Free Probability Theory: Deterministic Equivalents and Combinatorics

Dissertation
zur Erlangung des Grades
des Doktors des Naturwissenschaften
der Naturwissenschanftlich-Technischen Fakultäten
der Universität des Saarlandes

von
Carlos Vargas Obieta

Saarbrücken
2015

Tag des Kolloquiums:	04.05.2015
Dekan:	Prof. Dr. Markus Bläser
Prüfungsausschuss:	Vorsitzender
	Prof. Dr. Hannah Markwig
	Berichterstatter
	Prof. Dr. Roland Speicher
	Prof. Dr. Jörg Eschmeier
	Prof. Dr. Florent Benaych-Georges
	Akademischer Mitarbeiter
	Dr. Guillaume Cébron

This work is dedicated to all those stubbornly trying to bring education to people who live in remote and vulnerable places.

In solidarity with the families of the students of the "Raúl Isidro Burgos" Rural Teachers' College of Ayotzinapa:

Abel García Hernández
Abelardo Vázquez Peniten
Adán Abrajan de la Cruz
Antonio Santana Maestro
Benjamín Ascencio Bautista
Bernardo Flores Alcaraz
Carlos Iván Ramírez Villarreal
Carlos Lorenzo Hernández Muñoz
César Manuel González Hernández
Christian Alfonso Rodríguez Telumbre
Christian Tomas Colon Garnica
Cutberto Ortiz Ramos
Dorian González Parral
Emiliano Alen Gaspar de la Cruz.
Everardo Rodríguez Bello
Felipe Arnulfo Rosas
Giovanni Galindes Guerrero
Israel Caballero Sánchez
Israel Jacinto Lugardo
Jesús Jovany Rodríguez Tlatempa
Jonas Trujillo González
Jorge Álvarez Nava
Jorge Aníbal Cruz Mendoza
Jorge Antonio Tizapa Legideño
Jorge Luis González Parral
José Ángel Campos Cantor
José Ángel Navarrete González

José Eduardo Bartolo Tlatempa
José Luis Luna Torres
Jhosivani Guerrero de la Cruz
Julio César López Patolzin
Leonel Castro Abarca
Luis Ángel Abarca Carrillo
Luis Ángel Francisco Arzola
Magdaleno Rubén Lauro Villegas
Marcial Pablo Baranda
Marco Antonio Gómez Molina
Martín Getsemany Sánchez García
Mauricio Ortega Valerio
Miguel Ángel Hernández Martínez
Miguel Ángel Mendoza Zacarías
Saúl Bruno García,

kidnapped on Sept. 26th, 2014 in Iguala, Guerrero, Mexico,

and to the memory of the student Alexander Mora Venancio.

"We are missing our 43!".

Zusammenfassung

Gegenstand dieser Arbeit ist die freie Wahrscheinlichkeitstheorie. Ihr Hauptziel ist es, die asymptotische Eigenwertverteilung einer großen Klasse von Zufallsmatrizen zu verstehen. Für die in [SV12] diskutierten Modelle erhalten wir einen sehr allgemeinen Algorithmus zur graphischen Darstellung ihrer Verteilungen. Wir wenden auch Methoden aus [BSTV14] an, um einen allgemeinen numerischen Algorithmus zur Berechnung der asymptotischen Verteilungen anderer Typen von Matrizenmodellen formulieren zu können, wie etwa für die Block-linear modifizierten Verteilungen, die in [Aub12, BN12a, BN12b] betrachtet wurden.

Klassische, freie und nicht-kommutative Wahrscheinlichkeitstheorie können einheitlich über die Kombinatorik multiplikativer Funktionen bezüglich verschiedener Verbände von Partitionen von Mengen verstanden werden. Das zweite Ziel dieser Arbeit ist es, eine Übersicht über einige der grundlegenden kombinatorischen Strukturen in der freien Wahrscheinlichkeitstheorie zu geben. Unsere wesentliche Referenz hierfür ist [NS06]. Wir stellen neue Resultate vor, die die Berechnung der Kumulanten von Produkten freier und Boolesch unabhängiger Variablen mittels Posets k-teilbarer Partitionen ermöglichen [AV12]. Darüber hinaus geben wir Formeln an, die verschiedene Typen von Kumulanten zueinander in Verbindung setzen [AHLV14].

In Verbindung mit der Zufallsmatrizentheorie nutzen wir speziell den kombinatorischen Zugang zur operatorwertigen freien Wahrscheinlichkeitstheorie ([NSS02]), um die Cauchy-Stieltjes-Transformierten der asymptotischen Eigenwertverteilungen der in [SV12] eingeführten Matrizenensembles zu berechnen. Wir tun dies, um zu zeigen, dass unsere Definition eines freien deterministischen Äquivalents als konkreter Operator, wie er in [SV12] eingeführt wurde, mit dem weitverbreiteten Begriff des deterministischen Äquivalents übereinstimmt, wie er beispielsweise zur Beschreibung neuerer Matrizenmodelle in der drahtlosen Kommunikation verwendet wird [CD11].

Voiculescu begründete die freie Wahrscheinlichkeitstheorie im Jahr 1985 im Kontext von Operatoralgebren. 1991 fand er Realisierungen seiner freien Kreis-, Halbkreis- und Haar-unitären Operatoren als Grenzwerte von Eigenwertverteilungen bemerkenswerter unabhängigen Zufallsmatrizenmodelle, wie etwa (selbst-adjungierte und nicht selbst-adjungierte) Wigner Matrizen und Haar-unitäre Zufallsmatrizen. So ermöglichte Voiculescu das Verständnis des Wignerschen Halbkreisgesetzes als den einvariabligen Sonderfall eines wesentlich allgemeineren Phänomens bei gemeinsamen nicht-kommutativen Verteilungen großer Zufallsmatrizen.

1995 führte er die operatorwertige freie Wahrscheinlichkeitstheorie ein. Damit wurden verschiedene Zufallsmatrixmodelle auch durch Mittel der freien Wahrscheinlichkeitstheorie beschreibbar. Dasselbe gilt für Produkte von Block-Halbkreis Matrizen [BSTV14] und Block-modifizierte Zufallsmatrizen [ANV], welche in dieser Arbeit kurz betrachtet

werden. Viele hermiteschen Zufallsmatrixmodelle P ergeben sich durch Auswertung eines selbst-adjungierten Polynoms

$$P(x_1, \ldots, x_n)$$

nicht-kommutierender Variablen $x_1, \ldots, x_n, x_1^*, \ldots, x_n^*$ in zufälligen und deterministischen Matrizen. In dieser Arbeit beschäftigen wir uns mit derartigen Zufallsmatrizenmodellen, welche wir als "polynomiale Modelle" bezeichnen wollen.

Falls die Auswertung in unabhängigen (selbst-adjungierten oder nicht selbst-adjungierten) Wigner Matrizen, Wishart Matrizen, zufälligen Haar Unitären und deterministischen Matrizen erfolgt, kann man eine deterministische, operator-algebraische Vereinfachung $P^\square$ von P im Rahmen der freien Wahrscheinlichkeitstheorie betrachten, um damit eine Approximation der Eigenwertverteilung von P zu erhalten. Die Dimensionen der Matrizen $x_1, \ldots, x_n$ dürfen dabei unterschiedlich sein. Das freie deterministische Äquivalent $P^\square$ von P wurde in [SV12] eingeführt, basierend auf der in [BG09a, BG09b] beschriebenen Verallgemeinerung von [Voi91] auf rechteckige Räume und unter hauptsächlicher Verwendung der kombinatorischen Werkzeuge aus [NSS02].

Die Methode deterministischer Äquivalente (DE) wurde von Girko auf der Ebene der Cauchy-Stieltjes-Transformierten der betrachteten Matrizenmodelle eingeführt. Im Gegensatz zu seinen deterministischen Äquivalenten kann unsere Vereinfachung $P \to P^\square$ sehr leicht beschrieben werden und setzt darüber hinaus auch keine spezielle Gestalt des betrachteten Polynoms voraus. Es wird sich zeigen, dass unsere Definition aus [SV12] sehr effektiv mit allen Elementen der in [BMS13] beschriebenen Methode zur Berechnung der Verteilung selbst-adjungierter Polynome in quadratischen, asymptotisch freien Zufallsmatrizen kombiniert werden kann.

Im Verlauf dieser Arbeit werden wir anmerken, wie verschiedene Annahmen über die Verteilungen, die wir in die Zufallsmatrizenmodelle einsetzen, die Qualität und die Art der Konvergenz des Modells zu seinem freien deterministischen Äquivalent beeinflussen. Dies diskutieren wir insbesondere auf kombinatorischer Ebene: Die verschiedenen Annahmen über das Modell bestimmen die Klasse der Verteilungen (oder Kumulanten), die sich in den Matrizensummen zeigen, und damit die Momente und die Natur der Fixpunktgleichungen, die wir für die freien deterministischen Äquivalente erhalten.

Um auch numerisch effizient zu sein, muss man verstehen, wie sich Freeness einschränkt und fortsetzt zwischen verschiedenen operator-wertigen Ebenen. Zu diesem Zweck sind die Methoden von [NSS02] besonders wichtig.

Summary

The topic of this thesis work is free probability theory. The main goal is to understand asymptotic eigenvalue distributions of large classes of random matrices. For the models discussed in [SV12], we obtain a quite general algorithm to plot their distributions. We also apply the tools from [BSTV14] to give a general numerical algorithm to compute the asymptotic distribution of some other types of matrix models, such as the block-linearly modified models which have been considered in [Aub12, BN12a, BN12b].

Classical, free and non-commutative probability can be jointly understood through the combinatorics of multiplicative functions with respect to different lattices of set partitions. The second goal of this thesis is to survey on the basic topics on the combinatorics of free probability. Our basic reference is [NS06]. We present new results which allow to compute cumulants of products of free and Boolean independent random variables in terms of the posets of k-divisible set partitions [AV12]. We also find formulas relating the different types of cumulants in non-commutative probability [AHLV14].

In connection to random matrix theory, we make particular use of the combinatorial approach to operator-valued free probability ([NSS02]) to compute Cauchy-Stieljes transforms of the asymptotic eigenvalue distributions of the matrix ensembles introduced in [SV12]. We do this to show that our definition of a free deterministic equivalent as a concrete operator, introduced in [SV12], agrees with the more widespread notion of deterministic equivalents which are being used, for example, to describe recent matrix models for wireless communications [CD11].

Voiculescu introduced free probability in 1985 in the context of operator algebras. In 1991, he found realizations of his free circular, semi-circular and Haar-unitary operators through limits of eigenvalue distributions of quite remarkable random matrix models, such as independent (self-adjoint and non-self-adjoint) Wigner and Haar-unitary random matrices. This allowed to understand Wigner's semicircle law as a special, single-variable case of a very general phenomenon on joint non-commutative distributions of large random matrices.

In 1995 he introduced operator-valued free probability, where the limiting behaviors of much more general random matrix models can be realized. A rich class of random matrix models arises from considering a polynomial

$$P(x_1, \ldots, x_n)$$

in non-commutative indeterminates $x_1, \ldots, x_n, x_1^*, \ldots, x_n^*$ and evaluating it on random and deterministic matrices. In this work we are specially concerned about these kind of models. We refer to them as "polynomial models".

If the inputs are (self-adjoint or non-self-adjoint) Wigner matrices, Wishart matrices, and deterministic matrices, we may consider a deterministic operator $P^{\square}$ by evaluating

P on certain operators $(y_1, \ldots, y_n)$ in the context of Voiculescu's free probability theory. Provided that the size of the matrices is large (but not necessarily too large), the spectral measure of the simplified model $P^{\square}$ becomes a good approximation of the averaged eigenvalue distribution of P. The dimensions of the matrices can also be different. The free deterministic equivalent $P^{\square}$ of P was defined in [SV12], based on the generalizations of [Voi91] to rectangular spaces [BG09a, BG09b], and using mostly combinatorial tools from [NSS02].

The method of deterministic equivalents (DE) was introduced by Girko at the level of Cauchy-Stieltjes transforms of the considered matrix models. In contrast to DE, the simplification $P \to P^{\square}$ can be explained very easily and does not require the polynomial to have a specific form. It will turn out that our definitions from [SV12] can be very effectively combined with all the elements of method described in [BMS13] for the distributions of polynomials on self-adjoint, square, asymptotically free random matrices.

Throughout this work, we comment on how the different assumptions on the distributions that we input to the random matrix models affect the quality and the type of convergence of the model to its FDE. In particular, we discuss this at the combinatorial level: The different assumptions on the model determine the classes of partitions (or cumulants) that show up on the matrix sums, and hence the moments and the nature of the fixed point equations that we will get for its FDE.

To be numerically efficient one needs to understand how freeness restricts and extends to different operator-valued levels. For this, the combinatorial methods from [NSS02] are quite important.

Acknowledgements

I thank Prof. Dr. Roland Speicher for all his advice and support during these last 5 years and for his efforts towards constructing an active, cooperative and welcoming group on free probability at the University of the Saarland.

I thank my dear friend, Octavio Arizmendi, for several years of friendship and mathematical collaboration. Part of our published work [AV12] is directly addressed in this thesis work. I also thank my colleague Ion Nechita for allowing me to present the numerical part of our on-going work ([ANV]).

I would like to thank my coauthors: Serban Belinschi, Takahiro Hasebe, Franz Lehner, Roland Speicher, John Treilhard, for allowing me to include parts of our results [SV12, BSTV14, AHLV14] in this work. I also thank Tobias Mai for valuable discussions concerning the algorithm in [BMS13] and Mario Diaz, Cristos Ruiz and José Brito, for more general discussions.

I am very grateful to Prof. Víctor Pérez-Abreu, for his valuable mentoring, and to César Octavio Pérez Carrizales, for his unflagging efforts on igniting the passion for Mathematics on young students and teachers.

I thank also the current and former participants of our Free Probability Seminar in Saarbrücken: Moritz Weber, Pierre Tarrago, Jonas Wahl, Cedric Schönard, Simon Jäger, John Williams, Amaury Fréslon, Mehmet Mandesoy and Guillaume Cébron.

I thankfully appreciate the support from the Mexican National Council of Science and Technology through the project CONACYT-DAAD ref. 214839/310129., and from the Deutsche Forschungsgemeinschaft (DFG) through the project SP419/8-1.

Finally, I would like to thank my parents and my grandmother, Victoria, Francisco and Carmen, who I deeply love, respect and admire.

Saarbrücken, May 2015.

Contents

1 Introduction

The main topic of this thesis work is free probability and is divided in two parts: The combinatorics of free probability and its applications to random matrices.

More specifically, our main goal is to provide a direct and up-to-date access to many of the tools from free probability which can be used to understand eigenvalue distributions of large Hermitian random matrices. Our general algorithm allows to treat a large class of matrix models. We obtain it by stretching the applicability of the algorithm [BMS13] to more generalized operator-valued situations that arise in the context of rectangular spaces [BG09a, BG09b] and deterministic equivalents [SV12]. We test our algorithm against matrix models which have recently appeared in the literature [CD11].

We also address situations, such as the product of block-Wigner matrices [BSTV14] and block-linearly modified random matrices [ANV], where the distribution of the model is also given by means of free probability, but our general algorithm does not apply.

The author's intuition on random matrix theory is strongly influenced by the combinatorics of non-crossing partitions, which govern free probability. The relevance of cumulants and non-crossing partitions in random matrix theory can be observed already in the moment proofs for Wigner's semicircle law.

For this reason, we include a chapter on the combinatorics of free and non-commutative probability, where we survey on the basic combinatorial tools, developed by Speicher (see [NS06] for a comprehensive study), but we also present new formulas [AV12, AHLV14] which link non-commutative probability with a variety of combinatorial structures.

Our work will be motivated by random matrices from the very beginning, we will not follow Voiculescu's original, operator-algebraic approach to free probability.

Combinatorics of non-commutative probability

The second and shortest part of this work is based on Speicher's [Spe94] combinatorial approach to free probability, which relies on the notion of free cumulants and the Möbius inversion on the lattice of non-crossing partitions. The free cumulants, (and cumulants in general), are very important since they linearize convolutions of free non-commutative random variables.

One of our main contributions to the theory is the extension of the formulas which relate the different types of cumulants and certain classes of set partitions. Several interesting combinatorial objects were used in [AHLV14] to obtain these formulas.

In addition, we obtain a formula which allows us to compute the product of k-tuples of free (or Boolean independent) random variables in terms of free (or Boolean) cumulants and k-divisible non-crossing partitions. In [AV12] we used this results, together with

enumerative formulas for k-divisible partitions, to bound free multiplicative convolutions and re-derive certain limit theorems.

The first and main part of this thesis elaborates on the application of free probability theory to obtain eigenvalue distributions of several classes of Hermitian Random Matrix models.

Applications of free probability to Hermitian random matrix models

The starting point of our investigations is the article [SV12], where we defined the free-probabilistic equivalents (FDE) of polynomial random matrix models. The FDE is a concrete free-probabilistic operator $P^{\square}$, whose distribution is shown to match the one given by the usual, deterministic equivalent (DE). Deterministic equivalents are simplifications of distributions of large random matrix models. These simplifications are usually defined at the level of Cauchy-Stieltjes transforms and are now widely used in the community of random matrices and wireless communications.

As one of the main results in this work, we obtain a general algorithm, (based on the one in [BMS13]) to compute approximations of the eigenvalue distribution for most of the models which were introduced in [SV12]. Our algorithm covers in particular several recent examples from the literature (see [CD11], Chapter 6), but also many more. With respect to previous methods, one of our most notable gains is on the simplicity of the fixed point equations that we obtain to draw the distributions.

Another important question is to minimize the computational complexity of the algorithm to draw these densities. For general situations, the additive algorithm described in [BMS13] seems to be quite powerful. We address the case of block-modified random matrices [ANV], an example which can be solved numerically by a different type of fixed point equation, which is given in terms of operator-valued free multiplicative convolutions, which was developed in [BSTV14], where we also covered the case of products of operator-valued semi-circulars.

Although there are some examples in [CD11] which do not fit into our general definition of FDE's, the equations obtained by other methods still correspond to quite remarkable operators in free probability theory, as it was noticed already in [Shl96]. We expect that a more detailed analysis of the combinatorial structures which appear while computing operator-valued distributions will eventually explain all these non-polynomial cases more naturally (without going into the ad-hoc solutions that we sketch later in this work).

Free deterministic equivalents of Hermitian random matrix models

Many Hermitian random matrix models in the literature come from considering a (self-adjoint) polynomial

$$P(x_1, \ldots, x_n, x_1^*, \ldots, x_n^*)$$

in non-commutative indeterminates $x_1, \ldots, x_n$ (and their adjoints $x_1^*, \ldots, x_n^*$), and then evaluating it on some random and deterministic matrices.

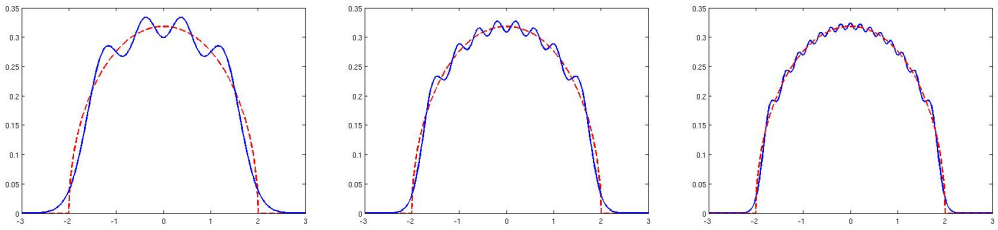

Fig. 1.1: Averaged eigenvalue distributions (AED) of normalized GUE (solid) vs Wigner's Law (dashed) for $N = 4, 8, 15$.

One of the most important objects related to a Hermitian random matrix is its averaged eigenvalue distribution (AED). The task of explicitly computing eigenvalue distributions of random matrices is quite hard in general.

A prominent example where such computation is possible is that of the standard Gaussian unitary ensemble (GUE). Such matrices $X_N = (x_{ij})_{i,j \leq N}$, consisting of independent (modulo the self-adjointness condition $X_N = X_N^*$) complex Gaussian entries, enjoy beautiful symmetries which ultimately lead to an expression for the AED μ_{X_N} in terms of the Hermite polynomials (defined recursively by $H_0(t) := 1, H_1(t) := t, H_k(t) := tH_{k-1}(t) - (k-1)H_{k-2}(t)$). One obtains:

$$d\mu_{X_N}(t) = \sum_{k=0}^{N-1} \frac{(H_k(t))^2 e^{-t^2/2}}{N\sqrt{2\pi}k!}.$$

If we re-normalize the entries ($\mathbb{E}(x_{ij}x_{ji}) = 1/N$), so that variance of the distributions μ_{X_N} remains fixed, it was noticed by Wigner [Wig58], that the probability measures μ_{X_N} converge as $N \to \infty$ to the semicircle law (see Fig 1.1).

GUE matrices belong to the larger class of Wigner random matrices, whose entries are i.i.d. but not necessarily Gaussian. For general Wigner matrices, there is no explicit expression for μ_{X_N}. As the size of the matrix grows large, however, the distributions μ_{X_N} converge towards the semicircle law. In fact, not only the AED, but also the empirical eigenvalue distributions (EED) converge to the semicircle law.

Of course, N needs to be quite large to see some structure on the EED's, whereas the convergence of the AED can be observed already for relatively small N. The speed of convergence does depend on the choice of the distribution (see Figs. 1.2 and 1.3).

Another important class of random matrices are Haar-distributed unitary matrices. One may construct more general Hermitian random matrices by adding and/or multiplying independent Wigner matrices, Haar-unitary random matrices and deterministic matrices.

Following a recent example (see [CHD11]), let us consider a model $P = QURU^*Q^* + SVTV^*S^*$, where Q, S, R, T are deterministic matrices of sizes 5×8, 5×4, 8×8 and 4×4, respectively, and $U \in \mathcal{U}(8)$, $V \in \mathcal{U}(4)$ are unitary matrices chosen independently with uniform distribution on the compact unitary groups $\mathcal{U}(8)$ and $\mathcal{U}(4)$. Suppose that

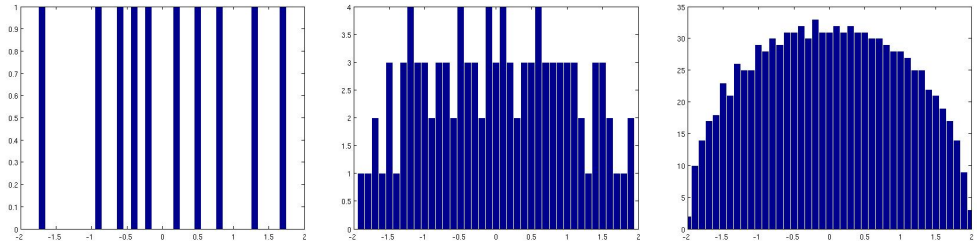

Fig. 1.2: Empirical eigenvalue distribution (EED) of one realization of a Wigner matrix with independent Bernoulli entries of size $N = 10, 100, 1000$.

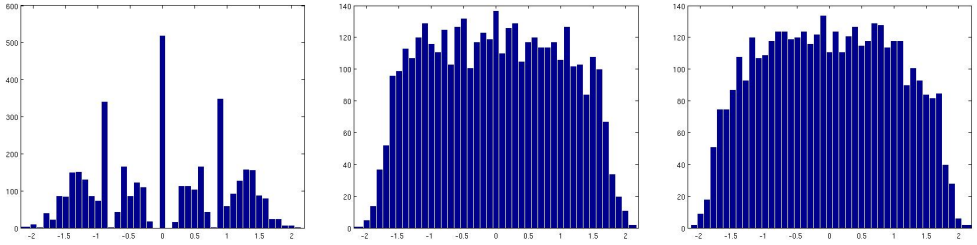

Fig. 1.3: 4000 Eigenvalues of Wigner matrices with independent Bernoulli entries of size $N = 5, 10, 20$.

$R = R^*$ and $T = T^*$, so that P is self-adjoint. An approximation of its distribution was computed in [CHD11] using the method of deterministic equivalents.

In the Wigner case, we decided to approximate the AED of our finite random matrix by the semicircular distribution, which is also the asymptotic distribution of the Wigner model. Hence, if our original Wigner matrix was large, the semicircular distribution is a good estimate for the AED. We would like to mimic this procedure.

However, we face now deterministic and Haar-distributed unitary matrices in our model and hence it is not even clear how to produce a growing sequence of random matrices whose asymptotic distribution converges to some distribution (which should additionally serve as an approximation of the distribution of the original model).

We can try blowing-up the model by considering $P_N = Q_N U_N R_N U_N^* Q_N^* + S_N V_N T_N V_N^* S_N^*$, where $A_N := A \otimes I_N$ for $A \in \{Q, R, S, T\}$ and letting $U_N \in \mathcal{U}(8N)$ $V_N \in \mathcal{U}(4N)$ be independent, with uniform distribution.

One observes that, indeed, the measures μ_{P_N} converge towards a deterministic shape. If we start with larger matrices, then the *original* model P_1 will be close to this asymptotic deterministic shape (see Fig. 1.4). In our example, the sizes for P_1 were maybe still too small, but we observe that the histograms of P_{10} and P_{40} are already very similar. The asymptotic distribution of the (quite artificially constructed) blown-up model P_N gives exactly the same distribution obtained in [CHD11] by the method of deterministic equivalents.

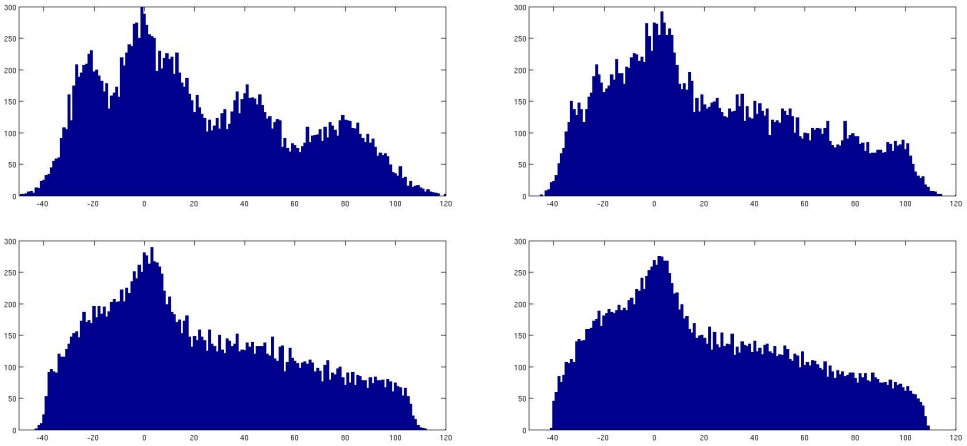

Fig. 1.4: 20000 Eigenvalues of $P_N = Q_N U_N R_N U_N^* Q_N^* + S_N V_N T_N V_N^* S_N^*$ for $N = 1$ (up-left), 3 (up-right), 10 (down-left), 40 (down-right).

