{R}$ containing $\mathbb{Z}$ such that all the atoms of $\mu_1$ and $\mu_2$ have weights contained in $\Gamma$. Then $\mathbb{C}(X)$ has the Atiyah property with group $\Gamma$.

**Proof.** Because [BBL21] Corollary 3.3 is not stated in this terms, let us give some details. Let $A \in M_m(\mathbb{C}(X))$. The matrix $A$ can be written as a polynomial expression in $I_m \otimes X_1$, $I_m \otimes X_2$ and some matrices $M \subset M_m(\mathbb{C})$. That is, let $R$ be a rational expression of formal variables $x_1, x_2$ as well as some variables corresponding to the finite set $M$ such that $R(I_m \otimes X_1, I_m \otimes X_2, M) = A$. Now let $(u, Q, v)$ be a formal linear representation of $R$, of dimension $n$ and with display $L$ in the sense of Definition 2.16. Thanks to Lemma 2.20 we have

$$\text{rank}(A) = \text{rank}(R(I_m \otimes X_1, I_m \otimes X_2, M)) = \text{rank}(L(I_m \otimes X_1, I_m \otimes X_2, M)) - mn.$$  

The matrix $L(I_m \otimes X_1, I_m \otimes X_2, M) \in M_{mn}(\mathbb{C}(X))$ being linear, [BBL21] Corollary 3.3] allows to say that its rank is an affine combination (with integer coefficients) of the weights of the atoms of $\mu_1$ and $\mu_2$. As a consequence, the rank of $L(I_m \otimes X_1, I_m \otimes X_2, M) \in M_{mn}(\mathbb{C}(X))$ belongs to $\Gamma$, and so does the rank of $A$. \hfill $\square$

### 4.3. Approximation for general free normal variables.

From Theorem 4.1 we see that the position and the size of an atom of $R(X_1, \ldots, X_d)$ do not depend on the non-atomic parts of the marginal distributions, if we assume that the variables $X_i$’s have rational weights for their atoms. Let us remove in the following this hypothesis on rational weights. Theorem 1.2 is the second item in the following theorem.

**Theorem 4.8.** Let $X = (X_1, \ldots, X_d)$ and $Y = (Y_1, \ldots, Y_d)$ be two tuples of $\ast$-free normal variables in tracial $W^*$-probability spaces such that, for each $1 \leq i \leq d$, we have $\mu_{X_i}^p = \mu_{Y_i}^p$. Then we have the following statements.

- **For each integer $m$ and each $A \in M_m(\mathbb{C}(x_1, \ldots, x_d))$, we have**

  $$\text{rank}(A(X)) = \text{rank}(A(Y)).$$

- **For each rational expression $R$, $R(X)$ is well-defined if and only if $R(Y)$ is well-defined, and in this case,**

  $$\mu_{R(X)}^p = \mu_{R(Y)}^p.$$
**Proof.** The second item is a consequence of the first one, thanks to Corollary 2.22. So let us prove the first item.

Thanks to Lemma 2.23, we can consider, for each $N$, two tuples of $*$-free normal random variables $X^{(N)} = (X_1^{(N)}, ..., X_d^{(N)})$ and $(Y_1^{(N)}, ..., Y_d^{(N)})$ such that:

- $\text{rank}(X_i^{(N)} - X_i) \leq C_X/N$ and $\text{rank}(Y_i^{(N)} - Y_i) \leq C_Y/N$, for all $i$;
- for all $\lambda \in \mathbb{C}$, we have $\mu_{X_i^{(N)}}(\{\lambda\}) = \mu_{Y_i^{(N)}}(\{\lambda\})$;
- the weights of the atoms of $X_k^{(N)}$ and $Y_k^{(N)}$ are rational.

Lemma 2.23 says that there exists some constant $c > 0$ such that

$$|\text{rank}(A(X) - A(X^{(N)}))| \leq c/N \quad \text{and} \quad |\text{rank}(A(Y) - A(Y^{(N)}))| \leq c/N.$$ 

Due to the subadditivity of rank, we have

$$|\text{rank}(A(X)) - \text{rank}(A(X^{(N)}))| \leq c/N \quad \text{and} \quad |\text{rank}(A(Y)) - \text{rank}(A(Y^{(N)}))| \leq c/N.$$ 

Moreover, according to Theorem 4.1, we have

$$\text{rank}(A(X^{(N)})) = \text{rank}(A(Y^{(N)})).$$

Hence we conclude that

$$|\text{rank}(A(X)) - \text{rank}(A(Y))| \leq 2c/N.$$ 

Letting $N$ tend to $\infty$, we get that

$$\text{rank}(A(X)) = \text{rank}(A(Y)).$$

\[\Box\]

Recall that Remark 2.9 says that the rational closure of a tuple $X = (X_1, \ldots, X_d)$ of random variables only depends on the $*$-distribution of $X$. In the case that $X$ is a tuple of free random variables, then this can be further improved in the sense that the rational closure of $X$ only depends on the point spectrum distribution of each variable $X_i$. That is, with the help of Corollary 2.22, we get the following corollary.

**Corollary 4.9.** Let $X = (X_1, \ldots, X_d)$ and $Y = (Y_1, \ldots, Y_d)$ be two $d$-tuples of $*$-free normal variables in tracial $W^*$-probability spaces such that, for each $1 \leq i \leq d$, we have $\mu_{X_i}^p = \mu_{Y_i}^p$. Then the rational closure of $X$ is isomorphic to the rational closure of $Y$.

We know that for the polynomial $P(x_1, x_2) = x_1 + x_2$, [BV98, Theorem 7.4] gave a complete description of the point spectrum distribution of $P(X_1, X_2)$ for any freely independent self-adjoint random variables with prescribed atoms. This result was recently promoted to a matrix-version in [BBL21] — this gives, via linearization, a possible analytic route for dealing with the point spectrum distribution for arbitrary polynomials in free variables.

In contrast to this analytic approach, we are able to provide here an algebraic method for the calculation of the point spectrum distribution. Of course, since we need to start with random variables with atoms of rational weights, we need to control the approximation process. First, notice that in our approximation scheme (Lemma 2.23) we have $\mu_{X_i}^p \leq \mu_{X_i}^p$, $i = 1, \ldots, d$. This means that for a rational expression $R$ for which $R(X)$ is well-defined, all the $R(X^{(N)})$ are actually also well-defined, by Theorem 5.1. Furthermore,
we have seen that \(\text{rank}(R(X^N))\) converges to \(R(X)\). Let us record this observation as a remark.

**Remark 4.10.** Let \(X = (X_1, \ldots, X_d)\) be a tuple of \(*\)-free normal variables in some tracial \(W^*\)-probability space. Then we can construct a sequence \(X^{(N)} = (X_1^N, \ldots, X_d^N)\) of tuples of \(*\)-free normal variables such that:

- each atom of \(\mu_{X_k}^p\) has a rational weight and \(\mu_{X_k}^p \leq \mu_{Y_k}^p\);
- for each matrix \(A\) over \(\mathbb{C}[x_1, \ldots, x_d]\) we have
  \[
  \text{rank}(A(X)) = \lim_{N \to \infty} \text{rank}(A(X^{(N)})),
  \]
  where \(\text{rank}(A(X^{(N)}))\) can be calculated algebraically as in Theorem 4.7;
- for each rational expression \(R\) such that \(R(X)\) is well-defined, we have
  \[
  \text{rank}(R(X)) = \lim_{N \to \infty} \text{rank}(R(X^{(N)})),
  \]
  where \(R(X^{(N)})\) is well-defined and \(\text{rank}(R(X^{(N)}))\) can be calculated algebraically as in Theorem 4.7.

As we will show in Section 6 and 7, this allows us to extend our results to arbitrary weights. Thus, in principle, we can calculate atoms of any polynomials (even rational functions) by doing algebraic inner rank calculations in the free field, according to Theorem 4.1. This calculation of the inner rank for a given polynomial can, for example, be done by a noncommutative Gauss elimination (that is, a diagonalization by row and column operators). Alternatively, we can also relate the calculation of the inner rank of a given polynomial to the calculation of the inner rank of its display (thanks to Lemma 2.20). The calculation of this inner rank might then be carried out with the help of various other means, see, for example, the notions and tools in [GGOW16]. However, whereas all those algebraic tools for the calculation of inner ranks as well as the analytic approach from [BBL21] have great conceptual power and can be used to develop algorithmic or numerical approaches for doing calculations in concrete examples, it is very hard to get general explicit results from there for the size of atoms for bigger classes of polynomials.

In our approach, on the other side, we can make this additional step, since our representation of freely independent random variables via elements in the free field allows us the compare the free situation with other situations (see Section 5), resulting in, often quite sharp, upper and lower bounds for the weights of atoms for general classes of polynomials.

5. **Comparison of atoms for rational functions: free variables versus general variables**

The main goal of this section is to prove the following theorem.

**Theorem 5.1.** Let \(Y = (Y_1, \ldots, Y_d)\) be a tuple of normal random variables in some tracial \(W^*\)-probability space. Let \(X = (X_1, \ldots, X_d)\) be a tuple of \(*\)-free normal random variables in some tracial \(W^*\)-probability space such that, for each \(k = 1, \ldots, d\), we have

\[
\mu_{X_k}^p \leq \mu_{Y_k}^p.
\]

Then we have the following.
For any $A \in M_n(\mathbb{C} \langle x_1, \ldots, x_d \rangle)$, we have
\[ \text{rank}(A(Y)) \leq \text{rank}(A(X)). \]

(2) For any rational expression $R$ in variables $x_1, \ldots, x_d$ such that $R(Y)$ is well-defined as an unbounded operator, $R(X)$ is also well-defined and
\[ \mu^p_{R(X)} \leq \mu^p_{R(Y)}. \]

In Subsection 5.2 we apply this theorem to deduce the universality property of the rational closures given by $*$-free normal random variables and in Section 5.3 we consider matrix approximations.

Theorem 5.1 shows that $*$-free variables realize an extreme situation for tuples of random variables with prescribed marginals: atoms of polynomials can not be smaller than the unavoidable sizes at the unavoidable locations of the $*$-free case. It is tempting to wonder how generic is this situation: in other words, how often do the inequalities in Theorem 5.1 hold to be equalities? Following the work of [BCD+10], where this genericity aspects are under consideration, one could attack this question by rotating each variable by unitary conjugation, and asks about the property of the set of unitaries realizing equalities in Theorem 5.1: is it norm-dense, is it a $G_\delta$-set? We leave those questions for further investigations.

5.1. Proof of Theorem 5.1. First we prove an intermediate step for Theorem 5.1. We start with a slightly different setting, where the analytic distribution of each random variable has finitely many atoms of rational weights.

Lemma 5.2. Let $Y = (Y_1, \ldots, Y_d)$ be a tuple of normal random variables in some $W^*$-probability space $(\mathcal{M}, \tau)$. Let $X = (X_1, \ldots, X_d)$ be a tuple of $*$-free normal random variables satisfying the following conditions.

- For each $k = 1, \ldots, d$, $X_k$ has finitely many atoms and each atom is of rational weight.
- For each $k = 1, \ldots, d$, we have $\mu^p_{X_k} \leq \mu^p_{Y_k}$.

Then we have the following.

(1) For any $A \in M_n(\mathbb{C} \langle x_1, \ldots, x_d \rangle)$, we have
\[ \text{rank}(A(Y)) \leq \text{rank}(A(X)). \]

(2) For any rational expression $R$ in variables $x_1, \ldots, x_d$ such that $R(Y)$ is well-defined as an unbounded operator, $R(X)$ is also well-defined and
\[ \mu^p_{R(X)} \leq \mu^p_{R(Y)}. \]

Proof. The second item is a consequence of the first one, thanks to Proposition 2.21. Let us prove the first item.