The advantage of our approach is that the blowing-up procedure can be done for any polynomial evaluated on deterministic matrices and Wigner and Haar-unitary random matrices. In addition, we can describe the asymptotic (non-commutative) joint distribution of the blown-up matrices $\{Q_N, R_N, S_N, T_N, U_N, V_N\}$ in the language of (operator-valued) free probability. By looking at Voiculescu's results [Voi91, Voi98] (and their generalization to rectangular spaces [BG09a, BG09b]), it becomes clear that the blowing-up procedure described above will always converge in AED distribution. In particular, we can construct operators q, r, s, t, u, v in a (rectangular) non-commutative probability space, which satisfy certain freeness relations and such that the limiting shape of Fig. 1.4 is the spectral distribution of the free deterministic equivalent $P^{\square} := quru^*q^* + svtv^*s^*$. After this simplification, we rely on the most advanced tools in operator-valued free probability to compute numerically the distribution $\mu_{P^{\square}}$ (see Fig 1.5).

After the pioneering result of Wigner, a long list of particular cases of this phenomenon where studied in the following years. These works usually fixed a specific polynomial P (and typically also the specific nature of the input matrices $X_1, \ldots, X_n$):

Marchenko and Pastur [MP67] considered the case $P(x) = xx^*$ where x is replaced by a non-self-adjoint Wigner Matrix. More complicated models were considered by Girko [Gir01], Bai and Silverstein [BS06] Tulino and Verdu [TV04], Müller [Mue02], Moustakas and Simon [MS05], Hechem, Loubaton and Najim [HLN07] Coulliet and Debbah [CD11]. Some cases were treated using operator-valued free probability [SV12], [BSTV14].

The general case for the polynomials $p(x, y) = x + y$ and $p(x, y) = xy^2x$ was studied by Voiculescu in [Voi86] and [Voi87], respectively. The general solution to the case of the free commutator, $p(x, y) = i(xy - yx)$, was given by Nica and Speicher in [NS98]. The general solution to the anti-commutator $p(x, y) = xy + yx$ was found by Vasilchuck [Vas03].

A notable step towards a general solution of this random matrix problem was found

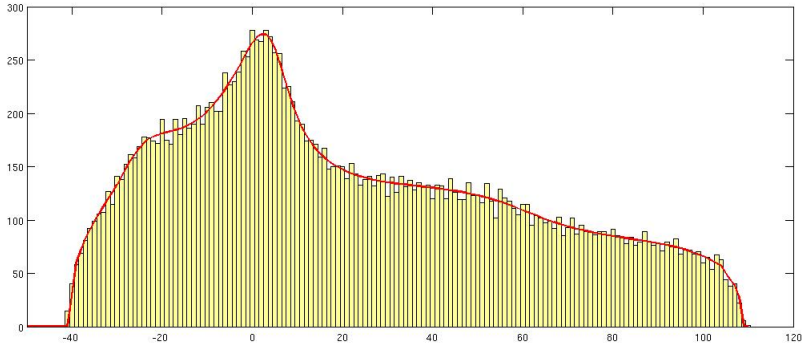

Fig. 1.5: 20000 eigenvalues from 100 realizations of P_{40} (histogram) vs distribution of the free deterministic equivalent (FDE) $P^{\square}$, computed with our algorithm (solid).

recently by Belisnchi, Mai, and Speicher. In [BMS13], they provided an algorithm to compute the limiting distribution of any polynomial evaluated on self-adjoint random matrices of the same size which are asymptotically free.

The main goal of this thesis work is to extend this result in several directions:

- We allow the matrices to be non-self-adjoint. We only require the polynomial $P(x_1, \ldots, x_n, x_1^*, \ldots x_n^*)$ to be self-adjoint.
- The deterministic matrices on which we evaluate the polynomial are allowed to preserve correlations. In particular, they need not to be all asymptotically free.
- The input matrices may have different sizes and deterministic matrices can be rectangular. We only ask that (after evaluating on $X_1, \ldots, X_n^*$) the monomials of P should be square and have the same size.

Our algorithm relies heavily on the recent trends on free probability. In particular, on the possibility of computing operator-valued free additive convolutions [BMS13].

In [Voi85], Voiculescu launched the theory of free probability in order to tackle problems on von Neumann algebras. He observed that, in the context of quantum probability spaces, some operators appearing in group von Neumann algebras satisfy a particular relation (which he called "freeness") that should be thought as a new, non-commutative notion of stochastic independence. In particular, replacing independence by freeness yields a free version of the Central limit theorem, where the role of the Gaussian distribution is now played by the semicircle law.

Some years later, in his seminal work [Voi91], he found out that some notable limiting distributions of random matrices are just the spectral distributions of certain combinations of free non-commutative random variables. For some cases, it was possible to use analytical tools [Voi86, Voi87] to obtain explicitly those spectral distributions.

During its earlier stages, free probability provided a fresh, new way to look at large random matrices, but failed to address successfully the latest, more sophisticated matrix models which were considered, for instance, for modeling wireless communications.

These recent models were usually treated using the method of deterministic equivalents. A free probabilistic interpretation of such deterministic equivalents was missing. In addition, no satisfactory limiting operator could be assigned to matrix models involving rectangular matrices.

The introduction of the operator-valued version of free probability [Voi95] (see also [Spe98]), and the corresponding, broader definition of free independence, allowed to extend this conceptual view to more general situations, such as band and block matrices [Shl96],[FOBS08], and rectangular matrices with different sizes [BG09a, BG09b].

In [SV12], we used Bennaych-Georges rectangular spaces to give a free probabilistic interpretation of deterministic equivalents. The usual method to obtain a deterministic equivalent involves some intuitive guess and hence it is not clear how to define it in a general situation. In contrast, our free deterministic equivalent is always defined as a concrete operator $P^{\square}$. We showed that the spectral distribution of $P^{\square}$ is, for all the considered models, exactly the distribution obtained by the method of deterministic equivalents. However, although we made clear how to compute the moments of $\mu_{P^{\square}}$, we still had to obtain the distribution in an ad-hoc way, depending on the specific matrix model.

Recent important developments [Bia98, Voi00, Voi02, Dyk06, HFS07, BSTV14] on the analytical tools of operator-valued free probability allowed to compute distributions of quite general operators. In combination with the linearizion trick [HT05, And11], these led to the general solution for the asymptotic eigenvalue distribution of arbitrary polynomials on asymptotically free square random matrices of the same size [BMS13].

For practical purposes, this asymptotic freeness condition is quite strong: it means that all the involved matrices should be in "generic position" with respect to each other. The latest models considered by the wireless communication community tend to break this assumption: One assumes that some collections of antennas (either from the transmitter or the receiver) remain always close and hence the noise is not strong enough to break the correlations between the corresponding channels. However, the channels corresponding to different (distant) collections are more affected by the noise and hence can be thought as being in general position (free) with respect to each other.

The condition of having all matrices of the same size is also challenged by the latest models, as the transmitters and receivers do not generally have the same number of antennas. Our generalization fixes both problems.

We show that the algorithm described in [BMS13] can be extended to compute the distribution of the free deterministic equivalent $P^{\square}$, associated with most matrix models formed by Wigner matrices, Haar-unitary random matrices and rectangular deterministic matrices of different sizes.

Main Result 1. *Let $P = P(x_1, x_1^*, \ldots, x_n, x_n^*) = P(x_1, \ldots, x_n)$ be a self-adjoint polynomial in non-commutative indeterminates $x_1, \ldots, x_n$ (and their adjoints $x_1^*, \ldots x_n^*$). For each $m \geq 1$, let $Y_1^{(m)}, \ldots, Y_n^{(m)}$ be independent (random and deterministic) matrices such that $Y_i^{(m)}$ is either an $mN_i \times mN_i$ Wigner or Haar-unitary random matrix or $Y_i^{(m)} = Y_i^{(1)} \otimes I_m$ is a $mN_i \times mM_i$ deterministic matrix. Let $P_m = P(Y_1^{(m)}, \ldots, Y_n^{(m)})$. Assume that P and $Y_1^{(m)}, \ldots, Y_n^{(m)}$ are such that all the summands in P_m have the same*

size and the sizes of consecutive matrices on each monomial fit. Let μ_m be the AED of the self-adjoint random matrix P_m. Then:

- *There exists a probability measure μ such that $\mu_m \to \mu$ weakly.*
- *There exist operators $y_1, \ldots, y_n$, such that μ is the spectral distribution of the free deterministic equivalent $P^\square = P(y_1, \ldots, y_n)$ of P.*
- *μ can be numerically computed.*

In practice, one is usually interested in computing the distribution of $P = P_1$ and not the asymptotic distribution of P_m. However, if one tries to obtain systems of equations defining (the Cauchy-Stieltjes transform of) μ_{P_1}, these are usually not closed and hence it is not possible to solve them with fixed point algorithms. The usual procedure up to now was to simplify these equations, inspired on the asymptotic behavior of the model, in order to obtain a closed system, called "deterministic equivalent", which could be then solved by fixed point algorithms.

We show that these deterministic equivalents and our free deterministic equivalents determine the same distributions. Furthermore, we completely understand the difference between the original model P_1 and its FDE $P^\square$ in terms of the (non-commutative, rectangular) joint distributions of $Y^{(N)} := (Y_1^{(N)}, \ldots, Y_n^{(N)})$ and $Y^\square := (y_1, \ldots, y_n)$:

- If $Y_i^{(1)}$ is Wigner (resp. Haar-unitary), y_i is a semicircular (resp. Haar-unitary) element when properly compressed in a rectangular probability space.
- If $(Y_i^{(1)})_{i \in I}$ are the deterministic matrices of the model, then $(y_i)_{i \in I}$ are just properly embedded copies of $(Y_i^{(1)})_{i \in I}$.
- The algebras $\langle y_i : i \in I \rangle$ and $\langle y_i : i \notin I \rangle$ are free (with amalgamation over a suitable algebra of projections). So are also $y_{i_1}, \ldots, y_{i_k}$ for $i_1, \ldots, i_k \notin I$.

Hence, each collection $Y^{(N)} = (Y_1^{(N)}, \ldots, Y_n^{(N)})$ in the sequence $(Y^{(N)})_{N \geq 1}$ is an improved version of $(Y^{(1)})$, where its deviations from freeness (e.g. contributions of crossing partitions) are penalized increasingly with N until freeness is achieved in the limit $Y^\square = (y_1, \ldots, y_n)$. If the norm of the deterministic matrices is fixed and the sizes of all the matrices are large, then the behavior of $(Y^{(1)})$ will already be close to that of $Y^\square$, and hence $P^\square$ will be a good approximation of P_1.

In order to reduce the number of variables which form $P^\square$ (and hence the number of iterations required to obtain $\mu_{P^\square}$), it will be convenient to recognize families of randomly rotated matrices which become asymptotically free due to these random rotations. We will illustrate our algorithm with several examples.

Organization of the work

To motivate our investigations, we give in Chapter 2 a proof of Wigner's semicircle law for the Gaussian case. This will allow us to single out important components of the

proof that we should have present when we discuss several generalizations along further chapters.

In Chapter 3 we introduce the basic concepts of free probability motivated by the natural occurrence of freeness among important matrix models. The main purpose of this chapter is to provide some intuition that will allow a better understanding for both the operator-valued analogues in Chapter 5, and our main definitions in Chapter 6.

In Chapter 4 we present the combinatorial approach to scalar-valued free probability. We include our results on cumulant-to-cumulant formulas and our description of the cumulants of products in terms of k-divisible partitions.

Chapter 5 covers the basics on Operator-valued free probability and recalls the numerical algorithm to compute operator-valued free additive convolutions. We specialize on rectangular and matrix-valued non-commutative spaces. We also compute the product of operator-valued semi-circulars [BSTV14] and the case of block-modified matrices [ANV].

In Chapter 6 we recall and extend the definition of the free deterministic equivalents from [SV12].

Chapter 7 explains our algorithm to obtain spectral distributions of FDE's. We show how it applies to the polynomial matrix models from [CD11].

2 Motivation: classical cumulants, Wick calculus and Wigner's semicircle law

In this Chapter we give the proof of the simplest version of Wigner's Semicircle Law, namely, the Gaussian case. This will serve us for future reference when dealing with several generalizations of this fundamental result.

2.1 Moments vs cumulants

Let $X : \Omega \to \mathbb{R}$ be a random variable in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The moments of X are the values

$$\mathbb{E}(X^n) = \int_\Omega (X(\omega))^n d\mathbb{P}(\omega).$$

For a large class of random variables, which includes Gaussian random variables and random variables with compact support, the moments of X determine its distribution. For several random variables $X_1, \ldots, X_k : \Omega \to \mathbb{R}$, we may consider the mixed moments

$$\mathbb{E}(X_1^{n_1} \cdots X_k^{n_k}) = \int_\Omega (X_1(\omega))^{n_1} \cdots (X_k(\omega))^{n_k} d\mathbb{P}(\omega).$$

If $X_1, \ldots, X_k$ are determined by their moments, then the stochastic independence of $X_1, \ldots, X_k$ is equivalent to the fact that, for all $n_1, \ldots, n_k \geq 0$, the mixed moments factorize

$$\mathbb{E}(X_1^{n_1} \cdots X_k^{n_k}) = \mathbb{E}(X_1^{n_1}) \cdots \mathbb{E}(X_k^{n_k}). \tag{2.1}$$

For a collection of random variables $X_1, \ldots, X_k$, we define the (multivariate) classical cumulants $K_n(X_{i_1}, \ldots X_{i_n})$, $n \leq 1$, $i_1, \ldots, i_n \leq k$, recursively as the collection of multilinear functionals $(K_n)_{\geq 1}$ which satisfy the moment-cumulant formula:

$$\mathbb{E}(X_{i_1} \cdots X_{i_n}) = \sum_{\pi \in \mathcal{P}(n)} K_\pi(X_{i_1}, \cdots, X_{i_n}),$$

where $\mathcal{P}(n) := \mathcal{P}([n])$ are the set partitions; $\mathcal{P}(n)$ is the power set of $[n] := \{1, 2, \ldots, n\}$ and each element $\pi = \{V_1, \ldots, V_{|\pi|}\} \in \mathcal{P}([n])$ decomposes the set $[n] = V_1 \cup V_2 \cup \cdots \cup V_{|\pi|}$ into non-empty, pairwise disjoint subsets ("blocks") $V_1, \ldots, V_{|\pi|}$ (see Chapter 4 for more definitions), and we write

$$K_\pi(X_1, \ldots, X_n) := \prod_{V \in \pi} K_{|V|}(X_V),$$

where we use the notation

$$K_{|V|}(X_V) := K_m(X_{v_1}, \ldots, X_{v_m})$$

for each block $V = \{v_1, \ldots, v_m\} \in \pi$, $v_1 < \cdots < v_m$. We highlight the case when we deal with a single random variable X with the notation:

$$\kappa_n(X) := K_n(X, X, \ldots, X)$$

For example, $K_1(X_i) = \mathbb{E}(X_i)$ is simply the mean and $K_2(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)$ is the covariance.

For an ordered tuple $\mathbf{X} = (X_1, \ldots, X_k)$ of random variables, we call

$$\Phi_m^{\mathbf{X}} := \Phi_m = \{(i_1, \ldots, i_m) \mapsto \mathbb{E}(X_{i_1} \ldots X_{i_m}) : i_1, \ldots, i_m \leq k\}$$

the m-th order mixed moments of $(X_1, \ldots, X_k)$. Analogously we define

$$\Psi_m^{\mathbf{X}} := \{(i_1, \ldots, i_m) \mapsto K_m(X_{i_1}, \ldots, X_{i_m}) : i_1, \ldots, i_m \leq k\}$$

the m-th order cumulant. The collection of moment maps $(\Phi_m^{\mathbf{X}})_{m \leq n}$ contains exactly the same information as the collection of mixed cumulants $(\Psi_m^{\mathbf{X}})_{m \leq n}$. However, cumulants seem to encode statistical information in a nicer way.

A real random variable X is the constant random variable $c \in \mathbb{R}$ iff all cumulants of degree $n \geq 2$ vanish and $\kappa_1(X) = c$. In fact, it is not hard to see that, if we input a constant in any of the arguments of a cumulant of order $k \geq 2$, then the cumulant must vanish, independently from the position of the constant argument and the rest of the arguments.

In terms of cumulants, the simplest (non-constant) random variables are the Gaussian random variables: X has the Normal distribution $\mathcal{N}(\mu, \sigma^2)$ iff all cumulants of degree $n \geq 3$ vanish, $\kappa_2(X, X) = \sigma^2$ and $\kappa_1(X) = \mu$.

Two random variables X, Y are independent if and only if all the mixed cumulants (i.e. $K_2(X, Y), K_2(Y, X), K_3(X, X, Y), K_3(X, Y, X), K_3(X, Y, Y), \ldots$ etc.) vanish. This implies in particular that the cumulants of $X + Y$ are simply

$$K_n(X + Y, X + Y, \ldots, X + Y) = K_n(X, \ldots, X) + K_n(Y, \ldots, Y).$$

Hence, cumulants can be used to compute additive convolutions (in fact, the cumulants are related to the coefficients of Fourier transforms, see Chapter 4).

If X, Y are independent standard Gaussian ($\mathcal{N}(0, 1)$) random variables, then $Z := X + iY$ has the standard complex Gaussian distribution, which can also be characterized in terms of the mixed cumulants of Z and $Z^* := \bar{Z}$. The only non-vanishing cumulants are

$$K_2(Z, Z^*) = K_2(Z^*, Z) = 1.$$

Wick's formula for independent complex Gaussians $Z_1, \ldots, Z_k$, states that, for any $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{1, *\}^n$, we have

$$\mathbb{E}(Z_{i_1}^{\varepsilon_1} \ldots Z_{i_n}^{\varepsilon_n}) = \sum_{\pi \in \mathcal{P}_2(n)} \kappa_\pi(Z_{i_1}^{\varepsilon_1}, \ldots, Z_{i_n}^{\varepsilon_n}), \tag{2.2}$$

where $\mathcal{P}_2(n) \subset \mathcal{P}(n)$ denotes the subset of matchings (i.e. partitions $\pi \in \mathcal{P}(n)$ such that every $V \in \pi$ has exactly 2 elements). Therefore, we observe that the Wick formula for this case can be easily derived from the moment-cumulant formula. The restriction to pairings $\mathcal{P}_2(n) \subset \mathcal{P}(n)$ follows from the fact that only the second order cumulants of standard complex Gaussians may not vanish.

However, going over cumulants to compute $\mathbb{E}(Z_{i_1}^{\varepsilon_1} \dots Z_{i_n}^{\varepsilon_n})$ seems more like a detour in this case. We could simply compute $\mathbb{E}(Z_{i_1}^{\varepsilon_1} \dots Z_{i_n}^{\varepsilon_n})$ by a direct application of the factorization in eq. (2.1). Our Wick formula (2.2) will be very useful when we go over to random matrices.

2.2 Gaussian matrices and Wigner's semicircle law

One of the simplest ways of constructing a random matrix is to let each entry be an independent copy of a given random variable X. The distribution of X induces a probability measure $\mathbb{P}$ on a set $\Omega \subseteq M_N(\mathbb{C})$ of matrices. In particular, we may consider an $N \times N$ matrix $Z = Z_N := (\frac{1}{\sqrt{N}} z_{ij})_{i,j \leq N}$ with independent standard complex Gaussian entries. Such random matrices are called non-self-adjoint Gaussian matrices (the choice of the normalization $\frac{1}{\sqrt{N}}$ will be clear later).

In [Wig58] Wigner described the asymptotic ($N \to \infty$) eigenvalue distribution of the (necessarily real) eigenvalues of $X_N = (Z_N + Z_N^*)/\sqrt{2}$ (where A^* denotes the Hermitian transpose of A).

Since $X(\omega) = (X(\omega))^*$ for any realization of X, we can diagonalize

$$X(\omega) = U(\omega) D(\omega) (U(\omega))^*,$$

where $D = diag(\lambda_1(\omega), \dots, \lambda_N(\omega))$ and hence for all $k \geq 0$, we have

$$\begin{aligned} \frac{1}{N}\mathrm{Tr}(X^k(\omega)) = \frac{1}{N}\mathrm{Tr}((U(\omega)D(\omega)(U(\omega))^*)^k) &= \frac{1}{N}\mathrm{Tr}(U(\omega)D(\omega)^k(U(\omega))^*) \quad (2.3) \\ &= \frac{1}{N}\mathrm{Tr}(D(\omega)^k) \quad (2.4) \\ &= \frac{1}{N}\sum_{i \leq N}(\lambda_i(\omega))^k. \quad (2.5) \end{aligned}$$

The expression $\frac{1}{N}\sum_{i \leq N}(\lambda_i(\omega))^k$ can be identified as the k-th moment of the real random variable $\Lambda(X(\omega))$, with discrete probability measure $\mu_{X(\omega)}$ which assigns a mass of $1/N$ to each eigenvalue of $X(\omega)$. The averaged eigenvalue distribution μ_X is the distribution of the random variable $\Lambda(X)$ obtained by averaging all such $\Lambda(X(\omega))$, $\omega \in \Omega$ against $\mathbb{P}$. More specifically, it is the probability measure μ_{X_N} with k-th moment $\frac{1}{N}\mathbb{E}(\sum_{i \leq N} \lambda_i^k)$. We want to describe μ_{X_N} when the size $N \to \infty$. Let us denote $\tau_N := \frac{1}{N}\mathrm{Tr}$.

Instead of computing the (rather complicated) joint distributions of $(\lambda_1, \lambda_2, \dots, \lambda_N)$, we compute the moments

$$\mathbb{E} \circ \tau_N(X^k(\omega)) = \frac{1}{N}\mathbb{E}(\sum_{i \leq N} \lambda_i^k),$$

which depend on the entries of our matrices in a polynomial way, hoping that we will be able to identify them as moments of a probability measure. Indeed, this will be the case, and for this, it will be convenient to work first directly with Z and Z^* and symmetrizing only at the very end.

In general, we would like to compute, for all $k \geq 1$ and every $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{1, *\}^k$

$$\mathbb{E} \circ \tau_N(Z^{\varepsilon_1}, \ldots, Z^{\varepsilon_k}).$$

For $k = 1$ we get

$$\mathbb{E} \circ \tau_N(Z) = \frac{1}{N^{3/2}} \sum_{i \leq N} \mathbb{E}(z_{ii}) = 0 = \mathbb{E} \circ \tau_N(Z^*).$$

If $\varepsilon = (1, *)$, we have

$$\mathbb{E} \circ \mathrm{Tr}(ZZ^*) = \frac{1}{N^2} \sum_{i_1, i_2 \leq N} \mathbb{E}(z_{i_1 i_2} \bar{z}_{i_1 i_2}) = 1. \tag{2.6}$$

If $\varepsilon = (1, 1)$, we have

$$\mathbb{E} \circ \tau_N(ZZ) = \frac{1}{N^2} \sum_{i_1, i_2 \leq N} \mathbb{E}(z_{i_1 i_2} z_{i_2 i_1}) = 0, \tag{2.7}$$

since $\mathbb{E}(z_{i_1 i_2} z_{i_2 i_1}) = \kappa_2(z_{i_1 i_2}, z_{i_2 i_1}) = 0$ by the characterization of the mixed cumulants of independent standard complex Gaussians.

Similarly one can see that $\mathbb{E} \circ \tau_N(Z^* Z) = 1$ and $\mathbb{E} \circ \tau_N(Z^* Z^*) = 0$. Hence, the mean of μ_{X_N} is $\frac{1}{\sqrt{2}} \mathbb{E} \circ \tau_N(Z + Z^*) = 0$. The variance of μ_{X_N} is then just the second moment

$$\frac{1}{2} \mathbb{E} \circ \tau_N((Z + Z^*)(Z + Z^*)) = \frac{1}{2} \mathbb{E} \circ \tau_N(ZZ^* + Z^* Z) = 1,$$

for all N. This explains the choice of normalization $Z = (\frac{1}{\sqrt{N}} z_{ij})_{i,j \leq N}$.

Let us now consider the fourth order mixed moment $\varepsilon = (1, *, *, 1)$, we have

$$\mathbb{E} \circ \tau_N(ZZ^* Z^* Z) = \frac{1}{N^3} \sum_{i_1, \ldots, i_4 \leq N} \mathbb{E}(z_{i_1 i_2} \bar{z}_{i_3 i_2} \bar{z}_{i_4 i_3} z_{i_4 i_1}) \tag{2.8}$$

By Wick's Formula,

$$\mathbb{E}(z_{i_1 i_2} \bar{z}_{i_3 i_2} \bar{z}_{i_4 i_3} z_{i_3 i_1}) = \sum_{\pi \in \mathcal{P}_2(4)} \kappa_\pi(z_{i_1 i_2}, \bar{z}_{i_3 i_2}, \bar{z}_{i_4 i_3}, z_{i_4 i_1}).$$

There are 3 pairings of $\{1, 2, 3, 4\}$, namely $\pi_1 = \{\{1, 2\}\{3, 4\}\}$, $\pi_2 = \{\{1, 3\}\{2, 4\}\}$, $\pi_3 = \{\{1, 4\}\{2, 3\}\}$. Since the z_{ij}'s are complex Gaussian random variables, π_3 will

vanish (independently of the choice of $i_1, \ldots, i_4$) since it will never match a z_{ij} with $\bar{z}_{ij}$, which is a necessary condition for the cumulant not to vanish. Hence

$$\begin{aligned} \mathbb{E} \circ \tau_N(ZZ^*Z^*Z) &= \frac{1}{N^3} \sum_{i_1,\ldots,i_4 \leq N} \kappa_{\pi_1}(z_{i_1 i_2}, \bar{z}_{i_3 i_2}, \bar{z}_{i_4 i_3}, z_{i_4 i_1}) \quad (2.9) \\ &+ \frac{1}{N^3} \sum_{i_1,\ldots,i_4 \leq N} \kappa_{\pi_2}(z_{i_1 i_2}, \bar{z}_{i_3 i_2}, \bar{z}_{i_4 i_3}, z_{i_4 i_1}) \quad (2.10) \end{aligned}$$

Since the entries are independent standard complex Gaussians, each partition imposes restrictions on the indices for the cumulants not to vanish, namely

$$\begin{aligned} \kappa_{\pi_1}(z_{i_1 i_2}, \bar{z}_{i_3 i_2}, \bar{z}_{i_4 i_3}, z_{i_4 i_1}) = \kappa_2(z_{i_1 i_2}, \bar{z}_{i_3 i_2})\kappa_2(\bar{z}_{i_4 i_3}, z_{i_4 i_1}) &= \delta_{i_1 i_3}\delta_{i_3 i_1}, \\ \kappa_{\pi_2}(z_{i_1 i_2}, \bar{z}_{i_3 i_2}, \bar{z}_{i_4 i_3}, z_{i_4 i_1}) = \kappa_2(z_{i_1 i_2}, \bar{z}_{i_4 i_3})\kappa_2(\bar{z}_{i_3 i_2}, z_{i_4 i_1}) &= \delta_{i_1 i_4}\delta_{i_3 i_2}\delta_{i_3 i_4}\delta_{i_2 i_1}, \end{aligned}$$

Hence we need only to count the number of free indices in order to obtain the contribution of each partition. For this case we obtain

$$\mathbb{E} \circ \tau_N(ZZ^*Z^*Z) = 1 + 1/N^2.$$

As $N \to \infty$ only the contribution of π_1 will survive. For a general moment of order k, an easy inductive argument shows that a pairing $\pi \in \mathcal{P}_2(k)$ can only contribute in the limit if $\pi \in \mathcal{NC}_2(k) \subset \mathcal{P}_2(k)$ is a non-crossing pairing (i.e. there is no quadruple $1 \leq a < b < c < d \leq k$ such that $a, c \in V_i, b, d \in V_j$ where $V_i \neq V_j$ are blocks of π). In addition, we must have that, for each matching $\{a, b\} \in \pi$, $\varepsilon_a \neq \varepsilon_b$ (in contrast to π_3 above). Hence, for computing the asymptotics of a general moment we need to find

$$\mathbb{E} \circ \tau_N(Z^{\varepsilon_1}, \ldots, Z^{\varepsilon_k}) = \frac{1}{N^{1+k/2}} \sum_{\substack{i_1,\ldots,i_k \leq N \\ \pi \in \mathcal{NC}_2(k)}} \kappa_\pi(z_{i_1 i_2}^{\varepsilon_1}, z_{i_2 i_3}^{\varepsilon_2}, \ldots, z_{i_k i_1}^{\varepsilon_k})) = |\mathcal{NC}_\varepsilon(k)|, \quad (2.11)$$

where $\pi \in \mathcal{NC}_\varepsilon(k) \subseteq \mathcal{NC}_2(k)$ iff $\varepsilon_a \neq \varepsilon_b$ for all $\{a, b\} \in \pi$.