First, let us replace $X$ in this equality by $\bar{X}$, which is an algebraic model for $X$. According to the assumption there exists $r \in \mathbb{N}$ such that $\mu_{X_k}(\{\lambda\}) \in \frac{1}{r}\mathbb{N}$ for each $1 \leq k \leq d$ and $\lambda \in \mathbb{C}$. That is, for each $1 \leq k \leq d$, the point spectrum distribution $\mu^p_{X_k}$ of $X_k$ can be written as
\[ \mu^p_{X_k} = \sum_{\ell=1}^{\ell_k} \frac{1}{r} \delta_{\alpha_k^{(\ell)}}. \]
where the multiset \( \{ \alpha_{\ell}^{(k)} \} \), \( \ell \subset \mathbb{C} \) is the set of atoms of \( X_k \). For each \( 1 \leq k \leq d \), we consider variables \( u_{ij}^{(k)} \) (\( 1 \leq i, j \leq r \)) and \( y_{\ell}^{(k)} \) (\( 1 \leq \ell \leq r - \ell_k \)). We denote by \( \mathcal{P} \) the polynomial ring formed by these \( dr^2 + dr - \sum_\ell \ell_k \) variables and by \( \mathcal{R} \) its universal field of fractions. For each \( k = 1, \ldots, d \) we define the matrix

\[
\hat{U}_k := (u_{ij}^{(k)})_{i,j=1}^r \quad \text{and} \quad \hat{D}_k := \text{diag}(\alpha_1^{(k)}, \ldots, \alpha_{\ell_k}^{(k)}, y_1^{(k)}, \ldots, y_{r-\ell_k}^{(k)}) \in M_r(\mathcal{P}).
\]

as in (1) and (2), and \( \bar{X}_k := U_k \hat{D}_k \hat{U}_k^{-1} \) as in Theorem 4.1. For simplicity, let us denote \( \hat{U} = (U_1, \ldots, U_d) \), \( \hat{D} = (\hat{D}_1, \ldots, \hat{D}_d) \) and \( \bar{X} = (\bar{X}_1, \ldots, \bar{X}_d) \).

Thanks to Theorem 4.1 we have

\[
\text{rank}(A(X)) = \frac{1}{\rho \mathcal{R}(A(\bar{X}))}.
\]

Next, we regard \( A(\bar{X}) \) as a rational expression in \( I_n \otimes \hat{U}_1, \ldots, I_n \otimes \hat{U}_d \) and \( I_n \otimes \hat{D}_1, \ldots, I_n \otimes \hat{D}_d \) and some matrices \( M \subset M_n(\mathbb{C}) \otimes I_r \) (since \( A \) is a polynomial in \( I_n \otimes x_1, \ldots, I_n \otimes x_d \) and some matrices in \( M_n(\mathbb{C}) \otimes I_r \)). That is, let \( \mathcal{R} \) be a rational expression of formal variables \( u_1, \ldots, u_d \) and \( z_1, \ldots, z_d \) as well as some variables corresponding to the finite set \( M \) such that \( \mathcal{R}(I_n \otimes \bar{U}, I_n \otimes \bar{D}, M) = A(\bar{X}) \). Now let \( (u, Q, v) \) be a formal linear representation of \( R \), of dimension \( m \) and with display \( L \) in the sense of Definition 2.16. Thanks to Lemma 2.20 we have

\[
\mathcal{R}(\bar{X}) = \mathcal{R}(\hat{U} \hat{D}_k \hat{U}_k^{-1}) = \mathcal{R}(\hat{D}_k \mathcal{R}(\hat{U} \hat{U}_k^{-1})) = \rho \mathcal{R}(A(\bar{X})),
\]

we define the matrix \( A(\bar{X}) \) as \( \rho \mathcal{R}(A(\bar{X})) \). Thus for each \( 1 \leq k \leq d \), we can write \( I_r \otimes Y_k \in M_r(\mathcal{M}) \) as \( I_r \otimes Y_k = U_k \hat{D}_k \hat{U}_k^{-1} \) where \( U_k \in M_r(\mathcal{M}) \) is a unitary matrix and

\[
D_k = \text{diag}(\alpha_1^{(k)}, \ldots, \alpha_{\ell_k}^{(k)}, \gamma_1^{(k)}, \ldots, \gamma_{r-\ell_k}^{(k)}) \in M_r(\mathcal{M}).
\]

We have

\[
\text{rank}(A(Y)) = \frac{1}{\rho \mathcal{R}(A(\bar{X}))}.
\]

We denote \( U = (U_1, \ldots, U_d) \) and \( D = (\hat{D}_1, \ldots, \hat{D}_d) \). Similar, with the help of Lemma 2.20 if \( A(I_r \otimes Y) = R(U, D, M) \) is well-defined, then

\[
\text{rank}(A(I_r \otimes Y)) = \text{rank}(R(I_n \otimes U, I_n \otimes D, M)) = \text{rank}(L(I_n \otimes U, I_n \otimes D, M)) - mnr.
\]

and so the desired claim follows if

\[
\text{rank}(L(I_n \otimes U, I_n \otimes D, M)) \leq \rho \mathcal{R}(L(I_n \otimes U, I_n \otimes D, M)).
\]

In order to see this let us denote by \( \Phi : \mathcal{P} \rightarrow \mathcal{M} \) the unital ring homomorphism given by extending \( \Phi(u_{ij}^{(k)}) = U_{ij}^{(k)} \) and \( \Phi(y_{\ell}^{(k)}) = \gamma_{\ell}^{(k)} \). So \( \Phi(\hat{U}_k) = U_k \), \( \Phi(\hat{D}_k) = D_k \) and thus it boils down to prove that

\[
\text{rank}(\Phi(L(I_n \otimes U, I_n \otimes D, M))) \leq \rho \mathcal{R}(L(I_n \otimes U, I_n \otimes D, M)).
\]

Note that \( \Phi \) extends to a homomorphism from \( M_{mnr}(\mathcal{P}) \) to \( M_{mnr}(\mathcal{M}) \) and thus for any matrix \( A \) over \( \mathcal{P} \),

\[
\rho \mathcal{M}(\Phi(A))) \leq \rho \mathcal{R}(A))
\]

Therefore, finally, the proof is completed by the fact

\[
\text{rank}(\Phi(A)) \leq \rho \mathcal{M}(\Phi(A))
\]
Proof of Theorem 5.1. The second item is a consequence of the first one, thanks to Proposition 2.21.

Using Lemma 2.28, we can consider, for each integer $N$, a tuple of $\ast$-free normal random variables $X^{(N)} = (X_1^{(N)}, \ldots, X_d^{(N)})$ such that for $k = 1, \ldots, d$

- $\text{rank}(X_k^{(N)} - X_k) \leq C_X/N$;
- for all $\lambda \in \mathbb{C}$, we have $\mu_{X_k^{(N)}}(\{\lambda\}) \leq \mu_{X_k}(\{\lambda\})$;
- the atoms of $X_k^{(N)}$ are finite and have rational weights.

For any matrix $A$ over polynomials $\mathbb{C}\langle x_1, \ldots, x_d \rangle$, Lemma 5.2 yields

$$\text{rank}(A(Y)) \leq \text{rank}(A(X^{(N)})),$$

and Lemma 2.23 says that

$$|\text{rank}(A(X)) - \text{rank}(A(X^{(N)}))| \leq c/N,$$

for some constant $c > 0$. This means that

$$\text{rank}(A(Y)) \leq \text{rank}(A(X)) + c/N.$$

Letting $N$ tend to $\infty$, we get that $\text{rank}(A(Y)) \leq \text{rank}(A(X))$. □

5.2. Universality of the free case. Here, we want to show that the rational closure of a free tuple of random variables with prescribed point spectrum distributions is the most general one in the sense that it satisfies a universal property.

In Subsection 2.3, we defined the category of $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebras whose arrows are given by specializations. Let $m_1, \ldots, m_d$ be $d$ sub-probability measures. We consider the sub-category $\mathcal{C}_m$ of $\mathbb{C}\langle x_1, \ldots, x_d \rangle$-algebras given by rational closures of $d$-tuples $X = (X_1, \ldots, X_d)$ of normal random variables in tracial $W^*$-probability spaces, such that, for all $1 \leq k \leq d$, we have

$$m_k \leq \mu_{X_k}^p.$$

The rational closure of a free-tuple is universal in the sense that it is an initial object in the category $\mathcal{C}_m$, as shown in the following corollary.

Corollary 5.3. Let $X = (X_1, \ldots, X_d)$ be a tuple of $\ast$-free normal random variables such that, for each $k = 1, \ldots, d$, $m_k = \mu_{X_k}^p$. Let $Y = (Y_1, \ldots, Y_d)$ be a tuple of normal random variables in some $W^*$-probability space $(\mathcal{M}, \tau)$ such that, for all $1 \leq k \leq d$, we have

$$m_k \leq \mu_{Y_k}^p.$$

Then, there exists a specialization from the rational closure of $X$ to the rational closure of $Y$, and in particular, the rational closure of $Y$ is a quotient of a subalgebra of the rational closure of $X$.

Proof. It is a direct consequence of Proposition 2.21 and Theorem 5.1. □

As pointed out in Remark 4.2, the normality required in the above Corollary is due to the way how we model the random variables $X_1, \ldots, X_d$. So in the simple case that $X_1, \ldots, X_d$ are free matrices in $\bigotimes_{i=1}^d M_n(\mathbb{C})$, we can remove the normality requirement on $X_1, \ldots, X_d$ (and thus the normality of $Y_1, \ldots, Y_d$ is also not necessary).
5.3. Approximation with matrices. Let \( X = (X_1, \ldots, X_d) \) be a tuple of \(*\)-free normal random variables. (Again, as pointed out in Remark 4.2, normality can be removed if \( X_1, \ldots, X_d \) are already in the matrix form, for example, if they are free matrices in \( *_{i=1}^d M_n(\mathbb{C}) \).) We define for each \( \varepsilon > 0 \)

\[
\mathcal{X}_{m,\varepsilon} := \{ Y = (Y_1, \ldots, Y_d) \in M_m(\mathbb{C})^d : \mu_{X_k}(\lambda) - \varepsilon \leq \mu_{Y_k}(\lambda), \forall \lambda \in \mathbb{C} \}
\]

and \( \mathcal{X}_\varepsilon := \prod_{m=1}^\infty \mathcal{X}_{m,\varepsilon} \). Then, thanks to Theorem 5.1 for each polynomial \( P \in \mathbb{C}\langle x_1, \ldots, x_d \rangle \)

\[
\mu_{P(X)} \leq \inf_{Y \in \mathcal{X}_\varepsilon} \mu_{P(Y)}
\]

where \( X_\varepsilon \) is a copy of \( X \) with moving a \( \varepsilon \) weight from each atom of \( X_k \), \( k = 1, \ldots, d \) (as in Lemma 2.28). Then by an application of Theorem 2.30 we can see that

\[
\mu_{P(X)} \leq \liminf_{\varepsilon \to 0} \inf_{Y \in \mathcal{X}_\varepsilon} \mu_{P(Y)}
\]

Now, we prove that this inequality actually is an equality.

**Proposition 5.4.** Let \( X = (X_1, \ldots, X_d) \) be a tuple of \(*\)-free normal random variables. Then for any \( P \in \mathbb{C}\langle x_1, \ldots, x_d \rangle \)

\[
\mu_{P(X)} = \liminf_{\varepsilon \to 0} \inf_{Y \in \mathcal{X}_\varepsilon} \mu_{P(Y)}
\]

**Proof.** We only need to prove that for every \( \varepsilon > 0 \), \( \inf_{Y \in \mathcal{X}_\varepsilon} \mu_{P(Y)} \leq \mu_{P(X)} \). For any operator \( Z \) in a tracial \( W^\ast\)-probability space, we know that

\[
1 - \mu_{P}(\{\lambda\}) = \text{rank}(\lambda - Z) = \frac{1}{2} \text{rank} \begin{pmatrix} \lambda - Z & 0 \\ 0 & \lambda - Z^* \end{pmatrix} = \frac{1}{2} \text{rank} \begin{pmatrix} 0 & -Z \\ \lambda - Z^* & 0 \end{pmatrix}
\]

Let us denote

\[
A_\lambda := \begin{pmatrix} 0 & \lambda - P \\ \lambda - P^* & 0 \end{pmatrix} \in M_2(\mathbb{C}\langle x_1, \ldots, x_d, x_1^*, \ldots, x_d^* \rangle).
\]

Then it suffices to prove that \( \inf_{Y \in \mathcal{X}_\varepsilon} \mu_{A_\lambda(Y)}(\{0\}) \leq \mu_{A_\lambda(X)}(\{0\}) \).

For that purpose, first for each \( k = 1, \ldots, d \) we can take a sequence \( D_k^{(m)} \) of diagonal matrices such that \( D^{(m)} = (D^{(m)}_1, \ldots, D^{(m)}_d) \in \mathcal{X}_{m,\varepsilon} \) and \( \mu_{D^{(m)}} \) converge weakly towards \( \mu_{X_k} \). Then we define \( Y^{(m)} := (Y^{(m)}_1, \ldots, Y^{(m)}_d) \in \mathcal{X}_{m,\varepsilon} \) with \( Y^{(m)} := U_k^{(m)} D_k^{(m)} U_k^{(m)*} \), where \( U_k^{(m)}, \ldots, U_d^{(m)} \) are a sample of independent Haar unitary random matrices of dimension \( m \). Clearly we have that \( Y^{(m)} \) converges in \(*\)-distribution towards \( X \) and thus \( \mu_{A_\lambda(Y^{(m)})} \) converges weakly towards \( \mu_{A_\lambda(X)} \). Hence

\[
\lim_{m} \inf \mu_{A_\lambda(Y^{(m)})}(\{0\}) \leq \limsup_{m} \mu_{A_\lambda(Y^{(m)})}(\{0\}) \leq \mu_{A_\lambda(X)}(\{0\})
\]

where we used Portmanteau Theorem in the last inequality. Hence the desired inequality follows as \( \inf_{Y \in \mathcal{X}_\varepsilon} \mu_{A_\lambda(Y)}(\{0\}) \leq \mu_{A_\lambda(Y^{(m)})}(\{0\}) \).

The main spirit of Proposition 5.4 is that if we prove a bound from below for the size of an atom for \( P(Y_1, \ldots, Y_d) \) when considering all tuples of matrices in \( Y_1, \ldots, Y_d \in \mathcal{X} \), this bound will hold for \( P(X_1, \ldots, X_d) \). On the contrary if we have one example of \( Y_1, \ldots, Y_d \in \mathcal{X} \) for which we can calculate the size of the atom, then this size is a bound from above.
for the size of the atom for $P(X_1, \ldots, X_d)$. In the next section we will use this observation repeatedly.

6. Applications to Polynomials

The purpose of this section is to apply the our general results in the particular but important case of polynomials. In this and the following section all our random variables and polynomials will be selfadjoint\footnote{A polynomial $P \in \mathbb{C}(x_1, \ldots, x_n)$ is selfadjoint if $P = P^*$. In our case this means that for any monomial $\lambda x_{i_1} x_{i_2} \ldots x_{i_n}$ appearing in $P$, the monomial $\overline{\lambda} x_{i_n} x_{i_{n-1}} \ldots x_{i_1}$ also appears in $P$} thus all analytic distributions will be probability measures on $\mathbb{R}$. Also, we will exclude trivial cases and then $P(x_1, \ldots, x_n)$ will contain all the variables $x_1, \ldots, x_n$.

We use Theorem 1.1 to bound from above the size of atoms for some polynomials. The proofs of our estimates will use operators with atomic distributions with rational sizes. However, all the bounds in this section remain true for general unbounded operators by suitable approximation.

More explicitly, we will consider $X = (X_1, \ldots, X_d)$, a tuple of $*$-free normal random variables, such that for $k = 1, \ldots, d$,

$$
\mu_{X_k} = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i^{(k)}},
$$

where $\lambda_i^{(k)} \in \mathbb{R}$, $i = 1, \ldots, n$.

Then we will choose a tuple $Y = (Y_1, \ldots, Y_d)$ of matrices in $M_n(\mathbb{C})$ such that for each $k = 1, \ldots, d$, $X_k$ has the same distribution as $Y_k$.

Again, thanks to Theorem 5.1, for each selfadjoint polynomial $P \in \mathbb{C}(x_1, \ldots, x_d)$ and $\lambda \in \mathbb{R}$

$$(3) \quad \mu_{P(X)}(\{\lambda\}) \leq \mu_{P(Y)}(\{\lambda\}).$$

Interestingly, in many cases, we will be able to calculate these weights exactly by comparing the bounds from above with general bounds from below that we derive in the next Subsection 6.1.

6.1. Some unavoidable atoms. Let us derive some lower bounds for the size of atoms for polynomials in any operators. The following propositions are basic lower bounds for the size of atoms. Notice that the proofs do not use any freeness assumptions for the involved operators. These are probably well known to the experts and easy to derive. We prove them for the convenience of the reader.

**Proposition 6.1.** Let $(\mathcal{M}, \tau)$ be a tracial $W^*$-probability space and consider a tuple $X = (X_1, \ldots, X_d)$ of selfadjoint operators in $\mathcal{M}$. Suppose that for each $i$, $\mu_i(\lambda_i) \geq t_i$ for some $\lambda_i \in \mathbb{R}$ and $t_i \in [0, 1]$. If $t_1 + \cdots + t_d > d - 1$ then for any selfadjoint polynomial $P \in \mathbb{C}(x_1, \ldots, x_d)$, the measure $\mu_{P(X)}$ has an atom at $P(\lambda)$ of size at least $t_1 + \cdots + t_d - (d-1)$, where $\lambda := (\lambda_1, \ldots, \lambda_d)$.

**Proof.** For each $i \in \{1, \ldots, d\}$, let $p_i$ be the orthogonal projection onto the eigenspace $\ker(\lambda_i - X_i)$. That is, $X_i p_i = \lambda_i p_i$ for all $i \in \{1, \ldots, d\}$ and $\tau(p_i) = t_i$. Then for
Let $X, Y_1, ..., Y_d$ be selfadjoint operators in some tracial $W^*$-probability space $(\mathcal{M}, \tau)$. Suppose that $\mu_X(\{0\}) \geq t$. Then for any selfadjoint polynomial $P$ of the form

$$P(x, y_1, ..., y_d) = \sum_{i=1}^{k} Q_{i,1}(x, y_1, ..., y_d)xQ_{i,2}(x, y_1, ..., y_d),$$

the analytic distribution of $P(X, Y_1, ..., Y_d)$ has an atom at 0 whose size is at least $\max(kt - (k-1), 0)$.

**Proof.** We bound the rank,

$$\text{rank}(P(X, Y_1, ..., Y_d)) \leq \sum_{i=1}^{k} \text{rank}(Q_{i,1}(X, Y_1, ..., Y_d)) \cdot \text{rank}(X) \leq k \text{rank}(X) = k(1-t),$$

from where we obtain

$$\mu_{P(X,Y_1,...,Y_d)}(\{0\}) = 1 - \text{rank}(P(X,Y_1,...,Y_d)) \geq 1 - k(1-t) = kt - (k-1).$$

Let us isolate two cases of polynomials of a special form which we will use later.

**Corollary 6.3.** Let $X, Y_1, ..., Y_d$ be selfadjoint operators in some tracial $W^*$-probability space $(\mathcal{M}, \tau)$.

1. Suppose that $\mu_X(\{0\}) \geq t > 1/2$. Then for any polynomial $P$ of the form

$$P(x, y_1, ..., y_d) = xQ(x, y_1, ..., y_d) + Q^*(x, y_1, ..., y_d)x$$

the analytic distribution of $P(X, Y_1, ..., Y_d)$ has an atom at 0 of size at least $2t - 1$.

2. For any polynomial $P$ of the form $P(x, y_1, ..., y_n) = Q_1(x, y_1, ..., y_n)xQ_2(x, y_1, ..., y_n)$, we have $\mu_{P(X,Y_1,...,Y_n)}(\{0\}) \geq \mu_X(\{0\})$

**Remark 6.4.** All the bounds above may also be derived, by proving the analog theorems for the specific case of matrices.

Indeed, if we prove such bound for free random variables $X, Y_1, ..., Y_d$, then the conclusion will also hold for operators in a finite von Neumann algebra, by our Theorem 5.1.

Now, by Proposition 6.2 and Theorem 2.24, we only need to prove this for $X, Y_1, ..., Y_d \in M_n(\mathbb{C})$.

### 6.2. Using diagonal matrices

Let $X_1, X_2$ be two freely independent selfadjoint random variables with

$$\mu_{X_1} = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i} \quad \text{and} \quad \mu_{X_2} = \frac{1}{n} \sum_{i=1}^{n} \delta_{\mu_i}$$

where $\lambda_i, \mu_i \in \mathbb{R}$, $i = 1, ..., n$. Let $Y_1$ and $Y_2$ be diagonal matrices with $\lambda_1, ..., \lambda_n$ and $\mu_1, ..., \mu_n$ on their diagonals respectively. Thanks to Theorem 5.1, for any $P \in \mathcal{C}(x_1, x_2)$ we have

$$\sigma_P(P(X_1, X_2)) \subset \sigma_P(P(Y_1, Y_2)) = \{P(\lambda_i, \mu_i) \mid i = 1, ..., n\}$$
as

\[
P(Y_1, Y_2) = \begin{pmatrix} P(\lambda_1, \mu_1) & 0 & \cdots & 0 \\ 0 & P(\lambda_2, \mu_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P(\lambda_n, \mu_n) \end{pmatrix}.
\]

Moreover, since we can reorder \(\lambda_i\)'s and \(\mu_j\)'s while \(\mu_X\) and \(\mu_Y\) stay invariant, we see that

\[
\sigma_p(P(X_1, X_2)) \subset \bigcap_{\sigma \in S_n} \{P(\lambda_i, \mu_{\sigma(i)}) \mid i = 1, \ldots, n\}.
\]

Let us define

\[
\Gamma_\sigma(\lambda) := \{i \in \{1, \ldots, n\} \mid p(\lambda_i, \mu_{\sigma(i)}) = \lambda\}, \quad \forall \lambda \in \mathbb{C}, \forall \sigma \in S_n
\]

and

\[
k(\lambda) := \min_{\sigma \in S_n} |\Gamma_\sigma(\lambda)|.
\]

Then we have \(\mu_{P(X_1, X_2)}(\{\lambda\}) \leq \frac{1}{n} k(\lambda)\). This can be done similarly for any tuple of operators \(X_1, \ldots, X_d\).

**Theorem 6.5.** Let \(X_1, \ldots, X_d\) be freely independent selfadjoint random variables and let \(P \in \mathbb{C} \langle x_1, \ldots, x_d \rangle\) be a selfadjoint polynomial. The possible atoms of \(P(X_1, \ldots, X_d)\) are contained in the set

\[
\{P(\rho_1, \ldots, \rho_d) \mid \rho_i \text{ is atom of } X_i, \ \forall i = 1, \ldots, d\}.
\]

Furthermore, if \(X_1, \ldots, X_d \in M_n(\mathbb{C})\) and \(\lambda \in \mathbb{R}\), then \(\mu_{P(X_1, \ldots, X_d)}(\{\lambda\})\) is smaller than \(k(\lambda)/n\), where

\[
k(\lambda) := \min_{\lambda_1, \ldots, \lambda_d \in \mathbb{R}} |\{j \in \{1, \ldots, n\} : P(\lambda_{\sigma_1(j)}, \ldots, \lambda_{\sigma_d(j)}) = \lambda\}|,
\]

where, for \(i = 1, \ldots, d\), \(\{\lambda_{j}^{(i)}\}_{j=1}^{n}\) denotes the eigenvalues of \(X_i\).