Now since $X = (Z + Z^*)/\sqrt{2}$, we have that

$$\mathbb{E} \circ \tau_N(X^k) = \frac{1}{N^{1+k/2} 2^{k/2}} \sum_{\substack{i_1,\ldots,i_k \leq N \\ \varepsilon = (\varepsilon_1,\ldots,\varepsilon_k) \in \{1,*\}^k \\ \pi \in \mathcal{NC}_2(k)}} \kappa_\pi(z_{i_1 i_2}^{\varepsilon_1}, z_{i_2 i_3}^{\varepsilon_2}, \ldots, z_{i_k i_1}^{\varepsilon_k})). \quad (2.12)$$

If we fix a non-crossing pairing $\pi \in \mathcal{NC}_2(k)$, there are $2^{k/2}$ non-vanishing choices for ε (for each block of $(a, b) \in \pi$, we can have $(\varepsilon_a, \varepsilon_b)$ equal to $(1, *)$ or $(*, 1)$). Hence, after summing over all free indices and all $\varepsilon \in \{1, *\}^k$, each non-crossing pairing contributes with 1 to the moment in the limit.

A nice way to count non-crossing pairings $\mathcal{NC}_2(2k) \subset \mathcal{P}_2(2k)$ is through the recursion

$$|\mathcal{NC}_2(2k)| = \sum_{i=1}^{k} |\mathcal{NC}_2(2i-2)||\mathcal{NC}_2(2(k-i))|,$$

which is obtained by fixing the pairing $(1, 2i)$. The non-crossing condition prevents the elements $\{2, 3 \ldots, 2i-1\}$ to be matched with elements in $\{2i+1, 2i+2, \ldots, 2k\}$ and hence the recursion follows. It is well-known that such recursion characterizes the Catalan numbers, hence $|\mathcal{NC}(2n)| = C_n := \frac{1}{n+1}\binom{n}{2n}$.

Since the Catalan numbers are the moments of the semicircular distribution supported on $[-2, 2]$ with density

$$d\mu(t) = \frac{1}{2\pi}\sqrt{4-t^2},$$

the assertion follows.

In the next chapter we introduce Voiculescu's free probability theory for non-commutative random variables, which allows to treat random matrices as random variables. The idea is to think of $\tau_N = \frac{1}{N}\mathbb{E} \circ \mathrm{Tr}$ as a generalization of $\mathbb{E}$ in classical probability.

3 Free probability theory and asymptotics of random matrices

In this chapter we introduce Voiculescu's notion of non-commutative probability spaces and free independence, starting with the most basic, algebraic approach. We realize that these abstract spaces model the asymptotic collective behavior of very concrete objects, such as random matrices.

3.1 Non-commutative probability spaces, Wigner and Haar-unitary random matrices

Definition 3.1.1. *A non-commutative probability space is a pair $(\mathcal{A}, \tau)$, where $\mathcal{A}$ is a complex algebra with unit and $\tau : \mathcal{A} \to \mathbb{C}$ is a unital linear functional. If $\mathcal{A}$ is endowed with an anti-linear involution $* : \mathcal{A} \to \mathcal{A}$ and $\tau(aa^*) \geq 0$, for all $a \in \mathcal{A}$, we call $(\mathcal{A}, \tau)$ a $*$-probability space.*

All the $*$-probability spaces considered in this work will be tracial (i.e. $\tau(ab) = \tau(ba)$ for all $a, b \in \mathcal{A}$).

An element $a \in \mathcal{A}$ is called a (non-commutative) random variable and $(\tau(a^k))_{k\geq 1}$ are known as the moments of a.

Example 3.1.2. *The algebra $\mathcal{A} := \mathcal{M}_N(\mathbb{C})$ of complex $N \times N$ matrices with the normalized trace $\tau = \tau_N = \frac{1}{N}\mathrm{Tr}$ is a $*$-probability space (with involution given by the Hermitian transpose). As we saw in the previous chapter, if $X \in \mathcal{M}_N(\mathbb{C})$ is Hermitian, we may diagonalize $X = UDU^*$ and realize that*

$$\tau(X^k) = \tau(UD^kU^*) = \tau(D^k) = \frac{1}{N}\sum_{i=1}^{N} \lambda_i(X)^k = \int_{\mathbb{R}} x^k \mu_X(dt),$$

where $\lambda_1(X) \leq \cdots \leq \lambda_N(X)$ are the eigenvalues of X and

$$\mu_X = \frac{1}{N}\sum_{i=1} \delta_{\lambda_i(X)}.$$

Hence $(\tau(a^k))_{k\geq 1}$ are the moments (in the usual, probabilistic sense) of the probability measure μ_X.

3.1.1 Wigner and Haar-unitary random matrices

A Wigner matrix X_N is an $N \times N$ self-adjoint random matrix $X_N = (N^{-1/2}x_{ij})_{i,j\leq N}$, where $x_{ji} = x_{ij}^*$, and such that the random variables

$$(x_{ii})_{i\leq N} \cup (\sqrt{2}\Re(x_{ij}))_{i<j\leq N} \cup (\sqrt{2}\Im(x_{ij}))_{i<j\leq N},$$

are independent and share the distribution of a given random variable x, with all moments. Note that the normalizations are such that $Var(x_{ij}) = \mathbb{E}(x_{ij}x_{ji}) = 1$.

Remark 3.1.3. *An $N\times N$ random matrix can be thought as an N^2-dimensional complex random variable X on a probability space $(\mathcal{M}_N(\mathbb{C}), \mathcal{Q}, \mathbb{P})$, where $\mathcal{Q}$ is the Borel σ-field of $\mathcal{M}_N(\mathbb{C})$ and $\mathbb{P}$ is a probability distribution. A random matrix X is self-adjoint iff $\mathbb{P}(X = X^*) = 1$.*

In view of Example 3.1.2, for each possible value $\omega \in \mathcal{M}_N(\mathbb{C})$ of X, we may consider the distribution μ_ω, and average these with the probability distribution $\mathbb{P}$.

The numbers $(\frac{1}{N}\mathbb{E}\circ\mathrm{Tr}(X_N^k))_{k\geq 0}$ (if they exist) are thus the moments of the averaged eigenvalue distribution μ_X which satisfies, for any Borel set $A \in \mathbb{R}$, that

$$\mu_X(A) = \int_{\omega\in\Omega} \mu_{X(\omega)}(A)d\mathbb{P}(\omega).$$

The averaged eigenvalue distribution (AED) μ_X is one of the most important objects of study in the field of random matrix theory.

For the Gaussian ensemble $(X_N)_{N\geq 1}$, we observed in Chapter 2 that the AED converges to the semicircle law.

More general versions of Wigner's Theorem state that the averaged eigenvalue distribution of a Wigner ensemble $(X_N)_{N\geq 1}$ converges weakly to the semicircle distribution, independently of the (identical) distribution of the entries (as long as x has moments of order $2 + \varepsilon$).

Another important example of random matrix are random unitary matrices.

The unitary group $\mathcal{U}(N) : \{U \in M_N(\mathbb{C}) : UU^* = 1 = U^*U\}$ is a compact Lie group, and hence there exists a unique (up to a scaling constant), Haar measure μ on the Borel sets of $\mathcal{U}(N)$, which is analogue to the Lebesgue measure in the sense that it is translation invariant and positive on any open subset. We may normalize so that $\mu(\mathcal{U}_N) = 1$ is a probability measure. A matrix $U \in \mathcal{U}(N)$ drawn uniformly according to μ will be called a Haar-distributed random matrix.

The uniform distribution of $U_N = (u_{ij})_{i,j\leq N}$ in $\mathcal{U}_N$ yields a (quite non-trivial) joint distribution of its entries. However, from the unitary invariance of μ we may deduce important properties of joint distribution of the u_{ij}. In particular, one can show that the mixed moments (see next section for a general definition) of $\{U_N, U_N^*\}$ are characterized by the fact that $U_N^*U_N = I_N = U_NU_N^*$ and

$$\mathbb{E}\circ\tau_N(U_N^k) = \delta_{0k}, \quad k \in \mathbb{Z}.$$

These two fundamental examples inspire the definition of some special non-commutative random variables.

Definition 3.1.4. *Let $(\mathcal{A}, \tau)$ be a $*$-probability space.*

An element $s = s^* \in \mathcal{A}$ is called semicircular, if $\tau(s^{2k+1}) = 0$ and $\tau(s^{2k}) = C_k$ for all $k \geq 1$.

An element $u \in \mathcal{A}$ is called a Haar-unitary, if $uu^* = u^*u = 1$ and $\tau(u^k) = 0$ for all $0 \neq k \in \mathbb{Z}$

So far we have only considered examples consisting of a single random matrix. We would like now to build more general random matrices by evaluating non-commutative polynomials on random and deterministic matrices. Voiculescu observed in his seminal paper [Voi91] that there is a rule, which he called "freeness", to compute the asymptotic mixed moments of several independent random matrices. Freeness is often refereed to as free independence, as it shares some similarities with the usual probabilistic notion of stochastic independence.

3.2 Asymptotic free independence of random matrices

One nice feature of Haar-distributed unitary random matrices is that they allow us to randomly rotate deterministic matrices. If A_N, B_N are a self-adjoint deterministic matrices, then μ_A and $\mu_{U_N A_N U_N^*}$ coincide. This means that the eigenvalues of A_N remain unchanged. However, the random rotation $U_N A_N U_N^*$ puts the eigenspaces of A_N in a "generic" position with respect to the eigenspaces of B_N.

Consider, for example, the case when A_N, B_N are diagonal deterministic matrices, with $\mu_{A_N} \to \mu_1$ and $\mu_{B_N} \to \mu_2$ as $N \to \infty$. The choice of A_N, B_N being diagonal puts the eigenspaces of A_N, B_N in very specific positions (namely, they coincide) and hence the eigenvalues of $A_N + B_N$ will be $\lambda_1(A_N) + \lambda_1(B_N), \ldots, \lambda_N(A_N) + \lambda_N(B_N)$. This means that the limiting distribution (if it exists) of the self-adjoint matrix $A_N + B_N$ depends not only on μ_1, μ_2, but also on the specific order in which the eigenvalues appear in each matrix.

3.2.1 Free independence and (non-commutative) joint distributions

Voiculescu found out that, surprisingly, the distribution $\mu_{U_N A_N U_N^* + B_N}$ converges towards a *deterministic* probability measure $\mu_1 \boxplus \mu_2$, which depends only on μ_1 and μ_2. The rule for computing the moments of $\mu_1 \boxplus \mu_2$ is based on the fact that $U_N A_N U_N^*$ and B_N behave asymptotically like free operators.

Definitions 3.2.1. *(1). Let $(\mathcal{A}, \tau)$ be a $*$-probability space and let $\bar{a} := a - \tau(a)1_{\mathcal{A}}$ for any $a \in \mathcal{A}$. The unital $*$-subalgebras $A_1, \ldots, A_k \subset \mathcal{A}$ are free iff, for all $m \geq 1$, and all tuples $a_1, \ldots, a_m \in \mathcal{A}$*

$$\tau(\bar{a_1}\bar{a_2}\cdots\bar{a_m}) = 0, \tag{3.1}$$

whenever $a_i \in A_{j(i)}$, with $j(1) \neq j(2) \neq \cdots \neq j(m)$ (note that it is allowed, for example, that $j(1) = j(3)$).

(2). Subsets $S_1, \ldots, S_k \subset \mathcal{A}$ are free if so are their generated unital $*$-subalgebras $\langle S_1 \rangle, \ldots, \langle S_k \rangle$.

We first observe that the freeness condition (3.1) can be applied recursively and turns into a rule for decomposing mixed moments into individual moments: If a, b are free, we have, for example:

$$\begin{aligned} 0 = \tau(\bar{a}\bar{b}) &= \tau((a-\tau(a))(b-\tau(b))) && (3.2)\\ &= \tau(ab) - 2\tau(a)\tau(b) + \tau(a)\tau(b) && (3.3)\\ &\Leftrightarrow && (3.4)\\ \tau(ab) &= \tau(a)\tau(b), && (3.5) \end{aligned}$$

and similarly (after some cancellations):

$$\begin{aligned} \tau(abba) &= \tau(a^2)\tau(b^2), && (3.6)\\ \tau(abab) &= \tau(a^2)(\tau(b))^2 + \tau(b^2)(\tau(a))^2 - (\tau(a))^2(\tau(b))^2. && (3.7) \end{aligned}$$

We note the non-commutative nature of freeness from the equations (3.6) and (3.7). One should be convinced that it is possible (by inductively solving for $\tau(a_1 \cdots a_m)$ in Eq. (3.1)) to express any mixed moment $\tau(a^{n_1}b^{m_1}\cdots a^{n_k}b^{m_k})$ in terms of individual moments.

Mixed moments of free random variables factorize following very precise combinatorial rules, developed by Speicher. In [Spe94] the freeness condition (3.1) was translated into a combinatorial relation which relies on the notions of free cumulants and non-crossing partitions (see Chapter 4, or [NS06] for a comprehensive exposition).

We recall that the classical, stochastic independence of random variables $X_1, \ldots, X_k$ with compact support (or, more generally, of random variables determined by their moments) is equivalent to the factorization of mixed moments:

$$\mathbb{E}(X_1^{n_1}X_2^{n_2}\cdots X_k^{n_k}) = \mathbb{E}(X_1^{n_1})\mathbb{E}(X_2^{n_2})\cdots\mathbb{E}(X_k^{n_k}).$$

In this sense, free probability is a realm, parallel to classical probability, where the factorization of expectations given by classical independence is replaced by freeness. Many fundamental theorems from classical probability, such as the convergence to the Central Limit or the Law of small numbers can be translated to the free setting.

In particular, by just replacing independence by free independence and working algebraically, the free Central Limit (i.e. the limiting distribution of $S_N = \frac{1}{\sqrt{N}}(\bar{a}_1 + \cdots + \bar{a}_N)$ for free self-adjoint, identically distributed elements $(a_i)_{i\geq 1}$) turns out to be Wigner's semicircle law, whereas the free analog of the law of small numbers is the Marchenko-Pastur distribution (also known as free Poisson), which is the (also universal) limit of singular-value distributions of Wishart matrices.

Voiculescu observed in [Voi91] that large random matrices provide approximations to quite intuitive *realizations* of free random variables. In order to consider several matrix ensembles, we need to organize the mixed moments.

Definition 3.2.2. *For an ordered tuple $a = (a_1, \ldots, a_k)$ of random variables in a $*$-probability space $(\mathcal{A}, \tau)$, we call*

$$\Phi_m^a := \Phi_m = \{(i_1, \ldots, i_m) \mapsto \tau(a_{i_1} \ldots a_{i_m}) : i_1, \ldots, i_m \leq k\}$$

the m-th order mixed moments of $(a_1, \dots, a_k)$. We call $\Phi(a) = \bigcup_{m\geq 0} \Phi^a_m$ the joint distribution of $a = (a_1, \dots, a_k)$. To avoid cumbersome notation, for $a = (b_1, b_1^*, \dots, b_k, b_k^*)$ we usually omit the adjoints and simply talk about the $*$-distribution $\Phi(b)^* := \Phi(a)$ of $b = (b_1, \dots, b_k)$.

If $k \geq 2$, one can very rarely associate a probability measure which encodes all the information from $\Phi(a)$. For this reason, whenever we talk about the distribution of $a = (a_1, \dots, a_k)$, we mean the collection of mixed moments $\Phi(a)$. Two k-tuples $(a_1, \dots, a_k) \in \mathcal{A}_1^k$ and $(b_1, \dots, b_k) \in \mathcal{A}_2^k$ in (possibly different) *-probability spaces $(\mathcal{A}_1, \tau_1)$, $(\mathcal{A}_2, \tau_2)$ have the same distribution iff

$$\tau_1(a_{i_1} \cdots a_{i_m}) = \tau_2(b_{i_1} \cdots b_{i_m}),$$

for all $m \geq 1$, $1 \leq i_1, \dots, i_m \leq k$. We denote this situation by writing $(a_1, \dots, a_k) \sim (b_1, \dots, b_k)$. If $(a_1, a_1^* \dots, a_k, a_k^*) \sim (b_1, b_1^*, \dots, b_k, b_k^*)$ we simply write $(a_1, \dots, a_k) \sim^* (b_1, \dots, b_k)$

Definition 3.2.3. *Let $(\mathcal{A}_N, \tau_N)$, $N \geq 1$, and $(\mathcal{A}, \tau)$ be $*$-probability spaces and let $(a_1^{(N)}, \dots, a_k^{(N)}) \in \mathcal{A}_N^k$, $(a_1, \dots, a_k) \in \mathcal{A}^k$ be such that*

$$\lim_{N\to\infty} \tau_N((a_{i_1}^{(N)}) \cdots (a_{i_m}^{(N)})) = \tau(a_{i_1} \cdots a_{i_m}),$$

for all $m \geq 1$, $1 \leq i_1, \dots, i_m \leq k$ Then we say that $(a_1^{(N)}, \dots, a_k^{(N)})$ converges in distribution to $(a_1, \dots, a_k)$ and we write $(a_1^{(N)}, \dots, a_k^{(N)}) \to (a_1, \dots, a_k)$. We denote the situation $(a_1^{(N)}, (a_1^{(N)})^*, \dots, a_k^{(N)}, (a_k^{(N)})^*) \to (a_1, a_1^* \dots, a_k, a_k^*)$ simply by $(a_1^{(N)}, \dots, a_k^{(N)}) \to^* (a_1, \dots, a_k)$

3.2.2 Asymptotic freeness

We give now a convenient reformulation of Voiculescu's striking generalization of Wigner's Semicircle law [Voi91] (and its stronger versions [Dyk93, Voi98, AGZ10, MS12]) which establishes the asymptotic freeness of large matrix ensembles.

Theorem 3.2.4. *For each $N \geq 1$, let $X_1^{(N)}, \dots, X_p^{(N)}$ and $U_1^{(N)}, \dots, U_q^{(N)}$ be independent Wigner and Haar-distributed unitary matrices. Let $D_1^{(N)}, \dots, D_r^{(N)}$ be deterministic matrices, such that, for any $k \geq 1$ and $1 \leq j_1, \dots, j_k \leq r$ there exist a constant $c(j_1, \dots, j_k) \in \mathbb{C}$ such that*

$$\lim_{N\to\infty} \frac{1}{N} \mathrm{Tr}((D_{j_1}^{(N)})(D_{j_2}^{(N)}) \cdots (D_{j_k}^{(N)})) = c(j_1, \dots, j_k). \tag{3.8}$$

Then, as $N \to \infty$,

$$(X_1^{(N)}, \dots, X_p^{(N)}, U_1^{(N)}, \dots, U_q^{(N)}, D_1^{(N)}, \dots, D_r^{(N)}) \to (s_1, \dots, s_p, u_1, \dots u_q, d_1, \dots d_r)$$

where $s_1, \dots, s_p, u_1, \dots u_q, d_1, \dots d_r$ are elements in some $*$-probability space $(\mathcal{A}, \tau)$ whose joint-distribution is determined by the following conditions:

- *s_i is a semicircular for all $i \leq p$.*
- *u_i is a Haar-unitary for all $i \leq q$.*
- *$\tau(d_{i_1} d_{i_2} \cdots d_{i_k}) = c(i_1, \ldots, i_k)$, for any $k \geq 1$, $1 \leq i_1, \ldots, i_k \leq r$.*
- *The algebras $\langle s_1 \rangle, \ldots, \langle s_p \rangle, \langle u_1, u_1^* \rangle, \ldots, \langle u_q, u_q^* \rangle, \langle d_1, \ldots, d_r \rangle$ are free.*

3.2.3 Ideas of the proofs

Theorem 3.2.4 generalizes Wigner's semicircle law in several directions. It allows us to compute the asymptotic mixed moments of

$$(X_1^{(N)}, \ldots, X_p^{(N)}, U_1^{(N)}, \ldots, U_q^{(N)}, D_1^{(N)}, \ldots, D_r^{(N)})$$

by means of the rule of free independence (3.1), in terms of the individual asymptotic moments of $X_1^{(N)}, \ldots, X_p^{(N)}, U_1^{(N)}, \ldots, U_q^{(N)}$ and the (given) asymptotic mixed moments of $(D_1^{(N)}, \ldots, D_r^{(N)})$. A nice way to understand how these mixed moments are calculated is in terms of free cumulants (see Chapter 4).

Our combinatorial proof of the Gaussian case in Chapter 2 will be our main reference to indicate how the different generalizations work.

- Relaxing Gaussian condition.

As we did in Chapter 2 for the Gaussian case, we may consider first non-self-adjoint Wigner matrices Z_i such that $\sqrt{2}X_i = Z_i + Z_i^*$.

For a single matrix $Z := Z_i$ we need to study again

$$\frac{1}{N}\mathbb{E} \circ \mathrm{Tr}(Z^{\varepsilon_1}, \ldots, Z^{\varepsilon_k}) = \frac{1}{N^{1+k/2}} \sum_{\substack{i_1, \ldots, i_k \leq N \\ \pi \in \mathcal{P}(k)}} \kappa_\pi(z_{i_1 i_2}^{\varepsilon_1}, z_{i_2 i_3}^{\varepsilon_2}, \ldots, z_{i_k i_1}^{\varepsilon_k}). \tag{3.9}$$

For the Gaussian case all cumulants of order different than 2 vanished and hence our sum ran over $\mathcal{P}_2(k)$. Then we noticed that only $\pi \in \mathcal{NC}_2(k)$ actually matter in the limit.

If Z is no longer Gaussian, we need to consider partitions $\pi \in \mathcal{P}(k)$ such that all blocks of π are of size greater or equal to 2, but $\pi \in \mathcal{P}(k)$ needs not in principle to be in $\mathcal{P}_2(k)$. Blocks of size 1 are still not allowed because the entries of Z are all centered. The important observation is that the number of different cumulants to be considered depends only on the size k of the moment in question (and does not grow with N, which only affects the number of choices for each free index).

If we allow π to have bigger blocks, it is intuitive that the contribution of π will vanish in the limit since it will imply that more indices need to be identified. This can be shown by induction.

- Several Wigner Matrices.

Going from one Wigner matrix to several is not hard. We now need to consider the more general expression, for $j_1, \ldots, j_k \leq p$

$$\frac{1}{N}\mathbb{E} \circ \mathrm{Tr}(Z_{j_1}^{\varepsilon_1}, \ldots, Z_{j_k}^{\varepsilon_k}) = \frac{1}{N^{1+k/2}} \sum_{\substack{i_1,\ldots,i_k \leq N \\ \pi \in \mathcal{P}(k)}} \kappa_\pi(z_{i_1 i_2}^{\varepsilon_1 ; j_1}, z_{i_2 i_3}^{\varepsilon_2 ; j_2}, \ldots, z_{i_k i_1}^{\varepsilon_k ; j_k})).$$

Where now the blocks of $\pi \in \mathcal{P}(n)$ should also respect the labels indicated by $j = (j_1, \ldots, j_k)$. Such restriction can be then carried out to the very end. In the limit, the sum will again run over non-crossing pairings $\pi \in \mathcal{NC}_\varepsilon(n) \subset \mathcal{NC}_2(n)$ with the additional condition that for all $(a, b) \in \pi$, $j_a = j_b$. We will see later in Chapter 4 that this characterizes a free family of semicircular operators.

- Wigner matrices and deterministic matrices.

If we now allow deterministic matrices to operate between our Wigner matrices we need to compute expressions of the form

$$\frac{1}{N}\mathbb{E} \circ \mathrm{Tr}(D_{j_0} Z_{j_1}^{\varepsilon_1} D_{j_2}, \ldots, Z_{j_{2k-1}}^{\varepsilon_k} D_{j_{2k}}).$$

More generally, we are interested in estimating

$$\frac{1}{N^{1+k/2}} \sum_{\substack{i_0,\ldots,i_{2k+1} \leq N \\ \pi \in \mathcal{P}(2k+1)}} \kappa_\pi(d_{i_0 i_1}^{(j_0)}, z_{i_1 i_2}^{\varepsilon_1}, d_{i_2 i_3}^{(j_2)}, z_{i_3 i_4}^{\varepsilon_2}, \ldots, z_{i_{2k-1} i_{2k}}^{\varepsilon_k}, d_{i_{2k} i_{2k+1}}^{(j_{2k})}).$$

Since the d_{ij}'s are constants, $\{\{1\}, \{3\}, \ldots, \{2k+1\}\} \subset \pi$, otherwise the cumulant vanishes. Therefore, we need only to consider partitions of $\mathcal{P}(2, 4, \ldots, 2k) \cong \mathcal{P}(k)$.