One natural question is whether \(k(\lambda)\) gives the precise size of the atom at \(\lambda\). Certainly if \(k(\lambda) = 0\) this is the case. A positive result is well-known for the special case \(P(x_1, x_2) = x_1 + x_2\). Actually, Theorem 7.4 in [BV98] implies that \(\alpha = \lambda + \mu\) is an atom of \(X_1 + X_2\) if and only if the sum of the multiplicity of \(\lambda\) and the multiplicity of \(\mu\) is larger than \(n\). Such a condition is sufficient for arbitrary polynomials \(P \in \mathbb{C} \langle x_1, x_2 \rangle\). That is, \(\alpha = P(\lambda, \mu)\) is an atom of \(P(X_1, X_2)\) if the sum of the multiplicity of \(\lambda\) and the multiplicity of \(\mu\) is larger than \(n\). Proposition 6.1 actually corresponds to this situation for any number of variables.

However, this description is not sufficient to give all atoms of \(P(X_1, X_2)\). For example, if \(P(x_1, x_2) = x_1 x_2 - x_2 x_1\), then \(\{\lambda \in \mathbb{C} \mid k(\lambda) > 0\} = \{0\}\) and \(k(0) = n\). But in Section 7 we will see in general for free random variables that \(X_1 X_2 - X_2 X_1\) may not have atoms even when \(X_1\) and \(X_2\) are purely atomic. Nevertheless, the Theorem [L3] shows that this condition could still be necessary under the assumption of certain injectivity properties of the polynomial.

Before proving this result, let us give some examples. First, Theorem [L3] allows to recover the well-known results about atoms of the free additive [BV98] and to generalize to the result in [Bel03] for multiplicative convolutions: indeed, we readily see that \(P(X, Y) = \)
The atom is at least $\max(X, Y)$.

Proof of Theorem 1.3. We want to prove that the atom is at most $\max(X, Y)$ for $X > 0$ and $Y > 0$. Another new example is $P(X, Y) = X^2YX + XYX^2$ for $a \neq 0$, and it is easy to exhibit other examples as Example 6.9 shows.

Let us state this

**Corollary 6.6.** Let $\mu, \nu \in \mathcal{P}$ be probability measures on $\mathbb{R}$.

1. $\mu \boxplus \nu$ has an atom at $a \in \mathbb{R}$ if and only if there exist $\lambda, \rho \in \mathbb{R}$ such that $\lambda = a$ and $\mu(\lambda) + \nu(\rho) > 1$. Moreover, if $\mu(\lambda) + \nu(\rho) > 1$, we have $\mu \boxplus \nu(a) = \mu(\lambda) + \nu(\rho) - 1$.
2. If $\mu \in \mathcal{P}(\mathbb{R}^+)$ then $\mu \boxdot \nu$ has an atom at $a \in \mathbb{R} \setminus \{0\}$ if and only if there exist $\lambda, \rho \in \mathbb{R}$ such that $\rho \lambda = a$ and $\mu(\lambda) + \nu(\rho) > 1$. Moreover, if $\mu(\lambda) + \nu(\rho) > 1$, we have $\mu \boxdot \nu(a) = \mu(\lambda) + \nu(\rho) - 1$.
3. If $P(X, Y) = X^2YX + XYX^2$, then $P^{\sqcup}(\mu, \nu)$ has an atom at $a \in \mathbb{R} \setminus \{0\}$ if and only if there exist $\lambda, \rho \in \mathbb{R}$ such that $P(\rho, \lambda) = a$ and $\mu(\lambda) + \nu(\rho) > 1$. Moreover, if $\mu(\lambda) + \nu(\rho) > 1$, we have $P^{\sqcup}(\mu, \nu)(a) = \mu(\lambda) + \nu(\rho) - 1$.

**Proof.** Given Theorem 1.3, the only thing to analyze is the atom at 0 in part (2). Corollary 6.6 already shows that $\mu \boxdot \nu \{0\} \geq \max(\mu \{0\}, \nu \{0\})$. To show that $\mu \boxdot \nu \{0\} \leq \max(\mu \{0\}, \nu \{0\})$, we may assume that $\mu \sim X_n^2$ and $\nu \sim Y_n$.

Consider the specific realization of $X_n$ and $Y_n$,

$$X_n = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \quad \text{and} \quad Y_n = \begin{pmatrix} \rho_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

with $\lambda_i \neq 0$ for $0 \leq i \leq m$ and $\rho_j \neq 0$, for $0 \leq j \leq l$. Then, if $r = \min(l, m)$, we have

$$X_n Y_n X_n = \begin{pmatrix} \lambda_1 \rho_1 \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

The result follows since $\text{Nullity}(X_n Y_n X_n) = \max\{\text{Nullity}(X_n), \text{Nullity}(X_n)\}$. \qed

Now we prove Theorem 1.3, which is one of the main results of this section.

**Proof of Theorem 1.3.** The atom is at least $\max(t + s - 1, 0)$ at $a$, by Proposition 6.1. We want to prove that the atom is at most $\max(t + s - 1, 0)$ at $a$. 

Case 1. $s + t - 1 > 0$. Consider the diagonal matrices,

$$X_n = \begin{pmatrix} \lambda & \cdots & \lambda \\ & \ddots & \vdots \\ & & \lambda_{i+1} \end{pmatrix} \quad \text{and} \quad Y_n = \begin{pmatrix} \rho_1 & \cdots & \rho_j \\ & \ddots & \vdots \\ & & \rho \end{pmatrix},$$

where $i = sn \geq j = n - tn$.

Then

$$P(X_n, Y_n) = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{pmatrix}$$

where

$$A_1 = \begin{pmatrix} P(\lambda, \rho_1) & \cdots & \cdots \\ & \ddots & \vdots \\ & & P(\lambda, \rho_i) \end{pmatrix}, \quad A_2 = \begin{pmatrix} P(\lambda, \rho) & \cdots & \cdots \\ & \ddots & \vdots \\ & & P(\lambda, \rho) \end{pmatrix},$$

and

$$A_3 = \begin{pmatrix} P(\lambda_{j+1}, \rho) & \cdots & \cdots \\ & \ddots & \vdots \\ & & P(\lambda_n, \rho) \end{pmatrix}.$$

By hypothesis, $P(\lambda, \rho_i) \neq a$ for $l \leq i$ and $P(\lambda_l, \rho) \neq a$ for $l > j$. We see that the size of the eigenspace associated to $a$ is $\dim(A_2) = i - j = (s + t - 1)n$.

Case 2. $s + t - 1 \leq 0$. We claim that we can reorder the eigenvalues in such a way that $P(\lambda_i, \rho_i) \neq a$. Then taking the matrices $X_n = \text{diag}(\lambda_1, \ldots, \lambda_n)$ and $Y_n = (\rho_1, \ldots, \rho_n)$, we see that $P(X_n, Y_n) = \text{diag}(P(\lambda_1, \rho_1), \ldots, P(\lambda_n, \rho_n))$ which has no eigenvalue equal to $a$, as desired.

Claim. There is an ordering of the eigenvalues of $X_n$ and $Y_n$ such that $P(\lambda_i, \rho_i) \neq a$ for all $i = 1, \ldots, n$.

We prove this by induction on $n$. For $n = 1$ and $n = 2$, it is clear. Consider a reordering of $\{\lambda_i, \rho_i\}$ such $P(\lambda_i, \rho_i) \neq a$ for $i \leq n - 1$ which is possible by induction. Now consider $\lambda_n, \rho_n$. If $P(\lambda_n, \rho_n) \neq a$ we are done. Otherwise, if $P(\lambda_n, \rho_n) = a$, then let

$$S = \{ j \in [n - 1] \mid \text{such that } \lambda_j = \lambda_n \text{ or } \rho_j = \rho_n \}.$$

If $|S| = n - 1$, by choosing for each $j$, $\rho_j$ or $\lambda_j$, together with $\lambda_n, \rho_n$ we have that $sn + tn \geq n + 1$, which yields a contradiction. Finally, if $|S| \leq n - 2$, there exists $j$ such that $\lambda_j \neq \lambda_n$ and $\rho_j \neq \rho_n$. By the injectivity of $P$ for $a$, $P(\lambda_j, \rho_n) \neq a$ and $P(\lambda_n, \rho_j) \neq a$, from where we get the desired reordering.

Remark 6.7. Note that the conditions in the previous theorem only depend on the evaluation of the polynomial in the atoms of the marginals and thus, under the above hypothesis, two injective non-commutative polynomials in two variables with the same values when evaluated in the atomic parts will have the same atoms when evaluated in free random
variables. For instance, $X + Y$ and $X + Y + i(XY - YX)$ have exactly the same atoms, for any $X$ and $Y$ which are free.

One of the implications of Theorem 1.3 is available for more variables, namely the analog of the bound from Proposition 6.1 is also achieved. For that purpose, first we generalize that condition on polynomials, as follows.

**Definition 6.8.** Let $P \in \mathbb{C}(x_1, \ldots, x_d)$ be given. For $a \in \mathbb{R}$, if for all $\rho_1, \ldots, \rho_i, \ldots, \rho_d \in \mathbb{C}$, such that $P(\rho_1, \ldots, \rho_i, \ldots, \rho_d) = a$ we have that

$$P(\rho_1, \ldots, \rho_i, \ldots, \rho_d) \neq P(\rho_1, \ldots, \tilde{\rho}_i, \ldots, \rho_d)$$

whenever $\rho_i \neq \tilde{\rho}_i$ for any given scalar values $\rho_1, \ldots, \rho_{i-1}, \rho_{i+1}, \ldots, \rho_d$, then we say $P$ is injective for $a$. If $P$ is injective for all $a \in \mathbb{R}$, then we simply say that $P$ is injective.

**Example 6.9.** For random variables supported on $\mathbb{R}^+$, a polynomial $P$ with positive coefficients such that any monomial has odd order in any variable is injective. Some examples in three variables are:

1. $X + Y + Z;
2. Y X Z + Z X Y;
3. X^3 + X Z X^2 + X^4 Z X + Y.$

**Theorem 6.10.** Let $P \in \mathbb{C}(x_1, \ldots, x_d)$ be injective for $a \in \mathbb{R}$ and let $a = P(\lambda_1, \ldots, \lambda_d)$. Suppose that $X_1, \ldots, X_d$ are free selfadjoint random variables and $X_i \sim \mu_i$ with $\mu_i(\{\lambda_i\}) = t_i$. If $t_1 + \ldots + t_d \geq d - 1$ then the distribution of $P(X_1, \ldots, X_d)$ has an atom at $a$ of size exactly $t_1 + \ldots + t_d - (d - 1)$.

**Proof.** Again, the lower bound is obtained by Proposition 6.1. We proceed by induction on the number of variables to get the upper bound. We assume that we can make a construction as in the proof of Theorem 1.3 for $d$ variables and show that we can make the same construction for $d + 1$. Suppose that $X_1, \ldots, X_{d+1}$ are free selfadjoint random variables, $X_i \sim \mu_i$ with $\mu_i(\{\lambda_i\}) = t_i$ and $t_1 + \ldots + t_{d+1} \geq d$. First note that, since $t_i \leq 1$, the condition $t_1 + \ldots + t_{d+1} \geq d$ implies that $t_1 + \ldots + t_d \geq (d - 1)$.