As an example, let us assume that we only have a single Gaussian matrix and consider the pairing $\{\{1,3\}\{2,7\}\{4,5\}\{6,8\}\}$ and $\varepsilon = (1, *, *, 1, *, 1, 1, *)$. Again, each block corresponds to a cumulant which imposses identifications of some indices:

$h_0 := i_0, \quad h_1 := i_1 = i_6, \quad h_2 := i_2 = i_5, \quad h_3 := i_4 = i_{13}, \quad h_4 := i_3 = i_{14},$
$h_5 := i_7 = i_{10}, \quad h_6 := i_8 = i_9, \quad h_7 := i_{11} = i_{16}, \quad h_8 := i_{12} = i_{15}, \quad h_9 := i_{17}$

The contribution of π will be $\kappa_\pi(z, \bar{z}, \bar{z}, z, \bar{z}, z, z, \bar{z})$ (which is one in this case, and does not depend on N in general), times $\mathrm{Tr}(D_{\pi,\varepsilon})$, where

$$D_{\pi,\varepsilon} = \sum_{h_0,\ldots,h_9 \leq N} d_{i_0 i_1}^{(j_0)} d_{i_2 i_3}^{(j_2)} \ldots d_{i_{2k} i_{2k+1}}^{(j_{2k})}.$$

For our example the sum $D_{\pi,\varepsilon}$ is given by:

$$\begin{aligned}
D_{\pi,\varepsilon} &= \sum_{h_0,\ldots,h_9 \leq N} d_{i_0 i_1}^{(j_0)} d_{i_2 i_3}^{(j_2)} d_{i_4 i_5}^{(j_4)} d_{i_6 i_7}^{(j_6)} d_{i_8 i_9}^{(j_8)} d_{i_{10} i_{11}}^{(j_{10})} d_{i_{12} i_{13}}^{(j_{12})} d_{i_{14} i_{15}}^{(j_{14})} d_{i_{16} i_{17}}^{(j_{16})} && (3.10) \\
&= \sum_{h_0,\ldots,h_9 \leq N} d_{h_0 h_1}^{(j_0)} d_{h_2 h_4}^{(j_2)} d_{h_3 h_2}^{(j_4)} d_{h_1 h_5}^{(j_6)} d_{h_6 h_6}^{(j_8)} d_{h_5 h_7}^{(j_{10})} d_{h_8 h_3}^{(j_{12})} d_{h_4 h_8}^{(j_{14})} d_{h_7 h_9}^{(j_{16})} && (3.11) \\
&= \sum_{h_0,\ldots,h_9 \leq N} (d_{h_0 h_1}^{(j_0)} d_{h_1 h_5}^{(j_6)} d_{h_5 h_7}^{(j_{10})} d_{h_7 h_9}^{(j_{16})}) (d_{h_3 h_2}^{(j_4)} d_{h_2 h_4}^{(j_2)} d_{h_4 h_8}^{(j_{14})} d_{h_8 h_3}^{(j_{12})}) (d_{h_6 h_6}^{(j_8)}) && (3.12) \\
&= D_{j_0} D_{j_6} D_{j_{10}} D_{j_{16}} \mathrm{Tr}(D_{j_4} D_{j_2} D_{j_{14}} D_{j_{12}}) \mathrm{Tr}(D_{j_8}), && (3.13)
\end{aligned}$$

Hence

$$\begin{aligned}\frac{1}{N^{1+k/2}}\mathrm{Tr}(D_{\pi,\varepsilon}) &= \frac{1}{N^5}\mathrm{Tr}(D_{j_0}D_{j_6}D_{j_{10}}D_{j_{16}})\mathrm{Tr}(D_{j_4}D_{j_2}D_{j_{14}}D_{j_{12}})\mathrm{Tr}(D_{j_8}) \\ &= \frac{1}{N^2}[\tau_N(D_{j_0}D_{j_6}D_{j_{10}}D_{j_{16}})\tau_N(D_{j_4}D_{j_2}D_{j_{14}}D_{j_{12}})\tau_N(D_{j_8})] \\ &\to \frac{1}{N^2}c(j_0,j_6,j_{10},j_{16})c(j_4,j_2,j_{14},j_{12})c(j_8) \\ &\to 0.\end{aligned}$$

The Gaussian case here is notably easier than the general Wigner case. Pair partitions are quite convenient because they identify pairs of indices, which allows us to order the d_{ij} in cycles as we did above. Then one can show inductively that only non-crossing pairings matter in the limit. For a non-crossing pairing, the map $\tau_N(D_{\pi,\varepsilon})$ turns out to be equal to the multiplicative extension of $\frac{1}{N}\mathrm{Tr}_{Kr(\pi)}(D_{j_0}, D_{j_2}, \ldots, D_{j_{2k}})$, where $Kr(\pi)$ is the Kreweras complement of π (which also separates the deterministic matrices according to parity, see Section 4).

For the general, non-Gaussian case, where cumulants of order greater than 2 are allowed, more than two indices may be identified and such cyclic reordering of the d_{ij} is not possible in general. One must show again that only non-crossing pair partitions yield a non-vanishing contribution.

One may associate a graph $G_\pi = (V, E)$ to each partition $\pi \in \mathcal{P}(k)$, where we put one vertex for each equivalence class of indices and an edge joining the classes corresponding to the vertices i_{2m}, i_{2m+1}. Equivalently, $G_\pi = (V, E)$ can be seen as a quotient graph: we start with the vertices $i_0, i_1, \ldots, i_{2k+1}$ and the edges $\{(i_0, i_1), (i_2, i_3), \ldots, (i_{2k}, i_{2k+1})\}$ and then we perform the identifications of the vertices indicated by the cumulant

$$\kappa_\pi(z^{\varepsilon_1}_{i_1 i_2}, \ldots, z^{\varepsilon_k}_{i_{2k-1} i_{2k}}).$$

Mingo and Speicher [MS12] bounded the order such products of matrix entries in terms of the operator norms of the matrices and the forest of two-edge connected components of the graph G_π. These sharp estimates allowed not only to prove that non-crossing pair partitions are the only non-vanishing contribution, but also stronger forms convergence which depend on more delicate conditions of the matrices.

Alternatively, one may use concentration of measure [AGZ10].

- Haar Matrices

We will briefly address the combinatorics behind the asymptotic freeness of Haar-unitary random matrices and deterministic matrices.

The joint distribution of the entries of a Haar-distributed unitary matrix $U_N = (u_{ij})_{i,j\leq N}$ is quite complicated. The entries u_{ij} are known to fulfill the Wick formula

$$\mathbb{E}(u_{i_1 j_1} \ldots u_{i_q j_q} \bar{u}_{i'_1 j'_1} \ldots \bar{u}_{i'_q j'_q}) = \sum_{\rho,\sigma \in S_q} \delta_{i_1 i'_{\rho(1)}} \ldots \delta_{i_q i'_{\rho(q)}} \delta_{j_1 j'_{\sigma(1)}} \ldots \delta_{j_q j'_{\sigma(q)}} Wg(N, \rho\sigma^{-1}),$$

where, for each N, the Weingarten function $Wg_N : S_q \to \mathbb{C}$ is some complicated class function whose values depend on characters of representations of symmetric groups.

The leading term of the Weingarten function can be obtained from the asymptotic expansion

$$Wg(N,\sigma) = N^{-n-|\sigma|}\prod_i (-1)^{l(\alpha_i)-1}C_{(l(\alpha_i)-1)} + O(N^{-n-|\sigma|-2}),$$

where the permutation $\sigma = \alpha_1 \dots \alpha_k \in S_n$ is a product of cycles $\alpha_1, \dots, \alpha_k$ of lengths $l(\alpha_i)$, C_k is the k-th Catalan number and $|\sigma|$ is the minimum number of transpositions required to express σ.

One may perform an asymptotic analysis, similar to the ones that we did for the Wigner case, to conclude hat the asymptotic mixed moments of the matrices are computed according to the rules described in Theorem 3.2.4 (see [Col03], [CS06]).

Remarks 3.2.5. *(1). In [Voi91, Voi98] Voiculescu proved first the asymptotic freeness of Gaussian matrices and deterministic matrices. By observing that independent Haar-distributed unitary matrices can be obtained by considering the unitary part of the polar decompositions of independent non-self-adjoint Gaussian random matrices, Voiculescu transferred the asymptotic freeness of independent non-self-adjoint Gaussian matrices and deterministic matrices to the analogous result with independent Haar-unitary matrices and deterministic matrices. In [Dyk93],[AGZ10] and [MS12] the Gaussian condition was relaxed.*

Alternatively, more direct, combinatorial proofs for both the Gaussian (or Wigner) case and the Haar-unitary case can be done using the Wick type formulas (see [NS06] and [Xu97]). It is clear that one can also combine these proofs to consider Wigner and Haar unitary random matrices simultaneously.

(2). A Non-self-adjoint Wigner matrix Y can be seen as the sum $X_1 + iX_2$ of two independent self-adjoint Wigner matrices. The limiting random variable $c = s_1 + is_2$ is called a circular element (observe that $c^* = s_1 - is_2$ is not self-adjoint). Clearly, they will also be asymptotically free from deterministic Matrices, and other independent Haar-unitary, Wigner and non-self-adjoint Wigner matrices. Our algorithms to compute distributions will be formulated for polynomial expressions on self-adjoint Wigner matrices / semicircular elements, which clearly include non-self-adjoint Wigner matrices / circular elements.

We will do the same with Wishart matrices: If p, c are free, c is a circular element and p is a projection, a Wishart matrix of parameter $\lambda > 0$ can essentially be viewed asymptotically as a scalar multiple of cpc^* or pcc^*p (depending on whether $\lambda \leq 1$ or not).

(3). There exist a stronger version of freeness, called second-order freeness, which allows to control fluctuations of random matrices. Second order freeness is achieved (see [AGZ10] or [MS12]) if we slightly strengthen the assumptions on our matrices, by either asking all Wigner random matrices to be Gaussian matrices, or by asking all deterministic matrices to be diagonal. Under such conditions, the empirical eigenvalue distribution of any polynomial $P(X_1^{(N)}, \dots, X_p^{(N)}, U_1^{(N)}, \dots, U_q^{(N)}, D_1^{(N)}, \dots, D_r^{(N)})$ converges almost surely to the spectral distribution of $P(s_1, \dots, s_p, u_1, \dots u_q, d_1, \dots d_r)$.

(4). In Chapter 6 we will typically be given a single collection of deterministic matrices $D_1, \ldots, D_r$ with some fixed (usually large) size s, along with some independent Haar-unitary and Wigner matrices. In order to meet the condition of Eq. (3.8) we need (sub-)sequences of matrices whose sizes grow arbitrarily large. We may obtain such sequences by considering deterministic matrices $D_i^{(N)} := D_i \otimes I_N$ and simply considering larger Haar Unitaries and Wigner matrices (of size sN). The normalized traces of $(D_1^{(N)}, \ldots, D_r^{(N)})$ do not vary with N and hence

$$c(i_1, \ldots, i_k) = \frac{1}{s}\mathrm{Tr}(D_{i_1} D_{i_2} \cdots D_{i_k}).$$

If we are given a self-adjoint random matrix model P, produced by evaluating a polynomial on Wigner, Haar-unitary matrices and deterministic matrices of size s, we can use the previous trick to blow-up each of the matrices and consider the corresponding model P_N. In the limit $N \to \infty$, we obtain a deterministic model with some spectral distribution μ_{P_∞}. If, for instance, the norm of the original deterministic matrices is kept fixed and s is large, the distribution μ_{P_∞} will be a good approximation of the original matrix model of size s.

One of the main purposes of this thesis is to successfully compute the distribution of μ_{P_∞} (and then provide a similar treatment for polynomials evaluated on matrices of different sizes).

Voiculescu's theorem allows us in principle to compute each moment of μ_{P_∞} in terms of moments of free semi-circulars, Haar-unitaries and deterministic matrices.

Going from the moments to the actual distribution μ_{P_∞} is usually not that easy and hence we need to gain some intuition on the nature of such mixed moments.

In particular, for many Hermitian matrix models, one can very often suppress the Haar-unitaries in the limit, provided that some "freeness" is imposed on some sub-collections of deterministic matrices. A first instance of this phenomenon comes from the study of the asymptotic joint distribution of $U_N A_N U_N^*$ and B_N, which will also serve as motivation for Voiculescu's free convolution.

3.3 Rotated families of deterministic matrices

If sub-algebras $\langle A, B\rangle \subset \mathcal{A}$ are $*$-free from a Haar unitary u, and $(a_1, \ldots, a_p) \in A^p$, $(b_1, \ldots, b_q) \in B^q$, then the joint distribution of $(\tilde{a}_1, \ldots, \tilde{a}_p, b_1, \ldots b_q)$, where $\tilde{a}_i = u a_i u^*$ is completely determined: $\{\tilde{a}_1, \ldots, \tilde{a}_p\}$ and $\{b_1, \ldots b_q\}$ are free and $(\tilde{a}_1, \ldots, \tilde{a}_p) \sim (a_1, \ldots, a_p)$. Roughly speaking, conjugating a family of variables by a free Haar-unitary does not alter the distribution of the family and makes it free from a second family of variables.

The fact that $(\tilde{a}_1, \ldots, \tilde{a}_p) \sim (a_1, \ldots, a_p)$ is trivial by the tracial property of τ. Hence we only need to show that $\langle u a_1 u^*, \ldots, u a_p u^*\rangle$ and $\langle b_1, \ldots, b_q\rangle$ are free.

Let $a^{(1)}, \ldots, a^{(k)} \in \langle a_1, \ldots, a_p\rangle$ and $b^{(1)}, \ldots, b^{(k)} \in \langle b_1, \ldots, b_q\rangle$. We note that $a^{(j)} \in \langle a_1, \ldots, a_p\rangle$ iff $u a^{(j)} u^* \in \langle u a_1 u^*, \ldots, u a_p u^*\rangle$. Since $\tau(u a^{(j)} u^*) = \tau(a^{(j)})$ and $\tau(u) = 0 = \tau(u^*)$, we have that $\overline{uau^*} = \overline{u}\overline{a^{(j)}}\overline{u^*}$.

Hence

$$\tau((\overline{ua^{(1)}u^*})\overline{b^{(1)}}\cdots(\overline{ua^{(k)}u^*})\overline{b^{(k)}}) = \tau(\overline{ua^{(1)}}\,\overline{u^*}\overline{b^{(1)}}\cdots\overline{ua^{(k)}}\,\overline{u^*}\overline{b^{(k)}}), \quad (3.14)$$

$$\tau(\overline{b^{(1)}}(\overline{ua^{(1)}u^*})\cdots\overline{b^{(k)}}(\overline{ua^{(k)}u^*})) = \tau(\overline{b^{(1)}}\,\overline{ua^{(1)}}\,\overline{u^*}\cdots\overline{ua^{(k-1)}}\,\overline{u^*}\overline{b^{(k)}}), \quad (3.15)$$

$$\tau((\overline{ua^{(1)}u^*})\overline{b^{(1)}}\cdots\overline{b^{(k-1)}}(\overline{ua^{(k)}u^*})) = \tau(\overline{ua^{(1)}}\,\overline{u^*}\overline{b^{(1)}}\cdots\overline{b^{(k-1)}}\,\overline{ua^{(k)}}\,\overline{u^*}), \quad (3.16)$$

$$\tau(\overline{b^{(1)}}(\overline{ua^{(1)}u^*})\cdots(\overline{ua^{(k-1)}u^*})\overline{b^{(k)}}) = \tau(\overline{b^{(1)}}\,\overline{ua^{(1)}}\,\overline{u^*}\cdots\overline{ua^{(k-1)}}\,\overline{u^*}\overline{b^{(k)}}). \quad (3.17)$$

By freeness of $\langle a_1,\ldots,a_p,b_1,\ldots,b_q\rangle$ and $\{u,u^*\}$ all the RHS expressions vanish and the freeness of $\langle ua_1u^*,\ldots,ua_pu^*\rangle$ and $\langle b_1,\ldots,b_q\rangle$ is established.

The same statement (with the same proof) holds for several randomly rotated collections:

Proposition 3.3.1. *Let $A_0,\ldots,A_k\subseteq\mathcal{A}$ be $*$-subalgebras of a $*$-probability space $(\mathcal{A},\tau)$ and let $u_1,\ldots,u_k\in\mathcal{A}$ be Haar-unitary elements, such that $\langle A_0,\ldots,A_k\rangle,\{u_1\},\ldots,\{u_k\}$ are $*$-free. For $0\le j\le k$, let $(a_1^{(j)},a_2^{(j)},\ldots,a_{p(j)}^{(j)})\in A_j^{p(j)}$. Then*

$$(a_1^{(0)},\ldots,a_{p(0)}^{(0)},u_1a_1^{(1)}u_1^*,\ldots,u_1a_{p(1)}^{(1)}u_1^*,\ldots,u_ka_{p(k)}^{(k)}u_k^*) \quad (3.18)$$

$$\sim\ (a_1^{(0)},\ldots,a_{p(0)}^{(0)},\tilde{a}_1^{(1)},\ldots,\tilde{a}_{p(1)}^{(1)},\ldots,\tilde{a}_{p(k)}^{(k)}), \quad (3.19)$$

where $\langle a_1^{(0)},\ldots,a_{p(0)}^{(0)}\rangle,\langle\tilde{a}_1^{(1)},\ldots,\tilde{a}_{p(1)}^{(1)}\rangle,\ldots,\langle\tilde{a}_1^{(k)},\ldots\tilde{a}_{p(k)}^{(k)}\rangle$ are free and $(\tilde{a}_1^{(j)},\ldots\tilde{a}_{p(j)}^{(j)})\sim(a_1^{(j)},\ldots a_{p(j)}^{(j)})$, for $j\le k$.

Theorem 3.2.4 states that, in particular, if $(A_N,B_N)\to(a,b)$, then $(A_N,B_N,U_N)\to(a,b,u)$, where $\{a,b\}$,$\{u\}$ are free. By the previous proposition $(U_NA_NU_N^*,B_N)\to(\tilde{a},b)$, where $\tilde{a},b$ are free and $\tilde{a}\sim a$.

If a,b are self-adjoint, so is $\tilde{a}+b$. By freeness, the distribution $\mu_{\tilde{a}+b}$ of $\tilde{a}+b$ depends only on the distribution $\mu_{\tilde{a}}(=\mu_a)$ of $\tilde{a}$ and the distribution μ_b of b. We write $\mu_a\boxplus\mu_b:=\mu_{\tilde{a}+b}$, the free additive convolution.

Analogously, the distribution of the self-adjoint element $b\tilde{a}b$, or, equivalently, the asymptotic distribution of $B_NU_NA_NU_N^*B_N$, depends again only on the distribution μ_a and μ_b. In fact, by the tracial property of τ, one can easily see that only the even moments of b appear in the computation. We write $\mu_a\boxtimes\mu_{b^2}:=\mu_{b\tilde{a}b}$ to denote the free multiplicative convolution.

The operation $(a_1,\ldots,a_k)\to(ua_1u^*,a_2\ldots,a_k)$ of conjugating a variable by a free Haar unitary simplifies the distribution of $(a_1,a_2,\ldots,a_k)$ by removing the correlations between a_1 and $(a_2,\ldots,a_k)$.

We will observe in Section 7.2 that the reason why the deterministic matrices R_i's and T_i of our models are not mixed in the fixed point equations is exactly because of this separation phenomenon.

The same separation happens if we consider the conjugation ca_1c^* with a free circular or semicircular random variable (but this time the conjugation ca_1c^* does change the distribution by its compound Poisson).

3.4 The Cauchy-Stieltjes transform

The basic analytical tool in non-commutative probability is the Cauchy-Stieltjes transform. For any Borel probability measure μ on $\mathbb{R}$, the Cauchy-Stieltjes transform is an analytic function

$$G_\mu(z) = \int_{\mathbb{R}} \frac{d\mu(t)}{z-t}$$

which maps the upper complex half-plane $\mathbb{C}^+$ to the lower half-plane $\mathbb{C}^-$. The probability measure can be recovered from its Cauchy-transform via the Stieltjes' inversion formula:

$$\mu((t_0,t_1]) = -\frac{1}{\pi}\lim_{\epsilon\to 0^+}\lim_{\delta\to 0}\int_{t_0+\delta}^{t_1+\delta}\Im(G_\mu(t+i\epsilon))dt.$$

3.4.1 The R and S transforms and the analytic subordination phenomena

Voiculescu [Voi86, Voi87] introduced some analytical tools to compute such free convolutions.

The reciprocal $F_\mu(z) = (G_\mu(z))^{-1}$ of the Cauchy-transform is an analytic self-map on $\mathbb{C}^+$. It was show in [BV93] that on certain Stolz domains

$$\Gamma_{\alpha,\beta} := \{z = x+iy : y>0, x<\alpha y, |z|>\beta\}$$

the map F_μ is injective and hence there exist a right inverse $F_\mu(F_\mu^{-1}(z)) = z$. We consider the analytic function $\phi_\mu(z) := F_\mu^{-1}(z) - z$.

Theorem 3.4.1. *[Voi86, BV93] Let μ,ν be probability measures on $\mathbb{R}$. For some α,β there exist a unique probability measure $\mu\boxplus\nu$ on $\mathbb{R}$ such that*

$$\phi_{\mu\boxplus\nu}(z) = \phi_\mu(z) + \phi_\nu(z), \quad z\in\Gamma_{\alpha,\beta}.$$

If both μ,ν are determined by their moments, the moments of $\mu\boxplus\nu$ are exactly the moments of $a+b$, where a,b are free and $\mu_a=\mu$, $\mu_b=\nu$.

Similarly, the analytic function $M_\mu(z) = z^{-1}G_\mu(z^{-1}) - 1$ is invertible in some domain of the form $\{z : |z|<\varepsilon\}\cap\mathbb{C}^+$. Let $S_\mu(z) := \frac{1+z}{z}M_\mu^{-1}(z)$.

Theorem 3.4.2. *[Voi87, BV93] Let μ,ν be probability measures on $\mathbb{R}^+$. For some $\varepsilon>0$ small enough, there exist a unique probability measure $\mu\boxtimes\nu$ on $\mathbb{R}^+$ such that*

$$S_{\mu\boxtimes\nu}(z) = S_\mu(z)S_\nu(z), \quad z\in\mathbb{C}^+, |z|<\varepsilon.$$

Once more, if both μ,ν are determined by their moments, the moments of $\mu\boxtimes\nu$ are exactly the moments of ab, where a,b are free and $\mu_a=\mu$, $\mu_b=\nu$. The element ab is not self-adjoint. If $\mathcal{A}$ is a C^*-algebra, any positive element $a>0$ has a unique positive square root $\sqrt{a}$ and the element ab has the same moments as the self-adjoint element $\sqrt{a}b\sqrt{a}$.

Explicit computations of additive and multiplicative convolutions are not possible in general, as both processes imply inverting analytic functions and this can only be done

explicitly for a small class of distributions. In the last years, an alternative approach by analytic subordination (which was developed in different contexts by Voiculescu and Biane [Bia98, Voi00, Voi02]).

Belinschi and Bercovici [BB07] introduced an approach that allows one to find the subordination functions with fixed point equations.

Theorem 3.4.3. *[BB07] Let μ, ν be probability measures on $\mathbb{R}$. There exist an analytic map $\omega : \mathbb{C}^+ \to \mathbb{C}^+$ such that $G_\mu(\omega(z)) = G_{\mu \boxplus \nu}(z)$. Furthermore, for any $z \in \mathbb{C}^+$ the subordination function $\omega(z)$ satisfies*

$$\omega(z) = \lim_{n \to \infty} f_z^{\circ n}(\lambda),$$

where, for any $z, \lambda \in \mathbb{C}^+$, $f_z(\lambda) = h_\nu(h_\mu(\lambda) + z) + z$ and h is the auxiliary analytic self-map $h_\mu(z) = G_\mu(z)^{-1} - z$ on $\mathbb{C}^+$.

One of our main tools for the computation of asymptotic distributions is the generalization [BMS13] of the previous theorem to the operator-valued level. Before going operator-valued, we want to point out some typical problems that we run into while dealing with asymptotic distributions of random matrices.

3.5 Limitations of scalar-valued free probability

The combination of the previous tools would let us compute the asymptotic distribution of some polynomials. For example, if a, b, c, d are free and self-adjoint, the distribution μ of the self-adjoint element $aba + cdc$ will be given by $\mu = (\mu_b \boxtimes \mu_{a^2}) \boxplus (\mu_d \boxtimes \mu_{c^2})$.

In this thesis we want to deal with arbitrary polynomials evaluated on Wigner, Haar-distributed unitary matrices and deterministic matrices, possibly with different sizes. Several problems arise when we consider such general polynomials:

1. Already for quite simple polynomials, such as $p = a_1 b_1 a_1 + a_2 b_2 a_2$, with $\{a_1, a_2\}$, $\{b_1, b_2\}$ free, we have that, in general, $a_1 b_1 a_1$ and $a_2 b_2 a_2$ are not free. We would be able to compute $\mu_{a_1 b_1 a_1}$ and $\mu_{a_2 b_2 a_2}$ as free multiplicative convolutions, but the distribution of p would still depend on the joint distribution of (a_1, a_2) and the joint distribution of (b_1, b_2) in a non-trivial way.

2. One can think of Wigner matrices of different sizes by simply considering large Wigner matrices of the same size and compressing them with projections. Unfortunately, the compressions of Haar-unitary Matrices are not Haar-unitary matrices of smaller size. If one considers independent Haar-unitary matrices of different sizes $U_1^{(N)} \in \mathcal{U}(N_1(N))$, $\dots$, $U_k^{(N)} \in \mathcal{U}(N_k(N))$, $N_1(N) \leq N_2(N) \leq \cdots \leq N_k(N)$ (all embedded in the upper-left corner of $M_{N_j(N)}(\mathbb{C})$), such that $N_k(N)/N_k \to \rho_j \in (0, 1]$, then $(U_1^{(N)}, \dots, U_k^{(N)})$ are not asymptotically free, unless $\rho_2 = \rho_j$. Of course this also means that randomly rotated matrices of different sizes are not free in general.

We will need to consider more general notions of freeness in order to overcome these problems. We will address the operator-valued situation in Chapter 5.

In the next chapter we present the basic topics on the combinatorics of free probability, which is based on the theory of Möbius inversion on the lattice of non-crossing partitions.

Although our main results and algorithms for eigenvalue distributions of random matrices can be understood independently from the combinatorics, a lot of intuition on the computation of first order moments can be obtained from understanding these non-crossing partitions.

We also present some recent results on k-divisible partitions and cumulant-to-cumulant formulas.

4 Combinatorics of free and non-commutative probability

An intriguing aspect of non-commutative probability is the existence of several notions of independence [Voi85, SW97, Mur01] with corresponding cumulants introduced in [Voi85, SW97, HS11b, Leh04] sharing many common features. In a certain sense (which can be made precise, see [Spe97, BGS02, Mur02]) these are the only "natural" notions of independence and the combinatorics of cumulants in particular show very close analogies between the different theories.

Cumulants provide a combinatorial description of independence of random variables. While Fourier analysis is the tool of choice for most problems in classical probability, cumulants are an indispensable ingredient for many investigations in non-commutative probability.

Let $(m_n)_{n\geq 1}$ be a (moment) sequence with $m_0 = 1$ and $F(z) = \sum_{n=0}^{\infty} \frac{m_n}{n!} z^n$ its exponential generating function and $M(z) = \sum_{n=0}^{\infty} m_n z^n$ the ordinary generating function.

1. The exponential generating function of the classical cumulants $(\kappa_n)_{n\geq 1}$ satisfies the identity
$$\sum_{n=1}^{\infty} \frac{\kappa_n}{n!} z^n = \log F(z).$$

2. The ordinary generating function of the free cumulants $(r_n)_{n\geq 1}$
$$R(z) = \sum_{n=1}^{\infty} r_n z^{n-1}$$
is called *R-transform* and satisfies the equivalent identities
$$1 + R(zM(z)) = M(z), \tag{4.1}$$
$$M(z/(1 + R(z))) = 1 + R(z). \tag{4.2}$$

3. The ordinary generating function of the Boolean cumulants $(b_n)_{n\geq 1}$
$$B(z) = \sum_{n=1}^{\infty} b_n z^n$$
satisfies the identity
$$M(z) = \frac{1}{1 - B(z)}. \tag{4.3}$$

The work by Bercovici and Pata [BP99] was the first instance where explicit bijections between large classes of random variables were found. One way of understanding such bijections is through cumulants.

The classical cumulants are the (scaled) coefficients of the logarithm of the Fourier transform of a probability measure. The n-th-cumulant of a measure is a polynomial in the first n moments. As we saw in Chapter 2, the precise relation between moments and cumulants can be understood via the lattices of set partitions $\mathcal{P}$.