Using the induction hypothesis, we assume that for the first $d$ variables we have diagonal matrices $Y_1, \ldots, Y_d$ such that:

1. There are $n(d - (t_1 + \ldots + t_d)) = l_d$ indices $m$ such that the following holds: There exists $i \in \{1, \ldots, d\}$ such that $(Y_j)_{mm} = \lambda_j$ for $j \neq i$ and $(Y_i)_{mm} \neq \lambda_i$.
2. The remaining $n - l_d$ indices, $m$, satisfy $(Y_j)_{mm} = \lambda_j$ for all $j$.

The base case $d = 2$ is done in Theorem 1.3. Without loss of generality, by a simple reordering, we may assume that case (1) is achieved in the indices $1, \ldots, l_d$.

We claim that if (1) and (2) are satisfied then the eigenspace of $a$ has dimension

$$n - l_d = n - n(d - (t_1 + \ldots + t_d)) = n(t_1 + \ldots + t_d - d - 1)$$

as desired. Indeed, for $m < l_d$, since $(Y_j)_{mm} = \lambda_j$ for $j \neq i$ and $Y_i \neq \lambda_d$, then $P((Y_1)_{mm}, \ldots, (Y_d+1)_{mm}) \neq a$. Conversely, for $m \geq l_d$ we have $(Y_j)_{mm} = \lambda_j$ and thus $P((Y_1)_{mm}, \ldots, (Y_d+1)_{mm}) = a$ in this range.

To prove (1) and (2) for $d + 1$, we consider the diagonal matrix $(Y_{d+1})_{mm} = \lambda_{d+1}$ for $m \leq n(t_{d+1})$ and $(Y_{d+1})_{mm} \neq \lambda_{d+1}$ for $m \leq n(t_d)$. 


Firstly, since \( t_1 + \ldots + t_{d+1} > d \) then \( n(t_{d+1}) > n(d - t_1 + \ldots + t_d) = l_d \). Thus, the indices \( m \in \{1, \ldots, l_d\} \) satisfy the property in (1) since by induction \((Y_j)_{mm} = \lambda_j\) for \( j \neq i \) and \( 1 \leq j \leq d \). \((Y_i)_{mm} = \lambda_j\) and we have \((Y_{d+1})_{mm} = \lambda_{d+1}\).

Secondly, for \( m > n(t_{d+1}) \), the indices \( m \), also satisfy property (2), since \( m > l_d \), by induction that \((Y_j)_{mm} = \lambda_j\) for \( j \leq d \) and \( Y_{d+1} \neq \lambda_{d+1} \).

Finally, the indices in \( m \in \{l_d, \ldots, n(t_{d+1})\} \) satisfy the property in (2), since again by induction \((Y_j)_{mm} = \lambda_j\) for all \( 1 \leq j \leq d \), and in this interval we have \((Y_{d+1})_{mm} = \lambda_{d+1}\). This proves that the number of indices satisfying (2) is \( n(t_{d+1}) - l_d = n - l_{d+1} \), which finishes the induction.

On the other hand, similar as for the commutator, for some polynomials we may exclude atoms outside of 0 (or any \( \lambda \in \mathbb{R} \), by a simple shifting). A selfadjoint polynomial \( P = P(x, y) \) in the variables \( x \) and \( y \), is called mixed if it consists only of monomials containing both \( x \) and \( y \) (this is equivalent to \( P(0, y) = 0 \) and \( P(x, 0) = 0 \)).

**Proposition 6.11.** Let \( X \) and \( Y \) be free random variables and suppose that \( X \sim \mu \) and \( Y \sim \nu \) with \( \mu = t\delta_0 + (1 - t)\bar{\mu} \) and \( \nu = s\delta_0 + (1 - s)\bar{\nu} \). If \( t + s \geq 1 \) then for all mixed polynomials \( P \), the distribution of \( P(X, Y) \) has no atoms outside of 0.

**Proof.** After suitable approximation, we may assume

\[
\mu = \frac{1}{n} \sum_{i=1}^{m} \delta_{\lambda_i} + \frac{n-m}{n} \delta_0 \quad \text{and} \quad \nu = \frac{m}{n} \delta_0 + \frac{1}{n} \sum_{i=1}^{n-m} \delta_{\mu_i}
\]

for some \( \lambda_i, \mu_j \in \mathbb{R} \).

By Proposition 5.4, it is enough to find matrices \( X_n, Y_n \) with distributions \( \mu \) and \( \nu \) such that \( P(X_n, Y_n) = 0 \). Consider the diagonal matrices,

\[
X_n = \begin{pmatrix}
\lambda_1 & \cdots & 0 \\
0 & \ddots & \cdots \\
0 & \cdots & \lambda_m
\end{pmatrix} \quad \text{and} \quad Y_n = \begin{pmatrix}
0 & \cdots & 0 \\
0 & \ddots & \cdots \\
\mu_1 & \cdots & \mu_{n-m}
\end{pmatrix}
\]

Notice that \( X_nY_n = X_nY_n = 0 \) and thus for any mixed monomial \( m \) we have that \( m(X_n, Y_n) = 0 \). By linearity, we have for any mixed polynomial that \( P(X_n, Y_n) = 0 \). \( \square \)

### 6.3. The algebra of two non-commuting projections

The idea of this section is to improve the bounds from the previous one by considering a more general situation where we allow blocks of size \( 2 \times 2 \) on the diagonal.
So let $X_n, Y_n \in M_{2n}(\mathbb{C})$ be matrices consisting of diagonal blocks of size $2 \times 2$, $A_1, \ldots, A_n$ and $B_1, \ldots, B_n$ on their diagonals, respectively. Similar as before,

$$p(X_n, Y_n) = \begin{pmatrix} p(A_1, B_1) & 0 & \cdots & 0 \\ 0 & p(A_2, B_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p(A_n, B_n) \end{pmatrix}.$$ 

We are interested in the eigenvalues of $p(A_i, B_i)$ with $A_i, B_i \in M_2(\mathbb{C})$. We may now write $A_i$ and $B_i$ in terms of the one rank one projections

$$R = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} t & \sqrt{(1-t)t} \\ \sqrt{(1-t)t} & 1 - t \end{pmatrix}.$$ 

That is, we may assume that

$$A_i = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = (\lambda_1 - \lambda_2) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$B_i = \begin{pmatrix} \frac{t \mu_1 + \mu_2(1-t)}{(\mu_1 - \mu_2)\sqrt{(1-t)t}} & \frac{(\mu_1 - \mu_2)\sqrt{(1-t)t}}{\mu_1(1-t) + t \mu_2} \\ \frac{t \mu_1 + \mu_2(1-t)}{(\mu_1 - \mu_2)\sqrt{(1-t)t}} & \frac{(\mu_1 - \mu_2)\sqrt{(1-t)t}}{\mu_1(1-t) + t \mu_2} \end{pmatrix}$$

$$= (\mu_1 - \mu_2) \begin{pmatrix} t \sqrt{(1-t)t} \\ \sqrt{(1-t)t} \end{pmatrix} + \mu_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$ 

In this way, $A = (\lambda_1 - \lambda_2)R + \lambda_2 I$ and $B = (\mu_1 - \mu_2)T + \mu_2 I$. So $P(A_i, B_i) = Q_i(R, T)$ for some polynomial $Q_i$.

In this way the invertibility of $P(A_i, B_i)$ is translated into the invertibility of $Q_i(R, T)$. The set of polynomials in $R$ and $T$ is well understood and rather easy to handle. Moreover, when considering $T$ as a function of $t$, for all $t \in (0, 1)$, the algebra generated by $R$ and $T$ is known to be isomorphic to the $C^*$-algebra generated by two non-commuting projections, see [RSS9].

To understand any polynomial in $R, T$ it is enough to consider certain quadratic polynomials. Indeed, since we have the relations $RTR = tR, TRT = tT, R^2 = R$ and $T^2 = T$, any polynomial in $R$ and $T$ may be written as a linear combination

$$aR + bT + cRT + dTR + eI,$$

which actually has the matrix form

$$\begin{pmatrix} a + e + (b + c + d)t & (b + c)\sqrt{(1-t)t} + c \\ (b + d)\sqrt{(1-t)t} & e + b(1-t) \end{pmatrix}.$$ 

The determinant of this matrix is given by $a(b(-t) + b + e) + be + cd^2 - cdt + cet + det + e^2$ which for $e = 0$ (that we may always assume), reduces to $(t - 1)(cdt - ab)$.

We aim to turn this calculation into a criterion to avoid situations as in Corollary 6.2 or even to characterize all the sizes of the atoms of some polynomials. To do this we introduce some notation.
For \( \lambda, \rho \in \mathbb{C} \), let us define define \( \text{Subs}_{\lambda, \rho} : C \langle x, y \rangle \rightarrow \mathbb{C}[t] \) which sends \( P \mapsto \text{Subs}_{\lambda, \rho}[P] \) as follows: For a monomial \( m = m(x, y) \), \( \text{Subs}_{\lambda, \rho}[m] \in \mathbb{C}[t] \), is given by the formula

\[
\text{Subs}_{\lambda, \rho}[m](t) = m(\lambda, \rho)t^{k-1},
\]

whenever \( m(x, y) \) is of any of the following forms: \( m(x, y) = x^iy^j \ldots x^i y^j \), \( m(x, y) = y^i x^j \ldots y^j x^i \).

We extend \( \text{Subs}_{\lambda, \rho} \) by linearity to any polynomial \( P \in C \langle x, y \rangle \) with no linear term.

**Definition 6.12.** Let \( P \in C \langle x, y \rangle \), be a polynomial with no constant term and write \( P = P_1 + P_2 + P_3 + P_4 \). where

\[
P_1 = xQ_1 y, \quad P_2 = yQ_2 x, \quad P_3 = xQ_3 x + ax \quad \text{and} \quad P_4 = yQ_4 y + by
\]

for some polynomials \( Q_i \in C \langle x, y \rangle \) and \( a, b \in \mathbb{C} \).

(1) We say that \( P \) satisfies the determinant condition on \((\lambda, \mu)\) if for some \( t \in [0, 1] \),

\[
t\text{Subs}_{\lambda, \rho}[P_1](t)\text{Subs}_{\lambda, \rho}[P_2](t) \neq \text{Subs}_{\lambda, \rho}[P_3](t)\text{Subs}_{\lambda, \rho}[P_4](t).
\]

(2) We say that \( P \) satisfies the strong determinant condition if it satisfies the determinant condition for all pairs \((\lambda, \rho)\) with \( \lambda, \rho \neq 0 \).

**Remark 6.13.** While condition (1) seems to be hard to check for specific \( \lambda \) and \( \mu \), there are actually many polynomials such that the strong determinant condition is satisfied and easy to check. Here we list some useful conditions which imply the strong determinant condition.

1) \( P_1 = 0 \) or \( P_2 = 0 \) and \( P_3, P_4 \neq 0 \).

2) \( P_3 = 0 \) or \( P_4 = 0 \) and \( P_1, P_2 \neq 0 \).

3) \( \deg(P_1P_2) \neq \deg(P_3P_4) + 1 \).

4) \( P_3 \) and \( P_4 \) have constant term.

**Proposition 6.14.** Let \( X \) and \( Y \) be free random variables and let \( t \geq s \) be such that \( t + s \geq 1 \). Suppose that \( X \sim \mu \) and \( Y \sim \nu \) with \( \mu = t\delta_0 + (1-t)\hat{\mu} \) and \( \nu = s\delta_0 + (1-s)\hat{\nu} \), with \( \hat{\mu} = \sum_{i=1}^l \frac{1}{n} \lambda_i \) and \( \hat{\nu} = \frac{1}{n} \sum_{i=1}^m \delta \rho_i \) with \( \lambda_i \neq 0 \) and \( \rho_i \neq 0 \). If there exist an injective function \( \sigma : [l] \rightarrow [m] \) such that \( \lambda_{\sigma(i)} \rho_{\sigma(i)} \) satisfies the determinant condition (1) for \( P(X, Y) \) then we have the following.

1) The atom of \( P(X, Y) \) at \( 0 \) has size at most \( 2t - 1 \).

2) If for all \( \rho \), \( P(0, \rho) \neq 0 \) then the atom at \( 0 \) has exactly size \( s + t - 1 \).

3) \( P(x, y) \) is mixed, then \( P(X, Y) \) has no more atoms.

**Example 6.15.**

1) Let \( X_1, X_2 \) have distributions of the form \( 1/2\delta_0 + 1/2\mu_1 \) and \( 1/2\delta_0 + 1/2\mu_2 \) with \( \mu_i \) having no atoms at \( 0 \). Then any polynomial of the form \( P(X, Y) = XQ_1 Y + YQ_2 X \) has no atoms. In particular, this is the case for \( XY + XY \) and \( i(XY - YX) \).

2) Consider the polynomial \( P(x, y) = yxxy + xyxy + yxyx + xxyy \), which is mixed and does not satisfy the determinant condition. One sees that

\[
P(x, y) = (y + 1)(xyx)(y + 1)
\]

and thus \( P(X, Y) \) has an atom at \( 0 \) if \( Y \) has an atom at \( 0 \) too (this is just free multiplicative convolution). Thus the determinant condition is not superfluous. Simpler examples like \( Q(x, y) = xyx, R(x, y) = xyx + xyxyx \) also work.
Proof of Proposition 6.14. Without loss of generality we may assume that $\sigma$ in the hypothesis is the inclusion, i.e. $\sigma(i) = i$.

Now, consider the block-diagonal matrices in $M_{2n}$ given by

$$
X_n = \begin{pmatrix}
A_1 & & \\
& \ddots & \\
& & A_{2l}
\end{pmatrix} \quad \text{and} \quad
Y_n = \begin{pmatrix}
B_1 & & \\
& \ddots & \\
& & B_{2m}
\end{pmatrix}
$$

where $A_i = \lambda_i R$ for $i = 1, \ldots, 2l$, $B_i = \rho_i T$ for $i = 1, \ldots, 2m$, where

$$
R = \begin{pmatrix}
1 & 0 \\
0 & 0
\end{pmatrix} \quad \text{and} \quad
T = \begin{pmatrix}
t & \sqrt{(1-t)t} \\
\sqrt{(1-t)t} & 1-t
\end{pmatrix}
$$

We see that

$$
P(X_n, Y_n) = \begin{pmatrix}
P(A_1, B_1) & & \\
& \ddots & \\
& & P(A_{2l}, B_{2l})
P(0, B_{2l+1}) & & \\
& \ddots & \\
& & P(0, B_{2m})
P(0, 0) & & \\
& \ddots & \\
& & 0
\end{pmatrix}
$$

which means that if $P(A_i, B_i)$ is invertible, for all $1 \leq i \leq 2l$, then $P(X_n, Y_n)$ has nullity at most $2n - 4l = 2n(1 - 2(1-t)) = 2n(2t - 1)$.

Let us fix $1 \leq i \leq 2l$ and notice that $A_i$ and $B_i$ satisfy the relations

$$
A_i B_i A_i = \lambda_i^2 \rho_i t A_i \quad \text{and} \quad B_i A_i B_i = \lambda_i \rho_i^2 t B_i \quad \text{and} \quad A_i^k = \lambda_i^k A_i \quad \text{and} \quad B_i^k = \rho_i^k B_i.
$$

Using these relations iteratively we arrive at the following:

$$
P(A_i, B_i) = \text{Subs}[P_1(t, \lambda_i, \rho_i)]RT + \text{Subs}[P_2(t, \lambda_i, \rho_i)]TR
+ \text{Subs}[P_3(t, \lambda_i, \rho_i)]R + \text{Subs}[P_4(t, \lambda_i, \rho_i)]T.
$$

Finally, for any $a_1, a_2, a_3, a_4 \in \mathbb{C}$, the determinant of $a_1 RT + a_2 TR + a_3 R + a_4 T$ is given by $(t - 1)(a_1 a_2 - a_3 a_4)$. Thus $\det[P(A_i, B_i)] \neq 0$ whenever the determinant condition holds for $\lambda_i$ and $\rho_i$.

For part (2), if $P(0, \rho) \neq 0$ then $P(0, B_i)$ for $2l < i \leq 2m$ has at least one non-zero eigenvalue and then the nullity decreases in at least $2(m - l)$. Thus the nullity is at most

$$
2n - 4l - 2(m - l) = 2n - 2m + 2l = 2n(1 - (1-s) - (1-t)) = 2n(s + t - 1).
$$

However, by Proposition 6.11, the size of the nullity is at least $2n(s + t - 1)$.

Finally, part (3) follows by Proposition 6.11. \qed
6.4. A rational function. To end this section let us consider two different rational expressions in \(x\) and \(y\), \(R_1(x, y) = xy^{-2}x + yx^{-2}y\) and \(R_2(x, y) = xy^{-2}x + x^{-1}y^2x^{-1}\).

We consider \(X\) and \(Y\) free random variables which are strictly positive, i.e. with support on \((0, \infty)\). We are interested in the atoms of \(R_1(X, Y)\) and \(R_2(X, Y)\). In such a case, \(R_1(X, Y)\) and \(R_2(X, Y)\) are well-defined due to Theorem 5.4 (for example, we can compare \(X, Y\) to their classically independent counterparts, which are clearly well-defined).

By considering the case where \(X\) and \(Y\) are commuting we see that similar as for polynomials, the atomic parts of \(R_1(X, Y)\) and \(R_2(X, Y)\) are both contained in the set

\[
\{\rho^2/\lambda^2 + \lambda^2/\rho^2 \mid \rho \text{ is an atom of } X, \lambda \text{ is an atom of } Y\}.
\]

Moreover, for given \(\rho\) and \(\lambda\) the functions \(f, g : \mathbb{R}^+ \to \mathbb{R}\), defined by \(f(x) = \rho^2/x^2 + x^2/\rho^2\), and \(g(x) = x^2/\lambda^2 + \lambda^2/x^2\), are both injective. Thus, arguing in the same way as in the proof of Proposition 1.3 one can prove that \(R_1(X, Y)\) has an atom at \(a\) if and only if there exist \(\lambda\) and \(\rho\) such that \(\mu_X(\{\lambda\}) + \mu_Y(\{\rho\}) > 1\) and \(a = \rho^2/\lambda^2 + \lambda^2/\rho^2\) and the mass at \(a\) is given by \(\mu_X(\{\lambda\}) + \mu_Y(\{\rho\}) - 1\). In particular, their atoms are inside \([2, \infty)\).

We see then that \(R_1(X, Y)\) and \(R_2(X, Y)\) have exactly the same atoms with the same weights. However, the full distributions \(\mu_{R_1(X,Y)}\) and \(\mu_{R_2(X,Y)}\) may differ, as one can see by calculating their second moments.

7. Examples

In this final section we show how to use the above results in two specific but important examples, the selfadjoint polynomials \(P_1(x, y) = i[x, y] = i(xy - yx)\) and \(P_2(x, y) = \{x, y\} = xy + yx\). For \(X\) and \(Y\) free selfadjoint variables we completely characterize the atoms of \(P_1(X, Y)\) and \(P_2(X, Y)\) in terms of the atoms of \(X\) and \(Y\).

7.1. Commutator. The following theorem describes exactly the size of the atoms of the commutator, \(P_1(x, y) = i[x, y] = i(xy - yx)\) of free random variables.

**Theorem 7.1.** Let \(X\) and \(Y\) be free random variables. Let \(t\) and \(s\) be the size of the largest atom of \(X\) and \(Y\), respectively, i.e.,

\[
t = \max\{\mu_X(\{a\}) \mid a \in \mathbb{R}\}
\quad \text{and} \quad
s = \max\{\mu_Y(\{b\}) \mid b \in \mathbb{R}\}.
\]

Then

1. \(i(XY - YX)\) has no atoms outside of 0.
2. \(i(XY - YX)\) has an atom at 0 of size given by \(\max(2t - 1, 2s - 1, 0)\).

**Proof.** Part 1 is easy. By taking \(X_n\) and \(Y_n\) diagonal matrices, we see that, since \(X_n Y_n - Y_n X_n = 0\), the only possible atom of \(i[X, Y]\) is at 0.

We will show that the size of this atom is given by \(r = \max(2t - 1, 2s - 1, 0)\), where \(t\) and \(s\) are the sizes of the largest atoms of \(X\) and \(Y\), respectively.

Now, already Corollary 6.3 tells us that \(\mu(\{0\}) \geq r\). Hence we only need to show that \(\mu(\{0\}) \leq \max(2t - 1, 2s - 1, 0)\). For this we will use Proposition 5.3 on block diagonal matrices.

Notice that, for any \(\lambda, \rho \in \mathbb{R}\), we have that \(i[X, Y] = i[X + \lambda Y + \rho]\), so we may assume without loss of generality that the largest of the atoms of \(X\) and \(Y\) are at 0. Also, since
and thus we may reorder the eigenvalues so that 

\[ \lambda \]

where 

\[ X \]

assume that 

\[ Y \]

s < multiplicity). Since 

\[ \lambda \]

\[ \rho \]

Our aim is to show that 

\[ \lambda_{2i-1} \neq \lambda_{2i} \]

Similarly, if \( \{\rho_i\}_{i=1}^{2n} \) is the multiset of eigenvalues of \( Y_n \) we may assume that \( \rho_{2i-1} \neq \rho_{2i} \)

Now, consider the block-diagonal matrices

\[
X_n = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_n \end{pmatrix}, \quad Y_n = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_n \end{pmatrix}
\]

where

\[
A_i = \begin{pmatrix} \lambda_{2i-1} & 0 \\ 0 & \lambda_{2i} \end{pmatrix} \quad \text{and} \quad B_i = \frac{1}{2} \begin{pmatrix} \rho_{2i-1} + \rho_{2i} & \rho_{2i-1} - \rho_{2i} \\ \rho_{2i-1} - \rho_{2i} & \rho_{2i-1} + \rho_{2i} \end{pmatrix}.
\]

As above

\[
[X_n, Y_n] = \begin{pmatrix} [A_1, B_1] & & \\ & \ddots & \\ & & [A_n, B_n] \end{pmatrix}
\]

and thus it is enough to prove that \([A_i, B_i]\) is invertible, for all \( i \).

This follows from

\[
[A_i, B_i] = \frac{1}{2} \begin{pmatrix} 0 & (\lambda_{2i} - \lambda_{2i-1})(\rho_{2i} - \rho_{2i-1}) \\ -(\lambda_{2i} - \lambda_{2i-1})(\rho_{2i} - \rho_{2i-1}) & 0 \end{pmatrix},
\]

whose determinant is \((\lambda_{2i} - \lambda_{2i-1})^2(\rho_{2i} - \rho_{2i-1})^2 \neq 0\).

Case 3. \( t > 1/2, t + s < 1 \). Let \( m = t(2n) \), which we may assume to be an integer. Our aim is to show that \( i(X_nY_n - Y_nX_m) \) has nullity \( 2m - 2n = (2t - 1)2n \), for some \( X_n \) and \( Y_n \) assuming that \( X_n \) has nullity \( m \) and \( Y_n \) has all eigenspaces of size at most \( n \), since \( s < 1/2 \). Again if \( \{\rho_i\}_{i=1}^{2n} \) is the multiset of eigenvalues of \( Y_n \) we may assume that \( \rho_{2i-1} \neq \rho_{2i} \)

On the other hand, for the eigenvalues of \( X_n \), since \( t > 1/2 \) and thus \( m > n \), we may assume that \( \lambda_{2i} = 0 \), for \( i = 1, \ldots, n \), and that \( \lambda_{2i+1} \neq 0 \) for \( i = 1, \ldots, 2n - m \) and \( \lambda_{2i+1} = 0 \) for \( i = 2n - m + 1, \ldots, n \).