The lattice $\mathcal{P}$ is in some sense too general. In practice, going over to the cumulants very often even seems as a detour to compute the moments of independent random variables, as these can be directly computed by simply factoring the expectation $\mathbb{E}$.

The free situation is very different. Speicher [Spe94] observed that the coefficients of Voiculescu's R-transform are obtained from the moments by performing a Möbius inversion on the sub-lattice $\mathcal{NC} \subset \mathcal{P}$ of non-crossing partitions (see section 4.1 for all definitions concerning partitions and cumulants). These partitions have much nicer combinatorial properties which yield unexpected coincidences between the free analogs of classical distributions.

Two of the most important distributions in classical probability are the Gaussian and the Poisson distributions, which are also known as the central limit and the law of small numbers. Just as the normal distribution, the Poisson random variable X also has nice cumulants, namely $\kappa_n(X) = 1$ for all $n \geq 1$.

The free counterparts of these classical distributions are respectively, Wigner's semicircle law and the Marchenko-Pastur distribution. Both distributions are universal in random matrix theory and satisfy that their free cumulants coincide with the classical cumulants of their classical counterparts. Since the number of non-crossing partitions of n elements $|\mathcal{NC}(n)|$ equals the number of non-crossing parings $|\mathcal{NC}_2(2n)|$ of $2n$ elements, the free Poisson happens also to be the square of a semicircular random variable. The square of the Gaussian distribution is the chi-squared distribution and hence the classical analog of this result does not hold.

Higher order moments of the i.i.d. distributions which we use to converge to the central limit are quite important when considering more quantitative estimates about their convergence. For example, the third and fourth moments appear in the estimates of Berry-Essen ([Ess56]) type theorems.

Convergence of the fourth moment of normalized distributions coming from a fixed chaos with respect to both Wiener integrals [NP05] and "Wigner" integrals [KNPS12] is enough to guarantee the convergence in distribution to the corresponding Gaussian distribution.

Examples of non-commutative random variables which have a particularly nice combinatorial behavior are the k-divisible elements, which have been studied by Arizmendi [Ari12]. The combinatorics of these elements are now governed by the posets of k-divisible and k-equal partitions (which are generalizations of $\mathcal{NC}$ and $\mathcal{NC}_2$), introduced by Edelman [Ede80]. Canonical examples of these variables are even elements (such as all standard self-adjoint and non-self-adjoint Gaussian distributions) and unitaries (such as uniform distributions at k-th roots of 1).

In [AV12] we observed that, surprisingly, the moments and cumulants of a product of

k free (and Boolean independent) random variables can be very nicely expressed in terms of Kreweras complements of k-divisible partitions. Our result is based on an equivalent, less symmetric formula for the case $k = 2$ [NS06].

Since the works of Belinschi and Nica [BN08b] several connections between the Boolean, free (and monotone) worlds have been found (see, for example [AH13]). Although the Boolean independence has still not found important realizations by itself (as with free random variables through large random matrices), its relation with other probabilities may be useful to transfer results from the Boolean world (where they may be simpler) to the other scenarios.

Examples of this are Arizmendi's fourth moment theorem for infinitely divisible measures [Ari13, AJ14] (á la Nualart-Pecatti) where the Boolean case essentially reduces to convergence to a constant random-variable. In [AV14] we obtain a uniform proof for the convergence in norm to the Boolean, monotone and free central limits (the free case was first proved by Bercovici and Voiculescu [BV95]). Since the classical central limit is unbounded, no classical analogue of this result exists. In [AV12] better bounds for the norm of *free* multiplicative convolutions are obtained using the *Boolean* cumulants.

In the spirit of these results and ideas, it seems important to be able to read how the different cumulants encode the information of a probability measure. The first formulas relating different types of cumulants were found by Lehner [Leh02] (see also [BN08a] and [JV13]).

In [AHLV14] we obtained the remaining cumulant-to-cumulant formulas, which will be presented in section 4.2. The proofs rely not only on partitions but also on a variety of combinatorial structures (colored partitions, nesting trees, crossing graphs, heaps, pyramids, Tutte polynomials) which are not immediately related to the Wick calculus on random matrices. For this reason, we omit these proofs.

We rather present full proofs of our results from [AV12] which we present in Section 4.3. The combinatorial objects that we use there (free cumulants, Kreweras complements, and in particular, Krawczyk and Speicher's [KS00] formula for cumulants of products) are quite useful in large random matrix theory.

4.1 Main definitions

Concepts on partitions and ordered partitions are summarized below, first let us recall some well known facts from the theory of *posets* (partially ordered sets). For details on the latter the standard reference is [Sta12].

Proposition 4.1.1 (Principle of Möbius inversion)**.** *On any poset $(P, \leq)$ there is a unique* Möbius function $\mu : P \times P \to \mathbb{Z}$ *such that for any pair of functions $f, g : P \to \mathbb{C}$ (in fact any abelian group in place of $\mathbb{C}$) the identity*

$$f(x) = \sum_{y \leq x} g(y) \tag{4.4}$$

holds for every $x \in P$ if and only if

$$g(x) = \sum_{y \leq x} f(y)\, \mu(y, x) \tag{4.5}$$

holds for every $x \in P$. In particular, if, given f, two functions g_1 and g_2 satisfy (4.4), then g_1 and g_2 coincide.

Definition 4.1.2. *1. A* partition *of a set is a decomposition into disjoint subsets, called* blocks. *The set of partitions of the set $[n] := \{1, \ldots, n\}$ is denoted by $\mathcal{P}(n)$. It is a lattice under refinement order with maximal element $\{[n]\}$ denoted by $\hat{1}_n$ and minimal element $\{\{1\}, \ldots, \{n\}\}$ denoted by $\hat{0}_n$.*

We write $\mathcal{P} = \bigcup_{n \geq 1} \mathcal{P}(n)$ (and similar notations will be used, such as $\mathcal{NC}$).

2. *Any partition defines an equivalence relation on $[n]$ and vice versa. Given $\pi \in \mathcal{P}(n)$, $i \sim_\pi j$ holds if and only if there is a block $V \in \pi$ such that $i, j \in V$.*

3. *A partition $\pi \in \mathcal{P}(n)$ is* non-crossing *if there is no quadruple of elements $1 \leq i < j < k < l \leq n$ such that $i \sim_\pi k$, $j \sim_\pi l$ and $i \not\sim_\pi j$. The non-crossing partitions of order n form a sub-poset which we denote by $\mathcal{NC}(n)$.*

4. *For two blocks V, W of a partition, we say V is an* inner block *of W or equivalently W is an* outer block *of V or V nests inside W if there are $i, j \in W$ such that $i < k < j$ for each $k \in V$.*

5. *A block V of a partition is called an* interval block *if V is of the form $V = \{k, k+1, \ldots, k+l\}$ for $k \geq 1$ and $0 \leq l \leq n - k$.*

6. *An* interval partition *is a partition π for which every block is an interval. The set of interval partitions of $[n]$ is denoted by $\mathcal{I}(n)$ and is a sub-lattice of $\mathcal{P}(n)$. Sometimes these are called* linear partitions *and in fact they are in obvious bijection with* compositions *of a number n, i.e., sequences of integers $(k_1, k_2, \ldots, k_r)$ such that $k_i > 0$ and $k_1 + k_2 + \cdots + k_r = n$.*

7. *The* non-crossing closure *$\bar{\pi}$ of a partition π is the smallest non-crossing partition which dominates π.*

8. *A partition π is* connected *if its non-crossing closure is equal to the maximal partition $\hat{1}_n$, or, equivalently, the diagram of π is a connected graph. The set of connected partitions is denoted by $\mathcal{P}_{\mathrm{conn}}(n)$.*

9. *The* connected components *of a partition π are the connected sub-partitions of π, i.e., the partitions induced on the blocks of the non-crossing closure $\bar{\pi}$.*

10. *The* interval closure *$\hat{\pi}$ of a partition π is the smallest interval partition which dominates π.*

11. *A partition $\pi \in \mathcal{P}(n)$ is* irreducible *if its interval closure is equal to the maximal partition $\hat{1}_n$. For a non-crossing partition of $[n]$ this is equivalent to the property that $1 \sim_\pi n$. Every partition π can be "factored" into irreducible factors which we denote by $\pi = \pi_1 \cup \cdots \cup \pi_r$. The factors π_j are sub-partitions induced on the blocks of the interval closure $\hat{\pi}$.*

    The sets of irreducible partitions and irreducible non-crossing partitions are respectively denoted by $\mathcal{P}_{\mathrm{irr}}(n)$ and $\mathcal{NC}_{\mathrm{irr}}(n)$.

Different types of partitions are shown in the following figure.

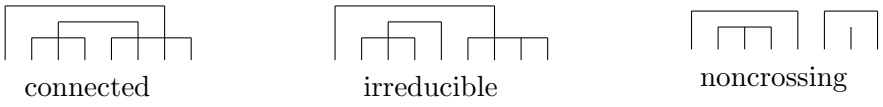

Fig. 4.1: Typical partitions

Definition 4.1.3. *1. An* ordered partition *is a pair (π, λ) of a set partition π and a linear order λ on its blocks. An ordered partition can be regarded as a sequence of blocks: $(\pi, \lambda) = (V_1, \ldots, V_k)$ by understanding that $V_i <_\lambda V_j$ iff $i < j$.*

2. A monotone partition *is an ordered partition (π, λ) with $\pi \in \mathcal{NC}(n)$ such that, for $V, W \in \pi$, $V >_\lambda W$ whenever V is an inner block of W.*

3. An ordered partition (π, λ) is irreducible *if π is irreducible. Let $\mathcal{M}_{irr}(n)$ denote the set of irreducible monotone partitions.*

Our treatment is purely algebraic. Let $(\mathcal{A}, \varphi)$ be a pair of a unital algebra over $\mathbb{C}$ and a unital linear functional on $\mathcal{A}$, i.e. $\varphi(1_\mathcal{A}) = 1$. We denote by K_n, H_n, B_n, R_n the multivariate classical, monotone, Boolean and free cumulants respectively. The univariate cumulants κ_n, h_n, b_n, r_n are obtained by evaluating the multivariate cumulants at n copies of a single variable.

Let A_n (resp., a_n) be one of the cumulant functionals K_n, B_n, H_n, R_n (resp., κ_n b_n, h_n, r_n). Given a partition $\pi \in \mathcal{P}(n)$ and $X, X_i \in \mathcal{A}$, we define the associated multivariate and univariate *partitioned cumulant functionals*

$$A_\pi(X_1, \ldots, X_n) := \prod_{V \in \pi} A_{|V|}(X_V), \qquad a_\pi(X) := A_\pi(X, \ldots, X) = \prod_{V \in \pi} a_{|V|}(X),$$

where we use the notation

$$A_{|V|}(X_V) := A_m(X_{v_1}, \ldots, X_{v_m})$$

for a block $V = \{v_1, \ldots, v_m\}$, $v_1 < \cdots < v_m$. The linear functional φ gives rise to the multi-linear functional

$$(X_1, \ldots, X_n) \mapsto \varphi(X_1 \cdots X_n)$$

on $\mathcal{A}^n$ for each n and φ_π is defined analogously.

The following formulas implicitly define the classical, free, Boolean and monotone cumulants.

Theorem 4.1.4.

$$\varphi_\pi(X_1,\cdots,X_n) = \sum_{\substack{\sigma\in\mathcal{P}(n)\\ \sigma\le\pi}} K_\sigma(X_1,\ldots,X_n), \quad \textit{[Sch47, Rot64]} \tag{4.6}$$

$$\varphi_\pi(X_1,\cdots,X_n) = \sum_{\substack{\sigma\in\mathcal{NC}(n)\\ \sigma\le\pi}} R_\sigma(X_1,\ldots,X_n), \quad \textit{[Spe94]} \tag{4.7}$$

$$\varphi_\pi(X_1,\cdots,X_n) = \sum_{\substack{\sigma\in\mathcal{I}(n)\\ \sigma\le\pi}} B_\sigma(X_1,\ldots,X_n), \quad \textit{[SW97]} \tag{4.8}$$

$$\varphi(X_1\cdots X_n) = \sum_{(\sigma,\lambda)\in\mathcal{M}(n)} \tfrac{1}{|\sigma|!} H_\sigma(X_1,\ldots,X_n). \quad \textit{[HS11a]} \tag{4.9}$$

The multiplicative extension of the monotone case (4.9) is not very useful because the summand would depend on both σ,π (but if π is an interval partition, the summand does not depend on σ; and this is used to prove Theorem 4.2.1).

Let $\mu_\mathcal{P},\mu_\mathcal{NC},\mu_\mathcal{I}$ be the Möbius functions on the posets $\mathcal{P},\mathcal{NC},\mathcal{I}$ respectively. The values are

$$\mu_\mathcal{P}(\hat{0}_n,\hat{1}_n) = (-1)^{n-1}(n-1)!, \quad \text{[Sch47, Rot64]} \tag{4.10}$$

$$\mu_\mathcal{NC}(\hat{0}_n,\hat{1}_n) = (-1)^{n-1}C_{n-1}, \quad \text{[Kre72]} \tag{4.11}$$

$$\mu_\mathcal{I}(\hat{0}_n,\hat{1}_n) = (-1)^{n-1}, \tag{4.12}$$

where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the Catalan number. The values $\mu(\pi,\sigma)$ for general intervals $[\pi,\sigma]$ are products of these due to the fact that in all lattices considered here any such interval is isomorphic to a direct product of full lattices of different orders, see [DRS72, Spe94].

An interesting observation is that the lattice of interval partitions of order n is anti-isomorphic to the lattice of subsets of a set with $n-1$ elements and formula (4.12) is equivalent to the inclusion-exclusion principle.

From the Möbius principle we may express the classical, free and Boolean cumulants as

$$K_\pi(X_1,\cdots,X_n) = \sum_{\substack{\sigma\in\mathcal{P}(n)\\ \sigma\le\pi}} \varphi_\sigma(X_1,\ldots,X_n)\,\mu_\mathcal{P}(\sigma,\pi), \tag{4.13}$$

$$R_\pi(X_1,\cdots,X_n) = \sum_{\substack{\sigma\in\mathcal{NC}(n)\\ \sigma\le\pi}} \varphi_\sigma(X_1,\ldots,X_n)\,\mu_\mathcal{NC}(\sigma,\pi), \tag{4.14}$$

$$B_\pi(X_1,\cdots,X_n) = \sum_{\substack{\sigma\in\mathcal{I}(n)\\ \sigma\le\pi}} \varphi_\sigma(X_1,\ldots,X_n)\,\mu_\mathcal{I}(\sigma,\pi). \tag{4.15}$$

Alternatively, univariate cumulants can be defined via generating functions as we did in the introduction to this chapter.

Definition 4.1.5. *The* nesting forest $\mathfrak{F}(\pi)$ *of a non-crossing partition π with k blocks is the forest of planar rooted trees on k vertices built recursively as follows.*

1. *If π is an irreducible partition, then $\mathfrak{F}(\pi)$ is the planar rooted tree, whose vertices are the blocks of π, the root being the unique outer block, and branches $\mathfrak{F}(\pi_i)$ where π_i are the irreducible components of π without the outer block.*

2. *If π has irreducible components $\pi_1,\pi_2,\ldots,\pi_k$, then $\mathfrak{F}(\pi)$ is the forest consisting of the rooted trees $\mathfrak{F}(\pi_1),\mathfrak{F}(\pi_2),\ldots,\mathfrak{F}(\pi_k)$.*

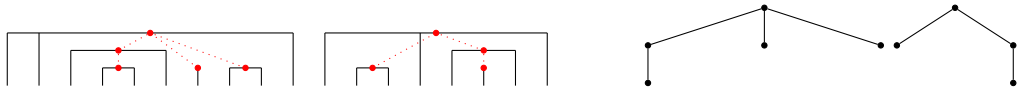

Fig. 4.2: A noncrossing partition and its nesting forest

Definition 4.1.6. *The* tree factorial $t!$ *of a finite rooted tree* t *is recursively defined as follows. Let* t *be a rooted tree with* $n > 0$ *vertices. If* t *consists of a single vertex, set* $t! = 1$*. Otherwise* t *can be decomposed into its root vertex and branches* t_1*,* t_2*,* $\ldots$*,* t_r *and we define recursively the number*

$$t! = n \cdot t_1! t_2! \cdots t_r!.$$

The tree factorial of a forest is the product of the factorials of the constituting trees.

For an arbitrary finite set $\mathcal{S}$ we denote by $\mathcal{P}(\mathcal{S})$ its set of partitions. Any bijection between $\mathcal{S}$ and $\{1, \ldots |\mathcal{S}|\}$ induces a poset isomorphism $\mathcal{P}(\mathcal{S})$ to $\mathcal{P}(|\mathcal{S}|)$. If $\mathcal{S}$ is totally ordered we consider the bijection which preserves this order and define $\mathcal{NC}(\mathcal{S}), \mathcal{I}(\mathcal{S})$ via this isomorphism.

Definition 4.1.7. *Let* $\pi \in \mathcal{P}(n)$*.*

1. *We define the crossing graph* $G(\pi) := (V, E)$ *of* π*, where the set of vertices* $V = \{V_1, \ldots, V_{|\pi|}\}$ *is indexed by the blocks*[1] *of* π *and an edge joins the vertices* V_i, V_j *if and only if they cross, i.e.,* $W = (V_i, V_j) \in (\mathcal{P}(V_i \cup V_j) \setminus \mathcal{NC}(V_i \cup V_j))$*.*

2. *Similarly, the vertices of the anti-interval graph* $\tilde{G}(\pi) := (V, E)$ *of* π *are just the blocks of* π*. An edge joining* (V_i, V_j) *is drawn if and only if* $W = (V_i, V_j) \in (\mathcal{P}(V_i \cup V_j) \setminus \mathcal{I}(V_i \cup V_j))$*. (For a non-crossing partition this is the nesting forest from Definition 4.1.5, augmented by the edges from all vertices to all their descendants).*

3. *For a finite graph* $G = (V, E)$ *and* $e \in E$*, we let* $G \setminus e = (V, E \setminus e)$*, and* $G/e = (V/e, E \setminus e)$ *be the graph obtained from removing* e *and identifying the endpoints of* e*. The* Tutte polynomial $T_G(x, y)$ *of* G *can be defined recursively by setting* $T_G(x, y) = 1$ *if* $E = \emptyset$ *and:*

$$T_G(x, y) = \begin{cases} xT_{G/e}(x, y) & \text{if } e \text{ is a bridge,} \\ yT_{G \setminus e}(x, y) & \text{if } e \text{ is a loop,} \\ T_{G/e}(x, y) + T_{G \setminus e}(x, y) & \text{otherwise.} \end{cases}$$

(where a bridge is and edge $e \in G$ *whose removal increases the number of connected components of* G*).*

[1] It should not cause confusion that we regard V_i simultaneously as a vertex of $G(\pi)$ and as a block of π

4.2 Cumulant-to-cumulant formulas

In [AHLV14] we extended the known relations between the different types of cumulants of a probability measure. The starting point are the following relations between cumulants, shown in [Leh02]:

$$b_n = \sum_{\pi\in\mathcal{NC}_{\mathrm{irr}}(n)} r_\pi, \tag{4.16}$$
$$r_n = \sum_{\pi\in\mathcal{P}_{\mathrm{conn}}(n)} \kappa_\pi, \tag{4.17}$$
$$b_n = \sum_{\pi\in\mathcal{P}_{\mathrm{irr}}(n)} \kappa_\pi.. \tag{4.18}$$

Relation (4.17) was used in [BBLS11] to attack the problem of free infinite divisibility of the normal law.

Relation (4.16) was extended by Belinschi and Nica in [BN08b] to the case of multivariate cumulants B_n, R_n. In addition, they obtained the inverse formula:

$$R_n = \sum_{\pi\in\mathcal{NC}_{\mathrm{irr}}(n)} (-1)^{|\pi|-1} B_\pi. \tag{4.19}$$

The extensions of (4.17) and (4.18) to the multivariate case can be shown by using the same proofs as in [Leh02] for the univariate case. An interesting inverse formula for (4.18) was proved recently by M. Josuat-Vergès [JV13], expressing classical cumulants in terms of free cumulants:

$$\kappa_n = \sum_{\pi\in\mathcal{P}_{\mathrm{conn}}(n)} (-1)^{1+|\pi|} T_{G(\pi)}(1,0) r_\pi, \tag{4.20}$$

where $G(\pi)$ is the crossing graph of π and $T_{G(\pi)}$ its Tutte polynomial. The proof of (4.20) in [JV13] is also valid for the multivariate case.

In [AHLV14] we completed the picture for the relations between classical, Boolean, free and monotone cumulants, extending some identities to the multivariate case. More precisely, we were able to prove the following cumulant identities.

Theorem 4.2.1. *The following identities hold for multivariate cumulants:*

$$B_n = \sum_{\pi\in\mathcal{M}_{\mathrm{irr}}(n)} \frac{1}{|\pi|!} H_\pi = \sum_{\pi\in\mathcal{NC}_{\mathrm{irr}}(n)} \frac{1}{\mathfrak{F}(\pi)!} H_\pi, \tag{4.21}$$
$$R_n = \sum_{\pi\in\mathcal{M}_{\mathrm{irr}}(n)} \frac{(-1)^{|\pi|-1}}{|\pi|!} H_\pi = \sum_{\pi\in\mathcal{NC}_{\mathrm{irr}}(n)} \frac{(-1)^{|\pi|-1}}{\mathfrak{F}(\pi)!} H_\pi. \tag{4.22}$$

By using Lenczewski's matricially free random variables and colored partitions [Len10, Len12] we obtained expressions for the univariate monotone cumulants.

Theorem 4.2.2. *The following identities hold for univariate cumulants:*

$$h_n = \sum_{\pi\in\mathcal{NC}_{\mathrm{irr}}(n)} \alpha_\pi r_\pi, \tag{4.23}$$
$$h_n = \sum_{\pi\in\mathcal{NC}_{\mathrm{irr}}(n)} (-1)^{|\pi|-1} \alpha_\pi b_\pi, \tag{4.24}$$
$$h_n = \sum_{\pi\in\mathcal{P}_{\mathrm{irr}}(n)} \alpha_{\bar{\pi}} \kappa_\pi, \tag{4.25}$$

where $\bar{\sigma} \in \mathcal{NC}(n)$ denotes the non-crossing closure of $\sigma \in \mathcal{P}(n)$ and α_π is the linear part of the number of non-increasing labellings of the nesting forest of π (which in the case of irreducible partitions consists of precisely one tree).

Remark 4.2.3. *Calculations indicate that a multivariate analogue of Theorem 4.2.2 also holds, but at present we do not know how to prove it.*

The proof of the Boolean-to-classical cumulant formula follows the techniques of the proof (4.20) which were used in [JV13] to prove more general formulas for cumulants of q-Gaussians.

Theorem 4.2.4.

$$K_n = \sum_{\pi \in \mathcal{P}_{\mathrm{irr}}(n)} (-1)^{1+|\pi|} T_{\tilde{G}(\pi)}(1,0) B_\pi, \tag{4.26}$$

where $\tilde{G}(\pi)$ is the anti-interval graph of π and $T_{\tilde{G}(\pi)}$ is its Tutte polynomial.

In [AHLV14] we also found another interpretation of classical cumulants in terms of Boolean cumulants via permutation statistics. Also, the values of the Tutte polynomials in (4.20) and (4.26) were interpreted as certain pyramids in the sense of Cartier-Foata [CF69]. In addition, we gave some partial results on the coefficients of the remaining, monotone-to-classical cumulant formula

$$K_n = \sum_{\pi \in \mathcal{P}(n)} \beta(\pi) H_\pi,$$

which seems to require a more detailed treatment. In particular, we were able to show that β is supported on $\mathcal{P}_{irr}$.

This concludes our section on cumulant formulas. Now we move to more specialized results for free case.

4.3 k-divisible partitions and products of random-variables

In this section we exploit the fact that the subposets of k-divisible and k-equal non-crossing partitions are linked, by the Kreweras complement, to partitions which are involved in the calculation of moments and free cumulants of the product of k free random variables.

It is helpful to picture partitions via their circular representation: We think of $[n]$ as the clockwise labeling of the vertices of a regular n-gon. If we identify each block of $\pi \in NC(n)$ with the convex hull of its corresponding vertices, then we see that π is non-crossing precisely when its blocks are pairwise disjoint (that is, they don't cross).

Out of two non-crossing partitions $\pi_1, \pi_2 \in NC(n)$ we can build the partition $\pi_1 \cup \pi_2$ by thinking $\pi_1 \in NC(\{1, 3, \ldots, 2n-1\})$, $\pi_2 \in NC(\{2, 4, \ldots, 2n\})$, and drawing them together. In general, the resulting partition may have crossings. Then, for a given $\pi \in NC(n) \cong NC(\{1, 3, \ldots, 2n-1\})$ we define its Kreweras complement

$$Kr(\pi) := \max\{\sigma \in NC(n) \cong NC(\{2, 4, \ldots, 2n\}) : \pi \cup \sigma \in NC(2n)\}.$$

The Kreweras complement [Kre72] satisfies many nice properties. The map $Kr : NC(n) \to NC(n)$ is an order reversing isomorphism. Furthermore, for all $\pi \in NC(n)$ we have that $|\pi| + |Kr(\pi)| = n + 1$.

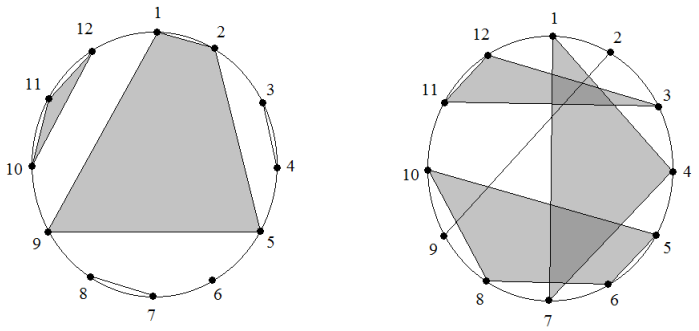

Fig. 4.3: The non-crossing partition $\{\{1,2,5,9\},\{3,4\},\{6\},\{7,8\},\{10,11,12\}\}$, and the crossing partition $\{\{1,4,7\},\{2,9\},\{3,11,12\},\{5,6,8,10\}\}$ of the set $[12]$ in their circular representations.

Definition 4.3.1. *We say that a non-crossing partition π is k-divisible if the sizes of all the blocks of are multiples of k. If, furthermore, all the blocks are exactly of size k, we say that π is k-equal. A partition $\pi \in NC(nk)$ is called k-preserving if all its blocks contain numbers with the same congruence modulo k. A k-preserving partition $\pi \in NC(nk)$ is called k-completing if it connects all blocks of the interval partition $\rho_k^n := \{\{1,2,\ldots,k\},\{k+1,\ldots,2k\},\ldots,\{k(n-1)+1,\ldots,kn\}\}$ (i.e. $\pi \vee \rho_k^n = \hat{1}_{nk}$).*

We will see that these concepts are closely related. We denote the set of k-divisible non-crossing partitions of $[kn]$ by $NC^k(n)$ and the set of k-equal non-crossing partitions of $[kn]$ by $NC_k(n)$.

The main tool used to derive our formulas is the following, proved by Krawczyk and Speicher [KS00].