Now, consider again the block-diagonal matrices

\[
X_n = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_n \end{pmatrix}, \quad Y_n = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_n \end{pmatrix}
\]

where

\[
A_i = \begin{pmatrix} \lambda_{2i-1} & 0 \\ 0 & \lambda_{2i} \end{pmatrix} \quad \text{and} \quad B_i = \frac{1}{2} \begin{pmatrix} \rho_{2i-1} + \rho_{2i} & \rho_{2i-1} - \rho_{2i} \\ \rho_{2i-1} - \rho_{2i} & \rho_{2i-1} + \rho_{2i} \end{pmatrix}.
\]
Remark 7.2. \(1\) When the variables \(X_1, \ldots, X_n\) are symmetric, the result may be deduced from Equation (1.29) in [NS98] and the description of atoms for the free additive [BV98] and multiplicative [Bel03] convolutions.

\(2\) In [EL21] the authors study the closure properties of the class of freely infinitely divisible laws, \(ID(\mathbb{H})\). They show that all the quadratic forms \(Q(x_1, \ldots, x_d)\) such that \(Q(x_1, \ldots, x_d) \in ID(\mathbb{H})\) whenever \(X_1, \ldots, X_d \in ID(\mathbb{H})\) coincide with the linear span of commutators of free random variables. They also show that their distribution is symmetric. Since measures in \(ID(\mathbb{H})\) have at most one atom, this tells that only possible atoms of \(Q(X_1, \ldots, X_d)\) are at 0 whenever \(X_1, \ldots, X_d\) belong to \(ID(\mathbb{H})\). Our results show that this is true for any free variables \(X_1, \ldots, X_d\).

7.2. Anticommutator. Now we consider the atoms of the anticommutator \(P_2(x, y) = \{x, y\} = xy + yx\), for free random variable.

Theorem 7.3. Let \(X\) and \(Y\) be free random variables and let \(Z = XY + YX\).

i) The size of the atom at 0 of \(Z\) is given by

\[
l := \max\{2t - 1, 2s - 1, s + u - 1, t + r - 1, 0\},
\]

where

1. \(t\) is the size of the atom at 0 of \(X\);
2. \(s\) is the size of the atom at 0 of \(Y\);
3. \(u\) is the size of the largest atom outside of 0 of \(X\);
4. \(r\) is the size of the largest atom outside of 0 of \(Y\).

ii) For any \(a \neq 0\), \(Z\) has an atom at \(a\) if and only if there exist weights \(s(a)\) and \(t(a)\) such that \(t(a) + s(a) - 1 > 0\), \(X\) has an atom at \(\lambda\) of size \(s(a)\) and \(Y\) has an atom at \(\rho\) of size \(t(a)\) and \(2\lambda\rho = a\).

Furthermore, the size of the atom of \(Z\) at \(a \neq 0\) is given by \(\max\{s(a) + t(a) - 1, 0\}\).

Proof. The size of atoms outside of 0 may be calculated from the commutative case. Indeed for \(\lambda, \rho \neq 0\) we see that \(P(\lambda, \rho) = 2\lambda\rho\) satisfies the hypothesis of Theorem 1.3.

The atom at 0 is more complicated and we need to use \(2 \times 2\) matrices. We will prove that the atom at 0 cannot be bigger than \(l\). The other inequality follows from Proposition 6.1 and Corollary 6.3.

Before continuing we notice the following calculations for \(2 \times 2\) matrices which we will use repeatedly. For complex numbers \(\lambda_i, \lambda_j\) and \(\rho_i, \rho_j\), consider the matrices

\[
A(\lambda_i, \lambda_j) = \begin{pmatrix} \lambda_i & 0 \\ 0 & \lambda_j \end{pmatrix}
\]

and

\[
B_1(\rho_i, \rho_j) = \begin{pmatrix} \rho_i & 0 \\ 0 & \rho_j \end{pmatrix}, \quad B_2(\rho_i, \rho_j) = \frac{1}{2} \begin{pmatrix} \rho_i + \rho_j & \rho_i - \rho_j \\ \rho_i - \rho_j & \rho_i + \rho_j \end{pmatrix}.
\]
A simple calculation shows that \( \{A, B_1\} \) is invertible if \( \lambda_i, \lambda_j, \rho_i, \rho_j \) are non-zero; \( \{A, B_2\} \) is invertible if \( \lambda_i \neq 0, \lambda_j = 0 \) and \( \rho_i \neq \rho_j \).

We take \( X_n, Y_n \in M_{n}(\mathbb{C}) \), where the non-zero eigenvalues of \( X_n \) are denoted by \( \lambda_1, \ldots, \lambda_{(1-t)n} \), the non-zero atom of \( Y_n \) with the largest weight \( r \) is denoted by \( \rho \) and the remaining eigenvalues of \( Y_n \) are denoted by \( \rho_1, \ldots, \rho_{(1-r)n} \). We consider the different cases depending on \( t \) and \( r \). Without loss of generality we assume that \( t \geq s \).

**Case 1.** \( t \leq 1/2, r+t \leq 1 \). We divide \( X_n \) into two parts, that is,

\[
X_n = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}, \quad Y_n = \begin{pmatrix} Y_1 & 0 \\ 0 & Y_2 \end{pmatrix},
\]

where \( X_1, Y_1 \in M_{2tn} \) and \( X_2, Y_2 \in M_{(1-2t)n} \) are defined as follows:

\[
X_1 = \begin{pmatrix} A(\lambda_1, 0) \\ & \ddots \\ & & A(\lambda_{tn}, 0) \end{pmatrix}, \quad Y_1 = \begin{pmatrix} B_2(\rho, \rho_1) \\ & \ddots \\ & & B_2(\rho, \rho_{tn}) \end{pmatrix},
\]

\( X_2 \) is a diagonal matrix that contains the remaining non-zero eigenvalues. If \( r \geq t \), then we define

\[
Y_2 = \begin{pmatrix} B_2(\rho, \rho_1) \\ & \ddots \\ & & B_2(\rho, \rho_{tn}) \end{pmatrix},
\]

where we may assume \( Y_1 \) contains all zero eigenvalues of \( Y_n \) as \( s \leq t \). Then we have \( \text{Nullity}(\{X_1, Y_1\}) = 0 \). If \( r < t \), then we have \( r < 1/2 \). In this case, we can order the eigenvalues of \( Y_n \) such the any pair of adjacent eigenvalues is distinct and thus in particular \( \text{Nullity}(\{X_1, Y_1\}) = 0 \). In both case, we let \( Y_2 \) be a diagonal matrix consisting of the remaining non-zero eigenvalues. Therefore we have \( \text{Nullity}(\{X_2, Y_2\}) = 0 \).

**Case 2.** \( t \leq 1/2, r + t > 1 \). We divide \( X_n \) and \( Y_n \) into three parts, that is,

\[
X_n = \begin{pmatrix} X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & X_3 \end{pmatrix}, \quad Y_n = \begin{pmatrix} Y_1 & 0 & 0 \\ 0 & Y_2 & 0 \\ 0 & 0 & Y_3 \end{pmatrix},
\]

where \( X_1, Y_1 \in M_{(r+t-1)n} \), \( X_2, Y_2 \in M_{(2-2r)n} \) and \( X_3, Y_3 \in M_{(r-t)n} \) are defined as follows:

\[
X_2 = \begin{pmatrix} A(\lambda_1, 0) \\ & \ddots \\ & & A(\lambda_{(1-r)n}, 0) \end{pmatrix}, \quad Y_2 = \begin{pmatrix} B_2(\rho, \rho_1) \\ & \ddots \\ & & B_2(\rho, \rho_{(1-r)n}) \end{pmatrix},
\]

\( X_1 = 0 \), \( X_3 \) is a diagonal matrix that contains the remaining non-zero eigenvalues and

\( Y_3 \) contains no zero eigenvalues (which is possible since we can let \( Y_1 \) and \( Y_2 \) together has \( tn \) many zero eigenvalues while we know \( s \leq t \)). Hence we see that \( \text{Nullity}(\{X_2, Y_2\}) = \text{Nullity}(\{X_3, Y_3\}) = 0 \) and thus \( \text{Nullity}(\{X_n, Y_n\}) = (r + t - 1)n \).
Case 3. $t \geq 1/2$, $r \geq t$. We write $t = (2t - 1) + (r - t) + (1 - r)$ and divide $X_n$ into three parts that contain correspondingly many zero eigenvalues. Consider the matrices

$$X_n = \begin{pmatrix} X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & X_3 \end{pmatrix}, \quad Y_n = \begin{pmatrix} Y_1 & 0 & 0 \\ 0 & Y_2 & 0 \\ 0 & 0 & Y_3 \end{pmatrix}$$

where $X_1, Y_1 \in M_{(2t-1)n}$, $X_2, Y_2 \in M_{(2r-2)t}$, and $X_3, Y_3 \in M_{(2-2r)n}$ are defined as follows: $X_1 = 0$,

$$X_2 = \begin{pmatrix} A(\lambda_1, 0) \\ \vdots \\ A(\lambda_{(r-t)n+1}, 0) \end{pmatrix}, \quad X_3 = \begin{pmatrix} A(\lambda_{(r-t)n+1}, 0) \\ \vdots \\ A(\lambda_{(1-t)n}, 0) \end{pmatrix}$$

and

$$Y_1 = Y_2 = \begin{pmatrix} \rho \\ \vdots \\ \rho \end{pmatrix}, \quad Y_3 = \begin{pmatrix} B_2(\rho, \rho_1) \\ \vdots \\ B_2(\rho, \rho_{(1-r)n}) \end{pmatrix}.$$ 

Thus one sees that

$$\text{Nullity}([X_1, Y_1]) = (2t - 1)n, \quad \text{Nullity}([X_2, Y_2]) = (r - t)n, \quad \text{Nullity}([X_3, Y_3]) = 0,$$

and so $\text{Nullity}([X_n, Y_n]) = (r + t - 1)n$.

Case 4. $t \geq 1/2$, $r < t$. We proceed similarly as in Case 3 but we do not need $X_2, Y_2$. Correspondingly, we let $X_3, Y_3$ have size $(2 - 2t)n$. If $s \geq 1 - t$ or $r \geq 1 - t$, then we can build $Y_3$ with block matrix $B_2(0, \rho_i)$ or $B_2(\rho, \rho_i)$ respectively. Hence $\text{Nullity}([X_3, Y_3]) = 0$. Otherwise we have $s < 1/2$ and $r < 1/2$, in which case we can order the eigenvalues of $Y_n$ such the any pair of adjacent eigenvalues is distinct and thus in particular $\text{Nullity}([X_3, Y_3]) = 0$. So the desired result follows similarly as in Case 3. □

Remark 7.4. One may want to include both calculations in this section in the more general situation $e^{it}XY + e^{-it}YX$. It turns out that, whenever $\cos(t) \neq 0$, the situation is exactly the same as in the anticommutator case, by just changing $2\mu \lambda$ to $2\cos(t)\mu \lambda$. Indeed, considering $\tilde{X} = e^{it}X$ and using the same notation as in Theorem 7.3, one sees that $AB_1 + B_1^*A$ is invertible if $\lambda_i, \lambda_j, \rho_i, \rho_j$ are non-zero and $AB_2 + B_2^*A$ is invertible if $\lambda_i \neq 0, \lambda_j = 0$ and $\rho_i \neq \rho_j$. The rest of the proof then mutatis mutandis.

7.3. Multiplicative commutator. Now we consider the rational function $XYX^{-1}Y^{-1}$ where $X$ and $Y$ are free invertible elements. The particular case, when $X$ and $Y$ are unitaries is of special interest but for the calculation here, it is not needed.