Proposition 4.3.2 (Formula for products as arguments)**.** *Let $(\mathcal{A},\tau)$ be a non-commutative probability space and let $(R_n)_{n\geq 1}$ be the corresponding free cumulants. Let $m,n \geq 1$, $1 \leq i(1) < i(2) \cdots < i(m) = n$ be given and consider the partition*

$$\sigma = \{\{1,\ldots,i(1)\},\ldots,\{i(m-1)+1,\ldots,i(m)\}\} \in NC(n).$$

Consider random variables $a_1,\ldots,a_n \in \mathcal{A}$. Then the following equation holds:

$$R_m(a_1\cdots a_{i(1)},\ldots,a_{i(m-1)+1}\cdots a_{i(m)}) = \sum_{\substack{\pi\in NC(n)\\ \pi\vee\sigma=\hat{1}_n}} R_\pi(a_1,\ldots,a_n). \tag{4.27}$$

For certain special cases, the formula above runs over partitions with more structure which lead to nice interpretations. For example, if $a,b \in \mathcal{A}$ are free random variables, with free cumulants $R_n(a)$ and $R_n(b)$, respectively, one can calculate the free cumulants of ab by

$$R_n(ab) = \sum_{\pi\in\mathcal{NC}(n)} R_\pi(a)R_{Kr(\pi)}(b), \tag{4.28}$$

where $Kr(\pi)$ is the Kreweras complement of the non-crossing partition π. In particular, we are able to compute the free cumulants of the free multiplicative convolution of two probability measures with compact support μ, ν, such that $Supp(\mu) \subseteq [0, \infty)$ by

$$R_n(\mu \boxtimes \nu) = \sum_{\pi \in \mathcal{NC}(n)} R_\pi(\mu) R_{Kr(\pi)}(\nu). \tag{4.29}$$

In principle, this formula could be inductively used to provide the free cumulants and moments of the convolutions of k (not necessarily equal) positive probability measures. This approach, however, prevents us from noticing a deeper combinatorial structure behind such products of free random variables.

Our fundamental observation is that, when π and $Kr(\pi)$ are drawn together, the partition $\pi \cup Kr(\pi) \in \mathcal{NC}(2n)$ is exactly the Kreweras complement of a 2-equal partition (i.e. a non-crossing pairing). Furthermore, one can show using the previous correspondence that Equation (4.28) may be rewritten as

$$R_n(ab) = \sum_{\pi \in \mathcal{NC}_2(n)} R_{Kr(\pi)}(a, b, \ldots, a, b), \tag{4.30}$$

where $NC_2(n)$ denotes the 2-equal partitions of $[2n]$.

Since 2-equal partitions explain the free convolution of two variables, it is natural to try to describe the product of k free variables in terms of k-equal partitions.

The main result of this section is the following:

Theorem 4.3.3. *Let $a_1, \ldots, a_k \in (\mathcal{A}, \tau)$ be free random variables. Then the free cumulants and the moments of $a := a_1 \ldots a_k$ are given by*

$$R_n(a) = \sum_{\pi \in NC_k(n)} R_{Kr(\pi)}(a_1, \ldots, a_k), \tag{4.31}$$

$$\tau(a^n) = \sum_{\pi \in NC^k(n)} R_{Kr(\pi)}(a_1, \ldots, a_k), \tag{4.32}$$

where $NC_k(n)$ and $NC^k(n)$ denote, respectively, the k-equal and k-divisible partitions of $[kn]$.

Proof. See Section 4.3.1 □

The main argument is the following observation.

Proposition 4.3.4. *i) $\pi \in NC(kn)$ is k-preserving if and only if $\pi = Kr(\sigma)$ for some k-divisible partition $\sigma \in NC^k(n)$.*

ii) $\pi \in NC(kn)$ is k-completing if and only if $\pi = Kr(\sigma)$ for some k-equal partition $\sigma \in NC_k(n)$.

Proof. See Section 4.3.1 □

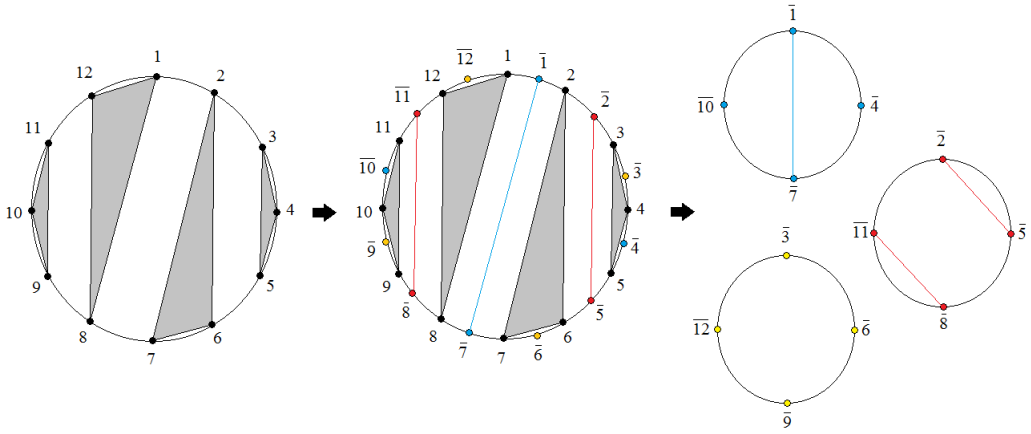

Fig. 4.4: The 3-equal partition $\pi = \{\{1,8,12\},\{2,6,7\},\{3,4,5\},\{9,10,11\}\}$ and its Kreweras complement $Kr(\pi) = \pi_1 \cup \pi_2 \cup \pi_3$, with $\pi_1 = \{\{1,7\},\{4\},\{10\}\}$, $\pi_2 = \{\{2,5\},\{8,11\}\}$ and $\pi_3 = \{\{3\},\{6\},\{9\},\{12\}\}$.

Remark 4.3.5. *In view of the previous characterization, for a k-divisible partition π, the Kreweras complement $Kr(\pi)$ may be divided into k partitions $\pi_1, \pi_2, \ldots, \pi_k$, with π_j involving only numbers congruent to j mod k. In this case we will write $\pi_1 \cup \cdots \cup \pi_k = Kr(\pi)$ for such decomposition.*

In [AV12], we used such formulas, together with enumerative results on k-divisible partitions to give a new proof of the fact (first proved by Kargin [Kar07]) that for positive measures centered at 1, the support of the free multiplicative convolution $\mu^{\boxtimes k}$ grows at most linearly. Moreover, our approach enabled us to generalize to the case $\mu_1 \boxtimes \cdots \boxtimes \mu_k$, as follows.

Theorem 4.3.6. *There exists a universal constant $C > 0$ such that for all k and any $\mu_1, \ldots, \mu_k$ probability measures supported on $[0, L]$, satisfying $E(\mu_i) = 1$ and $Var(\mu_i) \geq \sigma^2$, for $i = 1, \ldots, k$, the supremum L_k of the support of the measure $\mu_1 \boxtimes \cdots \boxtimes \mu_k$ satisfies*

$$\sigma^2 k \leq L_k < CLk.$$

In other words, for (not necessarily identically distributed) positive free random variables $(X_i)_{i \geq 1}$ such that $E(X_i) = 1$, $Var(X_i) \geq \sigma^2$ and $||X_i|| \leq L$, $i \geq 1$, we have that

$$\limsup_{n \to \infty} n^{-1} ||X_1^{1/2} \cdots X_{n-1}^{1/2} X_n X_{n-1}^{1/2} \cdots X_1^{1/2}|| < CL$$

and

$$\liminf_{n \to \infty} n^{-1} ||X_1^{1/2} \cdots X_{n-1}^{1/2} X_n X_{n-1}^{1/2} \cdots X_1^{1/2}|| \geq \sigma^2.$$

Let us point out that for the case $\mu_1 = \cdots = \mu_k$, the previous theorem can be proved as using the methods of [KS07]. However, the norm estimates given there are meant to address more general situations (where certain linear combinations of products are allowed) and hence, the constants obtained using these methods for our specific problem are far from optimal.

Our theorem also allowed us to provide a new proofs to the limit theorems of Sakuma and Yoshida [SY13].

4.3.1 Proofs of Theorem 4.3.3 and Proposition 4.3.4

Remark 4.3.7. *A useful characterization of non-crossing partitions is that, for any $\pi \in NC(n)$, one can always find an interval block $V = \{r+1, \ldots, r+s\}$ such that if one removes this block from π, the partition $\pi \setminus V \in NC(n-s)$ remains non-crossing.*

For a partition $\pi \in NC(n)$ will often write $r \sim_\pi s$, meaning that r, s belong to the same block of π.

Let us introduce two operations on non-crossing partitions. For $n, k \geq 1$ and $r \leq n$, we define $I_r^k : NC(n) \to NC(n+k)$, where $I_r^k(\pi)$ is obtained from π by duplicating the element in the position r, identifying the copies and inserting $k-1$ singletons between the two copies. More precisely, for $\pi \in NC(n)$, $I_k^r(\pi) \in NC(n+k)$ is the partition given by the relations:

1. For $1 \leq m_1, m_2 \leq r$,
$$m_1 \sim_{I_r^k(\pi)} m_2 \ \Leftrightarrow m_1 \sim_\pi m_2.$$
2. For $r+k \leq m_1, m_2 \leq n+k$,
$$m_1 \sim_{I_r^k(\pi)} m_2 \ \Leftrightarrow m_1 - k \sim_\pi m_2 - k.$$
3. For $1 \leq m_1 \leq r$ and $r+k+1 \leq m_2 \leq n+k$,
$$m_1 \sim_{I_r^k(\pi)} m_2 \ \Leftrightarrow m_1 \sim_\pi m_2 - k.$$
4. $r \sim_{I_r^k(\pi)} r+k$.

The operation $\tilde{I}_r^k : NC(n) \to NC(n+k)$ consists of inserting an interval block of size k between the positions $r-1$ and r in π. We will skip the explicit definition.

The importance of these operations is that they are linked by the relation

$$Kr(I_r^k(\pi)) = \tilde{I}_r^k(Kr(\pi)). \tag{4.33}$$

Our operations preserve properties of partitions, as shown in the following lemma.

Lemma 4.3.8. *Let $\pi \in NC(nk)$, $r \leq nk$, $s \geq 1$. Then*
i) π is k-preserving if and only if $I_r^{sk}(\pi)$ is k-preserving.
ii) π is k-completing if and only if $I_r^k(\pi)$ is k-completing.
iii) π is k-divisible if and only if $\tilde{I}_r^{sk}(\pi)$ is k-divisible.
iv) π is k-equal if and only if $\tilde{I}_r^k(\pi)$ is k-equal.

Proof. i) By definition of $I_r^k(\pi)$, the relations indicated by $I_r^k(\pi)$ are obtained by relations indicated by π, with possible shifts by ks (which do not modify congruences modulo k). Hence the equivalence follows.

ii) One should think of the block intervals of ρ_n^k as vertices of a graph. For $\pi \in NC(nk)$, an edge will join two vertices V, W, if there are elements $r \in V$, $s \in W$ such that $r \sim_\pi s$. Then $\pi \vee \rho_n^k = 1_{nk}$ if and only if the graph is connected.

It is easy to see that the effect of I_r^k on the graph of π is just splitting the vertex corresponding to the block V containing r into 2 vertices V_1, V_2. The edges between all other vertices are preserved, while the edges which were originally joined to V will now be joined either to V_1 or V_2. Finally, the last additional relation $r \sim_{I_r^k(\pi)} r + k$ means an edge joining V_1 to V_2. Therefore, it is clear that the connectedness of the two graphs are equivalent.

iii) and iv) are trivial. □

Now we want to show that we can produce all partitions of our interest by applying our operations to elementary partitions.

Lemma 4.3.9. *i) Let $\pi \in NC(kn)$ be k-preserving. Then there exist $m \geq 0$ and numbers $q_0, q_1, \ldots, q_m$, $r_1, \ldots, r_m$ such that*

$$\pi = I_{r_m}^{kq_m} \circ \cdots \circ I_{r_1}^{kq_1}(0_{q_0}). \tag{4.34}$$

ii) Let $\pi \in NC(kn)$ be k-completing. Then there exist $m \geq 0$ and numbers $r_1, \ldots, r_m$ such that

$$\pi = I_{r_m}^{k} \circ \cdots \circ I_{r_1}^{k}(0_{k}). \tag{4.35}$$

iii) Let $\pi \in NC(kn)$ be k-divisible. Then there exist $m \geq 0$ and numbers $q_0, q_1, \ldots, q_m$, $r_1, \ldots, r_m$ such that

$$\pi = \tilde{I}_{r_m}^{kq_m} \circ \cdots \circ \tilde{I}_{r_1}^{kq_1}(1_{q_0}). \tag{4.36}$$

iv) Let $\pi \in NC(kn)$ be k-equal. Then there exist $m \geq 0$ and numbers $r_1, \ldots, r_m$ such that

$$\pi = \tilde{I}_{r_m}^{k} \circ \cdots \circ \tilde{I}_{r_1}^{k}(1_{k}). \tag{4.37}$$

Proof. i) We use induction on n. For $n = 1$ the only k-preserving partition is 0_k, so the statement holds. So assume that i) holds for $n \leq m$. For $\pi \in NC^k(m)$ suppose that there exist $1 \leq r < r + sk \leq km$ such that $r \sim_\pi r + sk$ and $r + 1, \ldots r + sk - 1$ are singletons of π (if no such pair (r, s) exist, necessarily $\pi = 0_{mk}$ and we are done). Then its easy to see that $\pi = I_r^{sk}(\pi')$ for some $\pi' \in NC((n - s)k)$. By Lemma 4.3.8 i) π' is k-preserving. By induction hypothesis π' has a representation as in Equation (4.34) and hence, so does $\pi = I_r^{sk}(\pi')$.

The proof of ii) is similar. The proofs of iii) and iv) are trivial using Remark 4.3.7. □

Proof of Proposition 4.3.4. We only show the first implication of i). The converse and ii) are similar.

Let $\pi \in NC(kn)$ be k-preserving. Then by Lemma 4.3.9 i) we can express it as

$$\pi = I_{r_m}^{kq_m} \circ \cdots \circ I_{r_1}^{kq_1}(1_{q_0}).$$

But then we can apply Equation (4.33) at every step, obtaining

$$\begin{aligned} Kr(\pi) &= Kr(I_{r_m}^{kq_m} \circ \cdots \circ I_{r_1}^{kq_1}(0_{q_0})) && (4.38)\\ &= \tilde{I}_{r_m}^{kq_m} \circ Kr(I_{r_{m-1}}^{kq_{m-1}} \cdots \circ I_{r_2}^{kq_2} \circ I_{r_1}^{kq_1}(1_{q_0})) && (4.39)\\ &= \tilde{I}_{r_m}^{kq_m} \circ \tilde{I}_{r_{m-1}}^{kq_{m-1}} \circ Kr(I_{r_{m-2}}^{kq_{m-2}} \circ \cdots \circ I_{r_2}^{kq_2} \circ I_{r_1}^{kq_1}(1_{q_0})) && (4.40)\\ &\quad \vdots && (4.41)\\ &= \tilde{I}_{r_m}^{kq_m} \circ \cdots \circ \tilde{I}_{r_2}^{kq_2} \circ \tilde{I}_{r_1}^{kq_1}(Kr(1_{q_0})) && (4.42)\\ &= \tilde{I}_{r_m}^{kq_m} \circ \cdots \circ \tilde{I}_{r_2}^{kq_2} \circ \tilde{I}_{r_1}^{kq_1}(0_{q_0}), && (4.43) \end{aligned}$$

which, by Lemma 4.3.8 iii) is k-divisible. □

Now we can prove the main result of this section.

Proof of Theorem 4.3.3. By the formula for products as arguments, we have that

$$R_n(a) = \prod_{\substack{\pi \in NC(kn)\\ \pi \vee \rho_n^k = \hat{1}_{nk}}} R_\pi(a_1, \ldots, a_n).$$

Since the random variables are free, the sum runs actually over k-preserving partitions (otherwise there would be a mixed, hence vanishing cumulant). But then by Proposition 4.3.4 ii), the partitions involved in the sum are exactly the Kreweras complements of k-equal partitions, and the formula follows.

For the proof of (4.32), we use the moment-cumulant formula

$$\tau(a^n) = \sum_{\pi \in NC(kn)} R_\pi(a_1, \ldots, a_n).$$

Again, the elements involved are free, so only k-preserving partitions matter, and these are the Kreweras complements of k-divisible partitions by Proposition 4.3.4 i). Hence the result follows. □

Surprisingly, the Boolean cumulants of products of *free* random variables also satisfy the same formula.

Remark 4.3.10. *As pointed out in [BN08a], Equation (4.28) is also satisfied when we replace the free cumulants by the Boolean cumulants. Therefore, Formula (4.31) holds as well for Boolean cumulants $(b_n)_{n\geq 1}$, namely, if $a := a_1 \cdots a_k$ is a product of free random variables, we have*

$$b_n(a) = \sum_{\pi \in NC_k(n)} b_{Kr(\pi)}(a_1, \ldots, a_k). \tag{4.44}$$

5 Operator-valued free probability

The idea of operator-valued free probability [Voi95] is to generalize free probability, by replacing $\tau : \mathcal{A} \to \mathbb{C}$ by a conditional expectation $\mathbf{F} : \mathcal{A} \to \mathcal{B}$ onto a larger sub-algebra $\mathbb{C} \subseteq \mathcal{B} \subseteq \mathcal{A}$. This leads to a broader definition of freeness, which occurs in more general situations of random matrix theory, as observed first by Shlyakthenko [Shl96]. Many aspects of the theory of (scalar-valued) free probability can be lifted to the operator-valued level. The combinatorics of operator-valued free probability (see [Spe98]) remains the same provided that the nesting structure of non-crossing partitions is respected while operating. This makes cumulants particularly useful.

The free cumulants also allow to find good candidates for the smallest sub-algebra $\mathcal{B} \subset \mathcal{A}$ over which two given random variables are free. In [NSS02] more general results were found in this direction.

We include some examples of distributions of $dN \times dN$ matrix models which have been modified block-wise by a self-adjoint linear map $\varphi : M_d(\mathbb{C}) \to M_d(\mathbb{C})$. Such models are relevant in quantum information theory (see [Aub12, BN12a, ANV]). A general numerical solution to the problem of the asymptotic block-modified distribution can be obtained using the algorithm to compute free matrix-valued multiplicative convolutions from [BSTV14].

The question of finding explicit formulas for operator-valued distributions is closely related to the possibility of finding realizations of the distribution in terms of operators which are free over a commutative algebra. In [ANV], we use the criteria in [NSS02] to give sufficient conditions (in terms of the Choi matrix of the map φ) for such a realization to exist.

For the case of non-commutative algebras, the analytical side of free probability has the drawback of being extremely complicated if one tries to obtain exact distributions. On the other hand, numerical algorithms which rely on subordination, such as the ones in [BB07] for the computation of the additive and multiplicative free convolutions admit very effective generalizations (see [BSTV14, BMS13]).

Our main goal in this chapter is to understand freeness on rectangular spaces and matrix-valued probability spaces. Our FDE's for polynomial random matrix models will be defined there.

5.1 Rectangular and matrix-valued probability Spaces

Definitions 5.1.1. *(1). Let $\mathcal{A}$ be a unital $*$-algebra and let $\mathbb{C} \subseteq \mathcal{B} \subseteq \mathcal{A}$ be a $*$-sub-algebra. A $\mathcal{B}$-probability space is a pair $(\mathcal{A}, \mathbf{F})$, where $\mathbf{F} : \mathcal{A} \to \mathcal{B}$ is a conditional*

expectation, that is, a linear map satisfying:

$$\begin{aligned} \mathbf{F}\left(bab'\right) &= b\mathbf{F}(a)b', \qquad \forall b, b' \in \mathcal{B}, a \in \mathcal{A} \\ \mathbf{F}\left(1\right) &= 1. \end{aligned}$$

(2). Let $(\mathcal{A}, \mathbf{F})$ be a $\mathcal{B}$-probability space and let $\bar{a} := a - \mathbf{F}(a)1_{\mathcal{A}}$ for any $a \in \mathcal{A}$. The $*$-subalgebras $\mathcal{B} \subseteq A_1, \ldots, A_k \subseteq \mathcal{A}$ are $\mathcal{B}$-free (or free over $\mathcal{B}$, or free with amalgamation over $\mathcal{B}$) (with respect to $\mathbf{F}$) iff

$$\mathbf{F}(\bar{a_1}\bar{a_2}\cdots\bar{a_m}) = 0, \tag{5.1}$$

for all $m \geq 1$ and all tuples $a_1, \ldots, a_m \in \mathcal{A}$ such that $a_i \in A_{j(i)}$ with $j(1) \neq j(2) \neq \cdots \neq j(m)$.

(3). Subsets $S_1, \ldots, S_k \subset \mathcal{A}$ are $\mathcal{B}$-free if so are the $*$-subalgebras $\langle S_1, \mathcal{B}\rangle, \ldots, \langle S_k, \mathcal{B}\rangle$.

We extend Definition 3.2.2 to the operator-valued case.

Definition 5.1.2. *For a tuple $a = (a_1, \ldots, a_k)$ of random variables in a $\mathcal{B}$-probability space $(\mathcal{A}, \mathbf{F})$, the m-th order $\mathcal{B}$-valued moments of a are the collection of maps $\Phi_m^{\mathcal{B};a} := (\Phi_m^{(i_1,\ldots,i_m)})_{i_1,\ldots,i_m \leq k}$, where*

$$\Phi_m^{(i_1,\ldots,i_m)} : \mathcal{B}^{m-1} \rightarrow \mathcal{B} \tag{5.2}$$

$$(b_1, \ldots, b_{m-1}) \mapsto \mathbf{F}(a_{i_1} b_1 \ldots b_{m-1} a_{i_m}) \tag{5.3}$$

We call $\Phi(a) = \bigcup_{m\geq 0} \Phi_m^a$ the joint distribution of $a = (a_1, \ldots, a_k)$.

We always work with tuples of the form $a = (a_1, a_1^*, \ldots, a_k, a_k^*)$ but we will omit the adjoints.

5.1.1 Rectangular probability spaces

Let $(\mathcal{A}, \tau)$ be a tracial $*$-probability space endowed with pairwise orthogonal, non-trivial projections $p_1, \ldots, p_k \in \mathcal{A}$ adding up to one. Let $\mathcal{D} := \langle p_1, \ldots, p_k\rangle$ denote the $*$-algebra generated by $\{p_1, \ldots, p_k\}$. Then there exists a unique conditional expectation $\mathbf{F} : \mathcal{A} \rightarrow \mathcal{D}$ such that $\tau \circ \mathbf{F} = \tau$, which is given by the sum of compressions

$$\mathbf{F}\left(a\right) = \sum_{i=1}^{k} p_i \tau\left(p_i\right)^{-1} \tau\left(p_i a\right). \tag{5.4}$$

With this, $(\mathcal{A}, \mathbf{F})$ becomes a $\mathcal{D}$-valued probability space.

These kind of projection-valued spaces were introduced by Benaych-Georges in [BG09a] (see also [BG09b]) and are called rectangular probability spaces. We will denote by $\mathcal{A}^{(i,j)}$ the set of elements $a \in \mathcal{A}$ such that $a = p_i a p_j$. Elements in $\bigcup_{1\leq i,j\leq k} \mathcal{A}^{(i,j)}$ are called *simple* and we write $\mathcal{A}^{(i)} := \mathcal{A}^{(i,i)}$. Very often we will be interested in the compressed spaces $(\mathcal{A}^{(i)}, \tau^{(i)})$, where $\tau^{(i)}\left(a\right) = \tau(p_i)^{-1}\tau\left(a\right)$, for $a \in \mathcal{A}^{(i)}$.

We will make use of the caligraphic letter $\mathcal{D}$, to emphasize the situation when $\mathcal{D} = \mathcal{B} \subset \mathcal{A}$ will always mean freeness with amalgamation over the algebra of projections which defines the given rectangular probability space.

Definition 5.1.3. *Let $(\mathcal{A},\tau)$ be a $\mathcal{B}$-rectangular space and let $(\mathcal{A}_n,\tau_n)_{n\geq 1}$ be a sequence of $(p_1^{(n)},\dots,p_k^{(n)})$-rectangular spaces. Let $a_1,\dots,a_m\in\mathcal{A}$ and $a_1^{(n)},\dots,a_m^{(n)}\in\mathcal{A}_n$ be collections of simple elements. We say that $(a_1^{(n)},\dots,a_m^{(n)})$ converges in $\mathcal{D}$-distribution to $(a_1,\dots,a_m)$ if $(a_1^{(n)},\dots,a_m^{(n)},p_1^{(n)},\dots,p_k^{(n)})\in(\mathcal{A}_N^{m+k},\tau_N)$ converges in distribution to $(a_1,\dots,a_m,p_1,\dots,p_k)\in(\mathcal{A}^{m+k},\tau)$, and we write*

$$\left(a_1^{(n)},\dots,a_m^{(n)}\right)\xrightarrow{\mathcal{D}}(a_1,\dots,a_m)\qquad \text{as } n\to\infty.$$

If $a_1,\dots,a_m$ are $\mathcal{D}$-free, we say that $a_1^{(n)},\dots,a_m^{(n)}$ are asymptotically $\mathcal{D}$-free.

We reformulate some results by Benaych-Georges in [BG09b] generalizing, in the framework of rectangular spaces, Voiculescu's results on asymptotic freeness of square matrices.

Theorem 5.1.4 (Asymptotic freeness of rectangular matrix ensembles [BG09b])**.** *Let $k\geq 1$ be fixed. For each $N\geq 1$, let $P_1^{(N)},\dots,P_k^{(N)}$ be pairwise orthogonal projections, such that $I_N=P_1^{(N)}+\cdots+P_k^{(N)}$ and $N^{-1}\mathrm{Tr}(P_i^{(N)})\to\rho_i\in(0,1)$ for each $i\leq k$.*

Let $X_1^{(N)},\dots,X_p^{(N)}$ and $U_1^{(N)},\dots,U_q^{(N)}$ be independent random matrices, such that

- *for each $i\leq p$, $X_i^{(N)}=P_{j(i)}^{(N)}X_i^{(N)}P_{j(i)}^{(N)}$ is a Wigner Matrix when restricted to $(\mathcal{A}_N^{j(i)},\tau_N^{j(i)})$ for some $1\leq j(i)\leq k$.*
- *for each $i\leq q$, $U_i^{(N)}=P_{h(i)}^{(N)}U_i^{(N)}P_{h(i)}^{(N)}$ is a Haar-unitary random matrix when restricted to $(\mathcal{A}_N^{h(i)},\tau_N^{h(i)})$, for some $1\leq h(i)\leq k$.*

Let $D_1^{(N)},\dots,D_r^{(N)}$ be deterministic matrices, with $D_i^{(N)}=P_{h_1(i)}^{(N)}D_i^{(N)}P_{h_2(i)}^{(N)}$ for some $1\leq h_1(i),h_2(i)\leq k$, such that for any $m\geq 1$, $1\leq i_1,\dots,i_m\leq r$, there exist $c(i_1,\dots,i_m)\in\mathbb{C}$ such that

$$\lim_{N\to\infty}\frac{1}{N}Tr(D_{i_1}^{(N)}D_{i_2}^{(N)}\cdots D_{i_m}^{(N)})=c(i_1,\dots,i_m). \tag{5.5}$$

Then, as $N\to\infty$,

$$(X_1^{(N)},\dots,X_p^{(N)},U_1^{(N)},\dots,U_q^{(N)},D_1^{(N)},\dots,D_r^{(N)})\xrightarrow{\mathcal{D}}(s_1,\dots,s_p,u_1,\dots u_q,d_1,\dots d_r)$$

where $s_1,\dots,s_p,u_1,\dots u_q,d_1,\dots d_r$ are elements in some rectangular probability space $(\mathcal{A},\tau)$ (with orthogonal projections $p_1,\dots p_k$) whose joint $\mathcal{D}$-distribution is determined by the following conditions:

- *$\tau(p_i)=\rho_i$.*
- *For all $i\leq p$, $s_i=p_{j(i)}s_ip_{j(i)}$ is a semicircular in the compressed space $(\mathcal{A}^{j(i)},\tau^{j(i)})$.*
- *For all $i\leq q$, $u_i=p_{j(i)}u_ip_{j(i)}$ is a Haar-unitary in the compressed space $(\mathcal{A}^{h(i)},\tau^{j(i)})$.*

- *For all $i \leq r$, $d_i = p_{h_1(i)} d_i p_{h_2(i)}$. In addition, $\tau(d_{i_1} d_{i_2} \cdots d_{i_k}) = c(i_1, \ldots, i_k)$, for any $k \geq 1$, $1 \leq i_1, \ldots, i_k \leq r$.*
- *The algebras $\langle s_1 \rangle, \ldots, \langle s_p \rangle, \langle u_1, u_1^* \rangle, \ldots, \langle u_q, u_q^* \rangle, \langle d_1, d_1^* \ldots, d_r, d_r^* \rangle$ are free with amalgamation over $\langle p_1, \ldots p_k \rangle$.*

Benaych-Georges stated the Gaussian and Haar-unitary cases separately and the proof followed the same direction as Voiculescu's proof of the scalar case.

From the combinatorial point of view, the case of Wigner and deterministic matrices can be essentially treated as explained in Section 3.2.3. Again, we will need to compute the asymptotics of sums of products of deterministic matrices indexed by partitions $D_{\pi,\varepsilon}$. The main difference now is that each summand of $D_{\pi,\varepsilon}$ gets multiplied by $K_\pi(z_{i_1 i_2}^{\varepsilon_1}, z_{i_3 i_4}^{\varepsilon_2}, \ldots, z_{i_{2k-1} i_{2k}}^{\varepsilon_k})$, which is no longer fixed as we vary $i_1, \ldots i_{2k}$ because the variance of $z_{i_j i_{j+1}}$ is scaled according the different sizes of the Wigner matrices. However, since the ratios of the different sizes of the matrices converge, we can use an uniform bound for all such

$$K_\pi(z_{i_1 i_2}^{\varepsilon_1}, z_{i_3 i_4}^{\varepsilon_2}, \ldots, z_{i_{2k-1} i_{2k}}^{\varepsilon_k}),$$

and proceed as before.

The combinatorial arguments which show that only non-crossing pairings yield non-vanishing contributions remain valid for this case. So we are again left with non-crossing pairings. From there, it is an easy exercise to check that, from eq. (5.4) and Def. 5.1 one indeed obtains the same mixed moments for $(s_1, \ldots, s_p, d_1, \ldots d_r)$.

The adaptation of the combinatorial proofs for Haar-unitary matrices and deterministic matrices is explained in [SV12] and it is clear that both methods can be combined to consider mixed moments. One can lift all the arguments in Remark 3.2.5 (1)-(4) to the rectangular setting.

Example 5.1.5. *Consider the model*

$$\Phi = \sum_{i=1}^{K} R_i U_i T_i U_i^* R_i^*$$

of [CHD11] (see Section 7.3.1 for the precise assumptions on the matrices). If we let $M = N_0 + \cdots + N_K$, we may consider orthogonal projections $P_0, \ldots, P_K$ with $\mathrm{Tr}(P_i) = N_i$ and think that each matrix A of the model is embedded as a simple element $\tilde{A}$ in the $(P_0, \ldots, P_K)$-probability space $M_M(\mathbb{C})$, in such a way that $P_0 \tilde{R}_i P_i = \tilde{R}_i$, $P_i \tilde{U}_i P_i = \tilde{U}_i$ and $P_i \tilde{T}_i P_i = \tilde{T}_i$ as illustrated in Fig. 5.1.

Observe that the embedded matrices fit together, and we have that

$$\sum_{i=1}^{K} \tilde{R}_i \tilde{U}_i \tilde{T}_i \tilde{U}_i^* \tilde{R}_i^* = P_0 (\sum_{i=1}^{K} R_i U_i T_i U_i^* R_i^*) P_0 = P_0 \Phi P_0.$$

Hence, all the information about the distribution of Φ is contained in the joint $\langle P_0, \ldots, P_K \rangle$-distribution of $(\tilde{R}_1, \tilde{U}_1, \tilde{T}_1, \ldots, R_K, \tilde{U}_K, \tilde{T}_K)$.

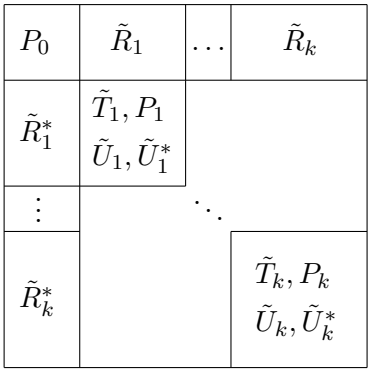

Fig. 5.1: Embedding of the Matrices of Φ on a Rectangular Space

In order to simplify notation for the rest of this example and all future rectangular matrix models, we will omit the tilde and the embedding discussions. In the context of a rectangular space, we will think directly that the matrices of the model are already embedded as above.

Now, for each $m \geq 1$, let us consider the blown-up model

$$\Phi^{(N,m)} = \sum_{i=1}^{K} R_i^{(m)} U_i^{(m)} T_i^{(m)} (U_i^{(m)})^* (R_i^{(m)})^* = \sum_{i=1}^{K} (R_i \otimes I_m) U_i^{(m)} (T_i \otimes I_m) (U_i^{(m)})^* (R_i \otimes I_m)^*,$$

where the matrices are now considered in the blown-up rectangular space generated by the orthogonal projections $(P_0 \otimes I_m), \ldots, (P_K \otimes I_m) \in M_{Mm}(\mathbb{C})$, where $(P_i \otimes I_m) U_i^{(m)} (P_i \otimes I_m) = U_i^{(m)}$ are independent and Haar-distributed in the compressed spaces $\mathcal{U}(mN_i) \subset (P_i \otimes I_m) M_{Mm}(\mathbb{C}) (P_i \otimes I_m)$.

By Theorem 5.1.4,

$$(R_1^{(m)}, U_1^{(m)}, T_1^{(m)}, \ldots, R_K^{(m)}, U_K^{(m)}, T_K^{(m)}) \quad \xrightarrow{\mathcal{D}} \quad (r_1, u_1, t_1, \ldots, r_K, u_K, t_K) \tag{5.6}$$

in some $\langle p_0, \ldots, p_K \rangle$-rectangular probability space $(\mathcal{A}, \tau)$ with $\tau(p_i) = N_i/M$, where $u_1, \ldots u_k, \langle r_1, t_1, \ldots, r_K, t_K \rangle$ are $\mathcal{D}$-free, $u_i|_{\mathcal{A}^{(i)}}$ is a Haar unitary and

$$\tau(r_{i_1}^{\varepsilon_1} t_{j_1}^{\delta_1} \cdots r_{i_s}^{\varepsilon_s} t_{j_s}^{\delta_s}) = \frac{1}{M} \mathrm{Tr}(R_{i_1}^{\varepsilon_1} T_{j_1}^{\delta_1} \cdots R_{i_s}^{\varepsilon_s} T_{j_s}^{\delta_s}),$$

for all $s \geq 1$, $i_1, j_1, \ldots, i_s, j_s \leq K$, and $\varepsilon_1, \delta_1, \ldots, \varepsilon_s, \delta_s \in \{0, 1, *\}$.

Hence $\Phi^{(N)} \to p_0 \Phi^{\square} p_0 = \Phi^{\square} := \sum_{i=1}^{K} r_i u_i t_i u_i^* r_i^* \in \mathcal{A}^{(0)}$. The nature of the joint $\mathcal{D}$-distribution of $(r_1, u_1, t_1, \ldots, r_K, u_K, t_K)$ falls into a general setting where our algorithm works (see Chapter 7), allowing us to compute the spectral distribution of $\Phi^{\square}$.

5.1.2 Matrix-valued probability spaces

Example 5.1.6 (Matrix-valued probability spaces)**.** *Let $(\mathcal{A}, \tau)$ be a $*$-probability space and consider the algebra $M_n(\mathcal{A}) \cong M_n(\mathbb{C}) \otimes \mathcal{A}$ of $n \times n$ matrices with entries in $\mathcal{A}$. The*

maps

$$\mathbf{F}_3 : (a_{ij})_{ij} \mapsto (\tau(a_{ij}))_{ij} \in M_n(\mathbb{C}),$$
$$\mathbf{F}_2 : (a_{ij})_{ij} \mapsto (\delta_{ij}\tau(a_{ij}))_{ij} \in D_n(\mathbb{C}),$$

and

$$\mathbf{F}_1 : (a_{ij})_{ij} \mapsto \sum_{i=1}^{n} \frac{1}{n}\tau(a_{ii}) I_n \in \mathbb{C} \cdot I_n$$

are respectively, conditional expectations onto the algebras $M_n(\mathbb{C}) \supset D_n(\mathbb{C}) \supset \mathbb{C} \cdot I_n$ of constant matrices, diagonal matrices and multiples of the identity.

Observe that $(M_n(\mathcal{A}), \mathbf{F}_1)$ is again a scalar-valued *-probability space. If $A_1, \ldots, A_k$ are free in $(\mathcal{A}, \tau)$, then the algebras $M_n(A_1), \ldots, M_n(A_k)$ of matrices with entries in $A_1, \ldots, A_k$ respectively are in general not free over $\mathbb{C}$ (w.r.t. $\mathbf{F}_1$). They are, however $M_n(\mathbb{C})$-free (w.r.t. $\mathbf{F}_3$). Below is a slightly more general assertion of this simple but fundamental result.

Proposition 5.1.7. *Let $(\mathcal{A}, \mathbf{F})$ be a $\mathcal{B}$-probability space, and consider the $M_n(\mathcal{B})$-valued probability space $(M_n(\mathbb{C}) \otimes \mathcal{A}, id \otimes \mathbf{F})$. If $A_1, \ldots, A_k \subseteq \mathcal{A}$ are $\mathcal{B}$-free, then $(M_n(\mathbb{C}) \otimes A_1), \ldots, (M_n(\mathbb{C}) \otimes A_k) \subseteq (M_n(\mathbb{C}) \otimes \mathcal{A})$ are $(M_n(\mathcal{B}))$-free.*

Proof. Let $a^{(1)}, \ldots, a^{(m)} \in M_n(\mathbb{C}) \otimes \mathcal{A}$ be such that $a^{(i)} \in M_n(\mathbb{C}) \otimes A_{j(i)}$ with $j(1) \neq j(2) \neq \cdots \neq j(m)$. Observe that

$$\overline{a^{(i)}} = a^{(i)} - (id \otimes \mathbf{F})(a^{(i)}) = ((a^{(i)}_{rs}) - \mathbf{F}(a^{(i)}_{rs}))_{rs \leq n} = (\overline{a^{(i)}_{rs}})_{rs \leq n}.$$

Hence

$$(id \otimes \mathbf{F})((\overline{a^{(1)}}) \cdots (\overline{a^{(m)}})) = \sum_{i_0, \ldots, i_m = 1}^{n} (\mathbf{F}((\overline{a^{(1)}_{i_0 i_1}})(\overline{a^{(2)}_{i_1 i_2}}) \cdots (\overline{a^{(m)}_{i_{m-1} i_m}})))_{i_0 i_m} = 0. \tag{5.7}$$

□

5.2 Combinatorics of operator-valued free probability

For $n \in \mathbb{N}$, a $\mathbb{C}$-multi-linear map $f : \mathcal{A}^n \to \mathcal{B}$ is called *$\mathcal{B}$-balanced* if it satisfies the $\mathcal{B}$-bilinearity conditions, that for all $b, b' \in \mathcal{B}$, $a_1, \ldots, a_n \in \mathcal{A}$, and for all $r = 1, \ldots, n-1$

$$\begin{aligned} f\left(ba_1, \ldots, a_n b'\right) &= bf\left(a_1, \ldots, a_n\right) b' \\ f\left(a_1, \ldots, a_r b, a_{r+1}, \ldots, a_n\right) &= f\left(a_1, \ldots, a_r, ba_{r+1} \ldots, a_n\right) \end{aligned}$$

A collection of $\mathcal{B}$-balanced maps $(f_\pi)_{\pi \in NC}$ is said to be multiplicative w.r.t. the lattice of non-crossing partitions if, for every $\pi \in \mathcal{NC}$, f_π is computed using the block structure of π in the following way:

- If $\pi = \hat{1}_n \in NC\,(n)$, we just write $f_n := f_\pi$.

- If $\hat{1}_n \neq \pi = \{V_1, \ldots, V_k\} \in \mathcal{NC}(n)$, then by a known characterization of NC, there exists a block $V_r = \{s+1, \ldots, s+l\}$ containing consecutive elements. For any such a block we must have

$$f_\pi(a_1, \ldots, a_n) = f_{\pi \backslash V_r}(a_1, \ldots, a_s f_l(a_{s+1}, \ldots, a_{s+l}), a_{s+l+1}, \ldots, a_n),$$

  where $\pi \backslash V_r \in NC(n-l)$ is the partition obtained from removing the block V_r.

We observe that a multiplicative family $(f_\pi)_{\pi \in \mathcal{NC}}$ is entirely determined by $(f_n)_{n \in \mathbb{N}}$. On the other hand, every collection $(f_n)_{n \in \mathbb{N}}$ of $\mathcal{B}$-balanced maps can be extended uniquely to a multiplicative family $(f_\pi)_{\pi \in NC}$.

The operator-valued free cumulants $\left(R_\pi^{\mathcal{B}}\right)_{\pi \in NC}$ are indirectly and inductively defined as the unique multiplicative family of $\mathcal{B}$-balanced maps satisfying the (operator-valued) moment-cumulant formulas

$$\mathbf{F}(a_1 \ldots a_n) = \sum_{\pi \in NC(n)} R_\pi^{\mathcal{B}}(a_1, \ldots, a_n)$$

By the cumulants of a tuple $a_1, \ldots, a_k \in \mathcal{A}$, we mean the collection of all cumulant maps

$$\begin{array}{rccc} R_{i_1,\ldots,i_n}^{\mathcal{B};a_1,\ldots,a_k} : & \mathcal{B}^{n-1} & \to & \mathcal{B}, \\ & (b_1, \ldots, b_{n-1}) & \mapsto & R_n^{\mathcal{B}}\left(a_{i_1}, b_1 a_{i_2}, \ldots, b_{i_{n-1}} a_{i_n}\right) \end{array}$$

for $n \in \mathbb{N}$, $1 \leq i_1, \ldots, i_n \leq k$.

A cumulant map $R_{i_1,\ldots,i_n}^{\mathcal{B};a_1,\ldots,a_k}$ is mixed if there exists $r < n$ such that $i_r \neq i_{r+1}$. The main feature of the operator-valued cumulants is that they characterize freeness with amalgamation [Spe98]: The random variables $a_1, \ldots, a_n$ are $\mathcal{B}$-free iff all their mixed cumulants vanish.

We recall two useful results from [NSS02]. The first proposition gives conditions for (operator-valued) cumulants to be restrictions of cumulants with respect to a larger algebra.

Proposition 5.2.1. *Let $1 \in \mathcal{D} \subset \mathcal{B} \subset \mathcal{A}$ be algebras such that $(\mathcal{A}, \mathbf{E})$ and $(\mathcal{B}, \mathbf{F})$ are respectively $\mathcal{B}$-valued and $\mathcal{D}$-valued probability spaces (and therefore $(\mathcal{A}, \mathbf{F} \circ \mathbf{E})$ is a $\mathcal{D}$-valued probability space) and let $a_1, \ldots, a_k \in \mathcal{A}$.*

Assume that the $\mathcal{B}$-cumulants of $a_1, \ldots, a_k \in \mathcal{A}$ satisfy

$$R_{i_1,\ldots,i_n}^{\mathcal{B};a_1,\ldots,a_k}(d_1, \ldots, d_{n-1}) \in \mathcal{D},$$

for all $n \in \mathbb{N}$, $1 \leq i_1, \ldots, i_n \leq k$, $d_1, \ldots, d_{n-1} \in \mathcal{D}$.

Then the $\mathcal{D}$-cumulants of $a_1, \ldots, a_k$ are just the restrictions of the $\mathcal{B}$-cumulants of $a_1, \ldots, a_k$, namely

$$R_{i_1,\ldots,i_n}^{\mathcal{B};a_1,\ldots,a_k}(d_1, \ldots, d_{n-1}) = R_{i_1,\ldots,i_n}^{\mathcal{D};a_1,\ldots,a_k}(d_1, \ldots, d_{n-1}),$$

for all $n \in \mathbb{N}$, $1 \leq i_1, \ldots, i_n \leq k$, $d_1, \ldots, d_{n-1} \in \mathcal{D}$.

The second proposition gives a characterization of operator-valued freeness by the agreement of operator-valued cumulants in different levels.

Proposition 5.2.2. *Let $1 \in \mathcal{D} \subset \mathcal{B}, \mathcal{N} \subset \mathcal{A}$ be algebras such that $(\mathcal{A}, \mathbf{E})$ and $(\mathcal{B}, \mathbf{F})$ are respectively $\mathcal{B}$-valued and $\mathcal{D}$-valued probability spaces. Then the first of the following statements implies the second.*

i) The subalgebras $\mathcal{B}, \mathcal{N}$ are free with amalgamation over $\mathcal{D}$.

ii) For every $k \in \mathbb{N}$, $n_1, \ldots, n_k \in \mathcal{N}$, $b_1, \ldots, b_{k-1} \in \mathcal{B}$, we have

$$R_n^{\mathcal{B}}(n_1 b_1, \ldots, n_{k-1} b_{k-1}, n_k) = R_n^{\mathcal{D}}(n_1 \mathbf{F}(b_1), \ldots, n_{k-1} \mathbf{F}(b_{k-1}), n_k).$$

Moreover, the two statements become equivalent if we add the faithfulness condition on $\mathbf{F} : \mathcal{B} \to \mathcal{D}$, that if $\mathbf{F}(b_1 b_2) = 0$ for all $b_2 \in \mathcal{B}$, then $b_1 = 0$.

5.3 Operator-valued convolutions via analytic subordination

Like in the scalar case, there are analytical tools to compute operator-valued free convolutions, which are based on the $\mathcal{B}$-valued Cauchy-transform

$$G_x^{\mathcal{B}}(b) = \mathbf{F}((b - x)^{-1}),$$

which maps the operatorial upper half-plane

$$\mathbb{H}^+(\mathcal{B}) := \{b \in \mathcal{B} : \exists \varepsilon > 0 \text{ such that } -i(b - b^*) \geq \varepsilon \cdot 1\}$$

into the lower half-plane $\mathbb{H}^-(\mathcal{B}) = -\mathbb{H}^+(\mathcal{B})$. In the usual settings coming from random matrix models (as we have seen above), our probability space $\mathcal{A}$ may have several operator-valued structures $\mathbf{F}_i : \mathcal{A} \to \mathcal{B}_i$ simultaneously, with $\mathbb{C} = \mathcal{B}_1 \subset \mathcal{B}_2 \subset \cdots \subset \mathcal{B}_k$, and $\mathbf{F}_i \circ \mathbf{F}_{i+1} = \mathbf{F}_i$. We are usually interested ultimately in the scalar-valued distribution, which can be obtained (via Stieltjes inversion) from the Cauchy-transform. The later in turn can be obtained from any "upper" $\mathcal{B}_i$-valued Cauchy transform, as we have that, for all $b \in \mathcal{B}_i$

$$\mathbf{F}_i(G_x^{\mathcal{B}_{i+1}}(b)) = \mathbf{F}_i \circ \mathbf{F}_{i+1}((b - x)^{-1}) = \mathbf{F}_i((b - x)^{-1}) = G_x^{\mathcal{B}_i}(b).$$

A drawback of the operator-valued setting is that, unless we ask $\mathcal{B}$ to be commutative, one can hardly compute explicit distributions: although $\mathcal{B}$-valued generalizations of the R and S-transforms exist ([Voi95], [Dyk06]), the task of explicitly inverting these operator-valued analytic maps is nearly impossible for any non-trivial situation (even for finite dimensional, relatively simple sub-algebras, like $\mathcal{B} = M_2(\mathbb{C})$).

In terms of moments, the operator-valued Cauchy transform is given by

$$G_x^{\mathcal{B}}(b) = \mathbf{F}((b - x)^{-1}) = \sum_{n \geq 0} \mathbf{F}(b^{-1}(xb^{-1})^n)$$

The operator-valued $\mathcal{R}$-transform is defined by

$$\mathcal{R}_x^{\mathcal{B}}(b) = \sum_{n\geq 1} R_n^{\mathcal{B}}(x, bx, \ldots, bx).$$

The vanishing of mixed cumulants for free variables implies the additivity of the cumulants, and thus also the additivity of the $\mathcal{R}$-transforms [Voi95]: If a_1 and a_2 are $\mathcal{B}$-free then we have for $b \in \mathcal{B}$ that $\mathcal{R}_{a_1+a_2}(b) = \mathcal{R}_{a_1}(b) + \mathcal{R}_{a_2}(b)$.

As in the scalar case, these transforms satisfy the functional equation

$$G_a^{\mathcal{B}}(b) = \left(\mathcal{R}_a^{\mathcal{B}}\left(G_a^{\mathcal{B}}(b)\right) - b\right)^{-1} \tag{5.8}$$

Rather than directly computing $\mathcal{B}$-valued R-transforms, a very powerful method to obtain $\mathcal{B}$-valued free convolutions is based on the analytic subordination phenomena observed by Biane ([Bia98], see also [Voi02]). In particular, the approach of [BB07] to obtain the subordination functions by iterating analytic maps can be very efficiently performed in the $\mathcal{B}$-valued context.

Theorem 5.3.1. *[BMS13] Let $(\mathcal{A}, \mathbf{F})$ be a $\mathcal{B}$-valued C^*-probability space and let $x, y \in \mathcal{A}$ be self-adjoint, $\mathcal{B}$-free. There exist an analytic map $\omega : \mathbb{H}^+(\mathcal{B}) \to \mathbb{H}^+(\mathcal{B})$ such that $G_x(\omega(b)) = G_{x+y}(b)$. Furthermore, for any $b \in \mathbb{H}^+(\mathcal{B})$ the subordination function $\omega(b)$ satisfies*

$$\omega(b) = \lim_{n\to\infty} f_b^{\circ n}(w),$$

where, for any $b, w \in \mathbb{H}^+(\mathcal{B})$, $f_b(w) = h_y(h_x(w) + b) + b$ and h is the auxiliary analytic self-map $h_x(b) = (E((b-x)^{-1}))^{-1} - b$ on $\mathbb{H}^+(\mathcal{B})$.

Numerically speaking, going from h_x to G_x and vice-versa is a simple operation. This means that one only needs the individual $\mathcal{B}$-valued Cauchy transforms of x, y (or good approximations of these) to obtain the $\mathcal{B}$-valued Cauchy transform of $x + y$, and hence, its probability distribution. The operator-valued multiplicative convolution can also be numerically approximated (see Section 5.5.2).

We will be interested in the particular setting where $(\mathcal{A}, \mathbf{F})$ is a rectangular probability space and hence our main space $(M_n(\mathbb{C}) \otimes \mathcal{A}, id \otimes \mathbf{F})$ consists of $n \times n$ matrices with entries in $\mathcal{A}$, endowed with the entry-wise evaluation of $\mathbf{F}$.

In Chapter 7 we will follow Anderson's self-adjoint linearization trick to obtain the distribution of a polynomial (such as $\Phi^{\square}$ from example 5.1.5) from the $M_n(\mathcal{D})$-distribution of an operator which depends linearly on the inputs of the polynomial. In the next section we show how to obtain the $M_n(\mathcal{D})$-valued Cauchy-transforms of such linear elements.

5.4 Linear elements

In a scalar-valued non-commutative probability space $(\mathcal{A}, \tau)$, we have the integral representation of the Cauchy-transform:

$$G_x(z) = \tau((z-x)^{-1}) = \int_{\mathbb{R}} (z-t)^{-1} d\mu_x(t).$$

Analogously, for linear, self-adjoint elements $c \otimes x$ in a $M_n(\mathbb{C})$-valued probability space $(M_n(\mathbb{C}) \otimes (\mathcal{A}), id_m \otimes \tau)$, we have:

$$G_{c\otimes x}(b) = (id_m \otimes \tau)((b - c \otimes x)^{-1}) = \int_{\mathbb{R}} (b - c \otimes t)^{-1} d\mu_x(t).$$

The previous integrals can be approximated, for example, by using matrix-valued Riemann sums. In particular, we are able to approximate the $M_n(\mathbb{C})$-valued Cauchy transform of any self-adjoint matrix which depends linearly on a semicircular element s. The same can be done if we start with a rectangular probability space.

Let $(\mathcal{A}, \mathbf{F})$ be a $\langle p_1, \dots p_k \rangle$-rectangular probability space and consider the $\mathcal{B}$-valued probability space $(M_m(\mathbb{C}) \otimes \mathcal{A}, \mathbf{F}_2)$, where $\mathbf{F}_2 = id_m \otimes \mathbf{F}$ and $\mathcal{B} = (M_m(\mathbb{C}) \otimes \langle p_1, \dots p_k \rangle)$

Consider $x \in \mathcal{A}$ of the form $x = \alpha_1 p_1 s_1 p_1 + \cdots + \alpha_k p_k s_k p_k$, where $s_i = p_i s_i p_i$ is a semicircular element when restricted to $\mathcal{A}^{(i)}$. Let $c \in M_m(\mathbb{C})$, and let $b = (d_{ij})_{i,j \leq m} \in \mathcal{B}$, with $d_{ij} = \beta_1^{ij} p_1 + \cdots + \beta_k^{ij} p_k$. Then we have

$$\begin{aligned} G^{\mathcal{B}}_{c\otimes x}(b) &= \mathbf{F}_2((b - c \otimes x)^{-1}) \\ &= (id_m \otimes \mathbf{F})((b - c \otimes x)^{-1}) \\ &= \frac{1}{2\pi} \int_{-2}^{2} [((\beta_1^{ij} - c_{ij}\alpha_1 t)p_1 + \cdots + (\beta_k^{ij} - c_{ij}\alpha_k t)p_k)_{ij}]^{-1} \sqrt{4 - t^2} dt. \end{aligned}$$

The case of deterministic matrices is simpler. If we assume that $M_n(\mathbb{C}) \subset \mathcal{A}$ and consider $x = x^* \in M_m(\mathbb{C}) \otimes M_n(\mathbb{C})$. Then $G_x^{\mathcal{B}}(b) = G_x^{\mathcal{B}}(b \otimes I_n)$ is just the partial trace $(id_m \otimes \mathbf{F}_2)((b \otimes I_n - x)^{-1})$.