Since when $X$ and $Y$ are commutative, clearly $XYX^{-1}Y^{-1} = 1$, then the only possible atom is at 1. In principle we could try to find some $2 \times 2$ matrix to have a similar calculations as for the commutator, however the following calculation shows that the size of the atom at 0 of $XY - YX$ is the same as the size of the atom at 1 for $XYX^{-1}Y^{-1}$.

$$\text{rank}(XY - YX) = \text{rank}((XY - YX)X^{-1}Y^{-1}) = \text{rank}(XYX^{-1}Y^{-1} - 1).$$

Thus arriving to the following conclusion.
Corollary 7.5. Let $X$ and $Y$ be free random variables that are invertible. Let $t$ and $s$ be the size of the largest atoms of $X$ and $Y$, then the only possible atom of $XYX^{-1}Y^{-1}$ is at 1 with mass $\max(2t - 1, 2s - 1, 0)$.

7.4. Universal Covers of Graphs. Now we present an application to the problem of relating the spectrum of the adjacency matrix $A(G)$ of a graph with the point spectrum of the density of states of its universal cover (or universal covering tree). We thank Jorge Garza-Vargas for leading our attention to this possible application.

Let $G = (V(G), E(G))$ be a finite simple undirected connected graph. The universal cover of $G$, that we denote by $T(G)$, is the graph constructed as follows:

1. Fix any vertex $v \in V(G)$. The set of vertices of $T(G)$ is given by non-backtracking walks in $G$ starting at $v$, including the empty walk. Here, a non-backtracking walk means a finite sequence of vertices $v_1, v_2, \cdots, v_n$, so that $v_i \sim v_{i+1}$, for $i = 1, \ldots, n - 1$ and $v_i \neq v_{i+2}$ for $i = 1, \ldots, n - 2$.
2. The set of edges of $T(G)$ is given by the pair of non-backtracking walks such that one of them can be obtained by appending a vertex.

$T(G)$ is independent of $v$, up to isomorphism. $T(G)$ is called the universal cover because it is the unique graph, up to isomorphism, that covers every other cover of $G$.

One can associate to $T(G)$ a probability measure known as the density of states of $T(G)$, we refer for precise definitions to $\[BGVM17\]$ and references therein. The relation with our results comes from the fact, that as observed in $[BC19]$ and $[GVK19]$, if $G$ has adjacency matrix $A = (A_{i,j})_{i,j}$ and $\{u_{i,j}\}_{0 \leq i,j \leq n}$ is a family of Haar distributed unitaries with $u_{i,j} = u_{j,i}^{-1}$ and $\{u_{i,j}\}_{i>j}$ freely independent, then the matrix $U(A)$ defined by $U(A)_{i,j} = A_{i,j}u_{i,j}$ has the same distribution as the density of states associated to the universal cover $T(G)$.

Proposition 7.6. $\[BGVM17\]$ Theorem 3.2] The point spectrum of $T(G)$ is contained in the set of eigenvalues of $A(G)$. Moreover if $\mu$ denotes the density of states of $T(G)$ and $\mu_{A(G)}$ denotes the empirical eigenvalue distribution of $A(G)$ then $\mu_{T(G)} \leq \mu_{A(G)}$.

Here we want to see how this fact may deduced from Theorem 5.1. To do this, we consider the matrix $R(A)$ on the variables $X := (x_{i,j})_{i>j}$ given by $R(A)(x)_{i,j} = A_{i,j}x_{i,j}$ if $i > j$, $R(A)(X)_{i,j} = A_{i,j}x_{j,i}^{-1}$ if $i < j$ and $R(A)(X)_{i,i} = 0$.

Now, $R(A)(X)$ does not belong to $M_n(\mathbb{C} \langle x_{2,1}, \ldots, x_{n,n-1} \rangle)$, nor is a rational expression in the variables $\{x_{i,j}\}_{i>j}$. However by following the same line of proof of part (2) in Theorem 5.1 one may arrive to the same conclusion for matrices which are rational expressions (in this case $x_{i,j}^{-1}$).

Now we only need to observe that the variables $u_{i,j}$ are normal and $\mu_{u_{i,j}}^p \leq \mu_p^1 = \delta_1$, since the variables $u_{i,j}$ are absolutely continuous. So $U = (u_{i,j})_{i>j}$ and $1 = (1)_{i>j}$ satisfy the hypothesis of Theorem 5.1 and we conclude that $\mu_{U(A)}^p \leq \mu_{R(A)}^p$. Proposition 7.6 i.e $\mu_{U(A)}^p \leq \mu_A^p = \mu_A$, now follows from the observation that $R(U) = U(A)$ and $R(1) = A$.

REFERENCES

[Ami66] S. Amitsur, Rational identities and applications to algebra and geometry, J. Algebra 3 (1966), no. 3, 304–359.

[BBL21] S.T. Belinschi, H. Bercovici, and W. Liu, The atoms of operator-valued free convolutions, Journal of Operator Theory 85 (2021), no. 1, 303–320.
Universality of Free Random Variables

[BC19] C. Bordenave and B. Collins, *Eigenvalues of random lifts and polynomials of random permutation matrices*, Annals of Mathematics 190 (2019), no. 3, 811–875.

[BCD+10] H. Bercovici, B. Collins, K. Dykema, W.-S. Li, and D. Timotin, *Intersections of Schubert varieties and eigenvalue inequalities in an arbitrary finite factor*, Journal of Functional Analysis 258 (2010), no. 5, 1579–1627.

[Bel03] S. T. Belinschi, *The atoms of the free multiplicative convolution of two probability distributions*, Integral Equations and Operator Theory 46 (2003), no. 4, 377–386.

[BMD+10] H. Bercovici, B. Collins, K. Dykema, W.-S. Li, and D. Timotin, *Intersections of Schubert varieties and eigenvalue inequalities in an arbitrary finite factor*, Journal of Functional Analysis 258 (2010), no. 5, 1579–1627.

[Bel03] S. T. Belinschi, *The atoms of the free multiplicative convolution of two probability distributions*, Integral Equations and Operator Theory 46 (2003), no. 4, 377–386.

[BGVM17] Jess Banks, Jorge Garza-Vargas, and Satyaki Mukherjee, *Point spectrum of periodic operators on universal covering trees*, International Mathematics Research Notices 2020 (2017), rnab152.

[BMS17] S.T. Belinschi, T. Mai, and R. Speicher, *Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem*, Journal für die reine und angewandte Mathematik (Crelles Journal) 2017 (2017), no. 732, 21–53.

[BV93] H. Bercovici and D. Voiculescu, *Free convolution of measures with unbounded support*, Indiana University Mathematics Journal 42 (1993), no. 3, 733–773.

[BV98] *Regularity questions for free convolution*, Nonselfadjoint operator algebras, operator theory, and related topics, Springer, 1998, pp. 37–47.

[CMM+21] B. Collins, T. Mai, A. Miyagawa, F. Parraud, and S. Yin, *Convergence for noncommutative rational functions evaluated in random matrices*, arXiv:2103.05962 (2021).

[Coh06] P. M. Cohn, *Free ideal rings and localization in general rings*, Cambridge: Cambridge University Press, 2006.

[CR99] P. M. Cohn and C. Reutenauer, *On the construction of the free field*, Int. J. Algebra Comput. 9 (1999), no. 3–4, 307–324.

[CS05] A. Connes and D. Shlyakhtenko, *L2-homology for von Neumann algebras*, J. Reine Angew. Math. 586 (2005), 125–168.

[CU16] G. Cébron and M. Ulrich, *Haar states and Lévy processes on the unitary dual group*, Journal of Functional Analysis 270 (2016), no. 7, 2769 – 2811.

[EL21] W. Ejsmont and F. Lehner, *Sums of commutators in free probability*, Journal of Functional Analysis 280 (2021), no. 2, 108791.

[GGOW16] A. Garg, L. Gurvits, R. Oliveira, and A. Wigderson, *A Deterministic Polynomial Time Algorithm for Non-Commutative Rational Identity Testing with Applications*, 2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS), 2016, pp. 109–117.

[GVK19] J. Garza-Vargas and A. Kulkarni, *Spectra of infinite graphs via freeness with amalgamation*, arXiv preprint arXiv:1912.10137 (2019).

[HMS18] J. W. Helton, T. Mai, and R. Speicher, *Applications of realizations (aka linearizations) to free probability*, J. Funct. Anal. 274 (2018), no. 1, 1–79.

[HS09] U. Haagerup and H. Schultz, *Invariant subspaces for operators in a general III-factor*, Publications Mathématiques de l’IHÉS 109 (2009), 19–111.

[HW00] U. Haagerup and C. Winslow, *The Effros–Marechal topology in the space of von Neumann algebras, II*, Journal of Functional Analysis 171 (2000), no. 2, 401–431.

[HW15] P. Hrubeš and A. Wigderson, *Noncommutative arithmetic circuits with division*, Theory Comput. 11 (2015), 357–393.

[JZ19] A. Jaikin-Zapirain, *The base change in the Atiyah and the Lück approximation conjectures*, Geometric and Functional Analysis 29 (2019), no. 2, 464–538.

[KLS17] A. Knebusch, P. Limnell, and T. Schick, *On the center-valued Atiyah conjecture for L2-Betti numbers*, Documenta Mathematica 22 (2017), 659–677.

[KR86] R.V. Kadison and J.R. Ringrose, *Fundamentals of the theory of operator algebras*, vol. II, Academic Press, New York, 1986.

[KV] D. S. Kaliuzhnyi-Verbovetskyi and V. Vinnikov, *Noncommutative rational functions, their difference-differential calculus and realizations*, Multidimensional Syst. Signal Process.

[Lüc02] W. Lück, *L2–Invariants: Theory and Applications to Geometry and K–Theory*, Berlin: Springer, 2002.
T. Mai, *Regularity of distributions of Wigner integrals*, arXiv preprint arXiv:1512.07593 (2015).

P. Malcolmson, *Determining homomorphisms to skew fields*, J. Algebra 64 (1980), no. 2, 399–413.

J.A. Mingo, M. Popa, and K. Szpojankowski, *On the partial transpose of a Haar unitary matrix*, 2021.

J.A. Mingo and R. Speicher, *Free Probability and Random Matrices*, Fields Institute Monographs, Springer New York, 2017.

T. Mai, R. Speicher, and S. Yin, *The free field: realization via unbounded operators and Atiyah property*, arXiv:1905.08187 (2019).

Alexandru Nica and Roland Speicher, *Commutators of free random variables*, Duke Mathematical Journal 92 (1998), no. 3, 553–592.

A. Nica and R. Speicher, *Lectures on the combinatorics of free probability*, Cambridge: Cambridge University Press, 2006.

A. Nica, D. Shlyakhtenko, and R. Speicher, *Operator-valued distributions: I. Characterizations of freeness*, International Mathematics Research Notices 2002 (2002), no. 29, 1509–1538.

K.K. Olesen, *The Connes Embedding Problem*, Københavns Universitet, Institut for Matematiske Fag, 2012.

H. Reich, *Group von Neumann algebras and related algebras*, PhD thesis, University of Göttingen (1998).

I. Raeburn and A.M. Sinclair, *The C*-algebra generated by two projections*, Mathematica Scandinavica (1989), 278–290.

A. Schofield, *Representations of rings over skew fields*, London Mathematical Society Lecture Notes 92, University Press, Cambridge, 1985.

D. Shylakhtenko, *Remarks on free entropy dimension*, Operator Algebras, Springer, 2006, pp. 249–257.

D. Shlyakhtenko and P. Skoufranis, *Freely independent random variables with non–atomic distributions*, Trans. Am. Math. Soc. 367 (2015), no. 9, 6267–6291.

D. Voiculescu, K. J. Dykema, and A. Nica, *Free random variables*, CRM Monograph Series 1 (1992).

D. Voiculescu, *The analogues of entropy and of Fisher’s information measure in free probability theory: VI. Liberation and mutual free information*, Advances in Mathematics 146 (1999), no. 2, 101–166.