One should be able to provide a similar trick to approximate Cauchy transforms for elements of the form $c \otimes u + c^* \otimes u^*$. For the moment, we find a way around this problem by removing Haar unitaries, as discussed in Section 3.3.

5.5 Operator valued free multiplicative convolution

In [BSTV14] we generalized to the operator-valued level the fixed point algorithm from [BB07] to compute free multiplicative convolutions. Let us first recall the eta transform (or Boolean cumulant transform)

$$\eta_x(b) = 1 - b(\mathbf{F}((b^{-1} - x)^{-1}))^{-1}.$$

Theorem 5.5.1. *Let $x > 0, y = y^* \in \mathcal{A}$ be two random variables with invertible expectations, free over $\mathcal{B}$. There exists a Fréchet holomorphic map $\omega_2 \colon \{b \in \mathcal{B} \colon \Im(bx) > 0\} \to \mathbb{H}^+(\mathcal{B})$, such that*

1. *$\eta_y(\omega_2(b)) = \eta_{xy}(b)$, $\Im(bx) > 0$;*

2. *$\omega_2(b)$ and $b^{-1}\omega_2(b)$ are analytic around zero;*

3. *For any $b \in \mathcal{B}$ so that $\Im(bx) > 0$, the map $g_b \colon \mathbb{H}^+(\mathcal{B}) \to \mathbb{H}^+(\mathcal{B})$, $g_b(w) = bh_x(h_y(w)b)$, where*

$$h_x(b) = b^{-1} - \mathbb{F}\left[(b^{-1} - x)^{-1}\right]^{-1}; \tag{5.9}$$

is well-defined, analytic and for any fixed $w \in \mathbb{H}^+(\mathcal{B})$,

$$\omega_2(b) = \lim_{n\to\infty} g_b^{\circ n}(w),$$

in the weak operator topology.

Moreover, if one defines $\omega_1(b) := h_y(\omega_2(b))b$, then

$$\eta_{xy}(b) = \omega_2(b)\eta_x(\omega_1(b))\omega_2(b)^{-1}, \quad \Im(bx) > 0.$$

The invertibility condition of $\mathbf{F}(Y)$ can be dropped if we restrict to finite dimensional algebras.

Proposition 5.5.2. *Let $\mathcal{B}$ be finite-dimensional and $x > 0$, $y = y^*$ free over B. There exists a function $g \colon \{b \in B \colon \Im(bx) > 0\} \times \mathbb{H}^+(B) \to \mathbb{H}^+(B)$ so that*

1. *$\omega_2(b) := \lim_{n\to\infty} g_b^{\circ n}(w)$ exists, does not depend on $w \in \mathbb{H}^+(B)$, and is analytic on $\{b \in B \colon \Im(bx) > 0\}$;*

2. $$\eta_y(\omega_2(b)) = \eta_{xy}(b), \quad b \in \{b \in B \colon \Im(bx) > 0\}.$$

Theorem 5.5.1 and Prop. 5.5.2 reduce the task of computing the distribution of the product of $\mathcal{B}$-free random variables x, y to the computation of the individual $\mathcal{B}$-valued Cauchy transforms of x and y, from which the h transforms can be easily derived. In [BSTV14] we used these theorems to compute the asymptotic distribution of certain matrix ensembles.

In particular, we treated the case of the product of shifted operator-valued semicirculars. In [BSTV14] we also solved some instance of the situation described in Section 3.5 of matrix ensembles of the form $a_1 b_1 a_1^* + \cdots + a_n b_n a_n^*$, where $\{a_1, \ldots, a_n\}$ and $\{b_1, \ldots, b_n\}$ are free.

This later case will not be included here since it can be also treated by the general method described in Chapter 7, which is based on [BMS13]. A drawback of our method to compute the multiplicative convolution is that one is in general not allowed to iterate it. While computing the free multiplicative convolution $\sqrt{x}y\sqrt{x}$ of x, y, we actually compute the transforms for the operator xy, which has the same scalar-valued moments as $\sqrt{x}y\sqrt{x}$. Since it is no longer true that $\sqrt{x}\sqrt{y}z\sqrt{y}\sqrt{x}$ and $\sqrt{x}yz\sqrt{x}$ have the same moments, we cannot proceed in this direction.

Another example which can be solved by both the multiplicative and additive convolutions is the case of block-linear transformations of a matrix ensemble which are used in quantum information theory to detect entanglement. Such ensembles were first considered by [Aub12], and further developed in [BN12a], [BN12b] and [ANV]. For this case, the multiplicative convolution seems not only to be the simplest way to get the distribution, but it also gives some hints on the conditions on the linear transformations which allow us to obtain explicit (non-numerical) descriptions of the distributions.

5.5.1 The product of two free operator-valued semi-circulars

Let s_1, s_2, s_3, and s_4 be free, semi-circular random variables, in some scalar-valued non-commutative probability space $(\mathcal{A}, \tau)$. Consider the matrices S_1 and S_2 defined by:

$$S_1 = \begin{pmatrix} s_1 & s_1 \\ s_1 & s_2 \end{pmatrix}, \; S_2 = \begin{pmatrix} s_3 + s_4 & 2s_4 \\ 2s_4 & s_3 - 3s_4 \end{pmatrix} \tag{5.10}$$

Since $s_1, \ldots, s_4$ are limits of independent Wigner matrices, the matrices $S_1, S_2 \in (M_2(\mathbb{C}) \otimes \mathcal{A}, id_2 \otimes \tau) := (\mathcal{M}, \mathbf{F})$ can be thought as limits of block Gaussian matrices.

We want to compute the spectral distribution of $(S_2 + cI_2)S_1$ in the scalar-valued probability space $(\mathcal{M}, \frac{1}{2}\mathrm{Tr} \otimes \tau)$, where c is some constant chosen large enough to make $S_2 + cI_2$ positive. By Prop. 5.1.7, the elements $(S_2 + cI_2), S_1$ are $M_2(\mathbb{C})$-free and hence we need only to compute the h-transforms h_{S_1} and h_{S_2}. We take care of the shift by observing that $G_{x+b_1}(b) = \mathbf{F}((b - (x + b_1))^{-1}) = G_x(b - b_1)$.

We can compute these h transforms numerically using the method described in [HFS07]. In brief, this involves expressing the Cauchy transform of the operator-valued semicircular in terms of the fixed point of a contraction mapping. Specifically, if we define

$$W(b) = \lim_{n \to \infty} \mathcal{F}_b^{\circ n}(W_0) \tag{5.11}$$

where $\mathcal{F}_b(W) = (-ib + \mathbf{F}[SbS])^{-1}$, then $G_S(b) = -iW(b)$.

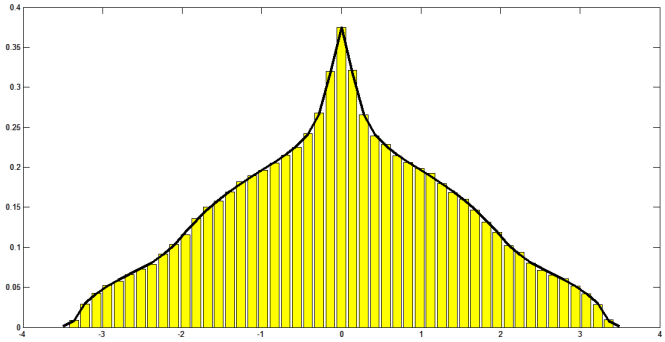

Fig. 5.2: Spectral distribution of S_1 random matrix simulations (histogram) compared with numerically calculated density, using fixed point method of [HFS07]

Note that we require the initial state W_0 to satisfy $\Im(W_0) > 0$; convergence of the above iteration scheme is ensured by arguments from [HFS07]. In our case, with $b = (b_{ij})_{i,j=1}^2$, we have

$$\mathbf{F}[S_1 b S_1] = \begin{pmatrix} b_{11} + b_{12} + b_{21} + b_{22} & b_{11} + b_{21} \\ b_{11} + b_{12} & b_{11} + b_{22} \end{pmatrix}$$

and

$$\mathbf{F}[S_2 b S_2] = \begin{pmatrix} 2b_{11} + 2b_{21} + 2b_{12} + 4b_{22} & 2b_{11} + -2b_{12} + 4b_{21} - 6b_{22} \\ 2b_{11} + 4b_{12} - 2b_{21} - 6b_{22} & 4b_{11} - 6b_{12} - 6b_{21} + 10b_{22} \end{pmatrix}$$

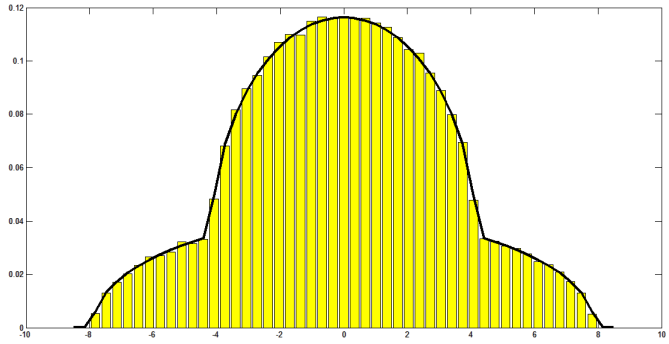

Fig. 5.3: Spectral distribution of S_2 random matrix simulations (histogram) compared with numerically calculated density, using fixed point method of [HFS07]

We compare the spectral distribution of S_1 and S_2 computed using this method and the Cauchy-Stieltjes inversion formula to random matrix simulations in Figures 5.2 and 5.3.

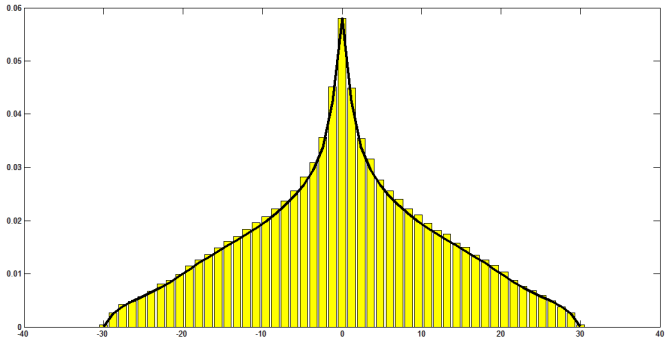

Fig. 5.4: Spectral distribution $(S_2 + 8.5I_2)S_1$ using our method.

Finally, using the numerically computed h transforms of S_1 and $S_2 + cI_2$ we used the iterative method discussed here to compute the h transform of their product. In Figure 5.4, we compare the distribution computed using our method to random matrix simulations of the ground truth spectral distribution of $(S_2 + cI_2)S_1$.

For the sake of variety, we consider another operator-valued semi-circular example. Let now $\{s_i\}_{i=1}^6$ be free semi-circular elements, and let:

$$S_1' = \begin{pmatrix} -10s_1 & 2s_2 & 30s_3 \\ 2s_2 & -4s_3 & 5s_1 \\ 30s_3 & 5s_1 & 16s_1 \end{pmatrix} \text{ and } S_2' = \begin{pmatrix} -2s_4 + 3s_6 & 3s_5 + 30s_6 & s_6 \\ 3s_5 + 30s_6 & s_4 + s_5 + s_6 & s_4 \\ s_6 & s_4 & 40s_4 \end{pmatrix} \tag{5.12}$$

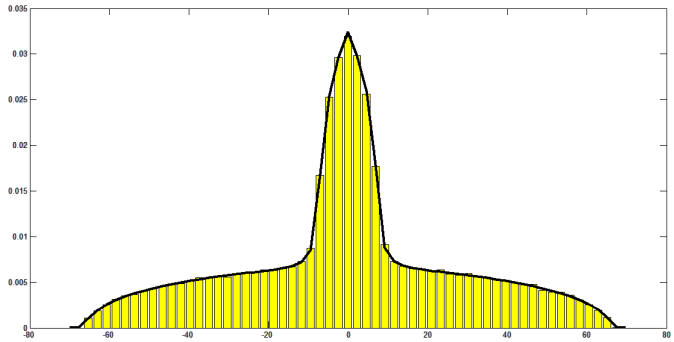

Fig. 5.5: Spectral distribution of S_1' random matrix simulations (histogram) compared with numerically calculated density using fixed point method of [HFS07].

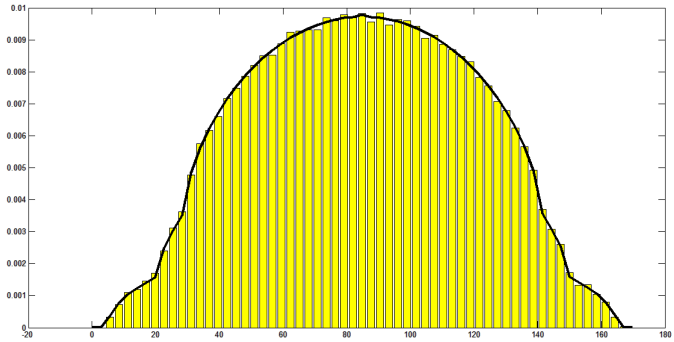

Fig. 5.6: Spectral distribution of S_2' - random matrix simulations (histogram) compared with numerically calculated density using fixed point method of [HFS07].

We follow the same pattern as previously: applying the numerical method proposed in [HFS07] to compute the individual h transforms of S_1' and S_2' (see Figure 5.6), and then using our iterative method to compute the spectral distributions of $(S_2'+85I_3)(S_1'+40I_3)$ and $(S_2'+85I_3)(S_1'+75I_3)$ (see Figure 5.8).

5.5.2 Block-linear transformations of random matrices

Now we show that the free multiplicative convolution can be used to compute the distribution of random matrices which have been deformed block-wise by a fixed self-adjoint linear transformation. Aubrun [Aub12] studied the case of the partial transposition, which is an example of a positive map which is not completely positive. These maps are relevant in quantum information theory to detect entanglement. In [BN12a, BN12b], Banica and Nechita computed the distributions of more general block-modified random matrices .

In [ANV] we continued these works. The new approach relies on operator-valued free

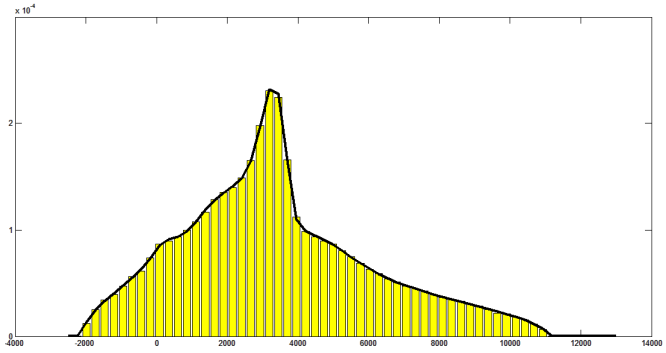

Fig. 5.7: Spectral distribution of $(S_2' + 85I_2)(S_1' + 40I_3)$ - random matrix simulations (histogram) compared with numerically calculated density using our method.

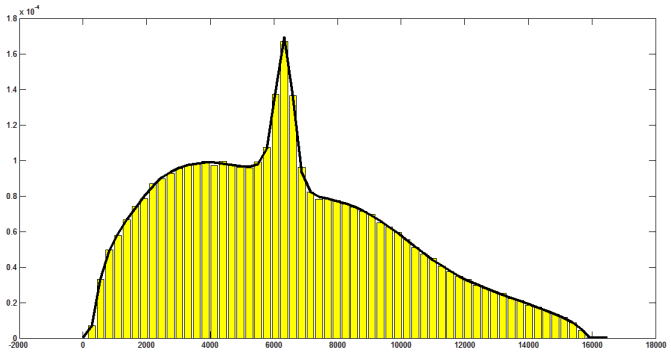

Fig. 5.8: Spectral distribution of $(S_2' + 85I_2)(S_1' + 75I_3)$ - random matrix simulations (histogram) compared with numerically calculated density using our method.

probability. In particular, we employed Theorem 5.5.1 to approximate the operator-valued free multiplicative convolution by iterating analytic maps, providing a general, numerical solution, which is explained in this section. In [ANV] we also found new classes of matrices for which we are able to write explicit limiting distributions.

Let W be a self-adjoint random matrix in a C^*-probability space $(\mathcal{A}_{dn}, \tau_{dn})$ of $dn \times dn$ random matrices, where τ_{dn} is the normalized expected trace, and let $\varphi : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$ be a linear map. We consider the block-modified random matrix

$$W^{\varphi} := (id_d \otimes \varphi)(W).$$

We want to understand the asymptotic eigenvalue distribution of W^{φ}

Let $\varphi : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$ be a self-adjoint linear map

$$\varphi(A) = \sum_{i,j,k,l=1}^{n} \alpha_{kl}^{ij} E_{ij} A E_{kl},$$

with $\alpha_{kl}^{ij} = \overline{\alpha_{ji}^{lk}} \in \mathbb{C}$.

The observation is that we may write

$$W^{\varphi} = \sum_{i,j,k,l=1}^{n} \alpha_{kl}^{ij} (I_d \otimes E_{ij}) W (I_d \otimes E_{kl}),$$

where $E_{ij} \in \mathcal{M}_n(\mathbb{C})$ are the matrix units.

Note that the collection $(I_d \otimes E_{ij})_{i,j=1}^n$ has (w.r.t. τ_{dn}) the same joint distribution as $(E_{ij})_{i,j=1}^n$ (w.r.t. τ_n) and then in the limit $d \to \infty$, $(I_d \otimes E_{ij})_{i,j=1}^n$ has a distribution. Thus, as $d \to \infty$, we can use Voiculescu's asymptotic freeness results between Wishart (or Wigner, or randomly rotated) matrices and deterministic matrices.

We will replace our deterministic matrices $(I_d \otimes E_{jl})_{i,j=1}^n$ by an abstract collection $(e_{ij})_{i,j \le n}$ of matrix units in a non-commutative probability space $(\mathcal{A}, \tau)$. The joint distribution of these matrix units is completely determined by the rules

$$e_{ij} e_{kl} = \delta_{jk} e_{il}, \quad \tau(e_{ij}) = n^{-1} \delta_{ij}, \quad \sum_{i=1}^{n} e_{ii} = 1, \quad e_{ij}^* = e_{ji}.$$

On the other hand, the random matrix W will be replaced by an element $w \in \mathcal{A}$ which is free from $(e_{ij})_{i,j \le n}$ and whose individual distribution is given by the asymptotic distribution of W as $d \to \infty$. Then the limiting distribution of W^{φ} is the same as the distribution of

$$w^{\varphi} := \sum_{i,j,k,l=1}^{n} \alpha_{kl}^{ij} e_{ij} w e_{kl}$$

Thus, the problem is reduced to study the distribution of such elements. In general, it is not true that the summands $e_{ij} w e_{kl}$ are free among themselves.

In order to construct an auxiliary self-adjoint matrix, let us consider the lexicographic order of two-letter words in an n-letter alphabet, and write, for $(i,j) < (l,k)$,

$$\begin{aligned} r e^{i\theta} e_{ij} w e_{kl} + r e^{-i\theta} e_{lk} w e_{ji} &= r^{1/2} (e^{i\theta/2} e_{ij} + e^{-i\theta/2} e_{lk}) w r^{1/2} (e^{-i\theta/2} e_{ji} + e^{i\theta/2} e_{kl}) \\ &\quad - r e_{ij} w e_{ji} - r e_{lk} w e_{kl} \\ &= (f_{kl}^{ij}) w (f_{kl}^{ij})^* - r e_{ij} w e_{ji} - r e_{lk} w e_{kl} \end{aligned}$$

so that we get

$$\begin{aligned} w^{\varphi} &= \sum_{\substack{1 \le i,j,k,l \le n \\ (i,j) < (l,k)}} (f_{kl}^{ij}) w (f_{kl}^{ij})^* + \sum_{1 \le i,j \le n} \beta_{ij} e_{ij} w e_{ji} \\ &= \sum_{\substack{1 \le i,j,k,l \le n \\ (i,j) < (l,k)}} (f_{kl}^{ij}) w (f_{kl}^{ij})^* + \sum_{1 \le i,j \le n} f_{ij} (-1)^{\varepsilon(i,j)} w f_{ij}^* \\ &= \sum_{\substack{1 \le i,j,k,l \le n \\ (i,j) \le (l,k)}} (f_{kl}^{ij}) \varepsilon_{kl}^{ij} w (f_{kl}^{ij})^*, \end{aligned}$$

where $\varepsilon_{ji}^{ij} = (-1)^{\varepsilon(i,j)}$ and $\varepsilon_{kl}^{ij} = 1$ for $(i,j) \neq (l,k)$.

From the elements f_{kl}^{ij}, $(i,j) \leq (l,k)$ we build a vector $f = (f_{11}^{11}, f_{12}^{11}, \ldots, f_{nn}^{nn})$ of size $m := n^2(n^2+1)/2$. We consider also the diagonal matrix $\Sigma = diag(\varepsilon_{11}^{11}, \varepsilon_{12}^{11}, \ldots, \varepsilon_{nn}^{nn})$ and we let $\tilde{w} := \Sigma \otimes w$.

We see that $f\tilde{w}f^* = w^{\varphi}$, so the desired distribution is the same (modulo a Dirac mass at zero of weight $1 - 1/m$) as the distribution of $f^*f\tilde{w}$ in the C^*-probability space $(M_m(\mathbb{C}) \otimes \mathcal{A}, tr_m \otimes \tau)$. Moreover, since w and each of the e_{kl}^{ij} are free, by Proposition 5.1.7, the matrices f^*f and $\tilde{w}$ are free with amalgamation over $M_m(\mathbb{C})$ (with respect to the conditional expectation $\mathbb{E} := id_m \otimes \tau$).

By Prop. 5.5.2, we can obtain the $M_m(\mathbb{C})$-valued Cauchy transform of $f\tilde{w}f^*$ numerically, provided that we can compute the $M_m(\mathbb{C})$-valued Cauchy transforms (or good approximations) of f^*f and $\tilde{w}$. The elements f^*f and $\tilde{w}$ fall in the cases which we can compute from Section 5.4.

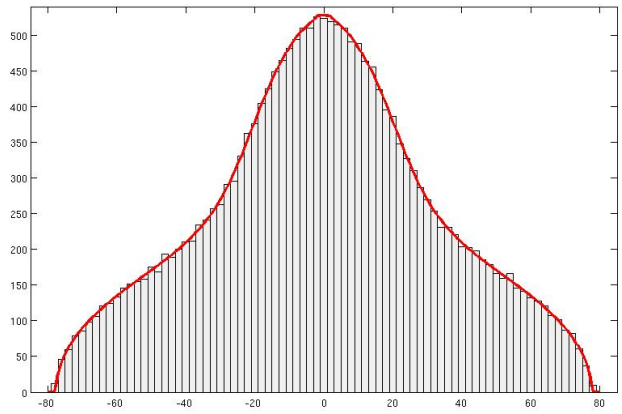

Fig. 5.9: Block-modified Wigner matrix

We compare the distributions obtained with our method and the empirical eigenvalue distributions of 20 realizations of 1000×1000 block-modified for the cases where w has a Wigner, Wishart and arcsine limiting distribution and the block transformation

$$\varphi((a_{ij})_{i,j\leq 2}) = \begin{pmatrix} 11a_{11} + 15a_{22} - 25a_{12} - 25a_{21} & 36a_{21} \\ 36a_{12} & 11a_{11} - 4a_{22} \end{pmatrix}$$

In [ANV] we also obtained conditions on φ which allow to express the asymptotic distribution of W^{φ} as the product of operators which are free with amalgamation over a commutative algebra. For such cases we are able to provide explicit distributions.

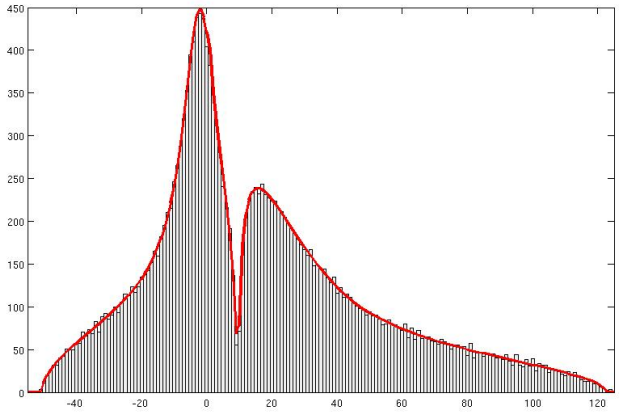

Fig. 5.10: Block-modified Wishart matrix

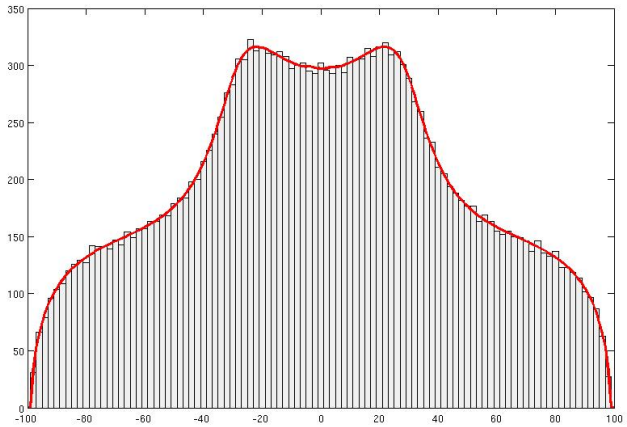

Fig. 5.11: Block-modification of a rotated arcsine matrix

6 (Free) deterministic equivalents

In the engineering literature the notion of a deterministic equivalent (apparently going back to Girko [Gir01], see also [HLN07]) has recently gained quite some interest. A deterministic equivalent is a simplification of the Cauchy transform G_P of the considered random matrix model (for which no analytic solution exists) by a function $\hat{G}_P$ which is defined as the solution of a specified system of equations, such that $\hat{G}_P$ and G_P are close if the matrices of the model are large. The specific form of those equations is determined in an ad-hoc way, depending on the considered problem, by making approximations for the equations of G_P, such that one gets a closed system of equations.

Most examples of deterministic equivalents are used to obtain approximations of Hermitian random matrix models involving Wigner, Wishart, Haar-unitary and deterministic matrices. For these cases, we showed in [SV12] that the Cauchy transform of our free deterministic equivalent $P^{\square}$ is the solution to the equations of the deterministic equivalents, i.e., that $\hat{G}_P = G_{P^{\square}}$ (see Section 6.2 for the definition of $P^{\square}$).

We want to point out that a first instance of these phenomena was essentially also observed in [NS95] in the context of the so-called CPA approximation (a kind of mean-field approximation) for the Anderson model from statistical physics. In our present language their result can be rephrased as saying that the free deterministic equivalent of the Anderson model is given by the CPA approximation.

We think that our definition of a deterministic equivalent gives a more conceptual approach (which can be easily extended to arbitrary polynomials) but more importantly, it gives an algebraic framework in which the individual and collective behavior of large random and deterministic matrices can be better organized.

From the free probabilistic nature of $P^{\square}$ we can easily associate an operator $(P_N)_{N\geq 1}$ which approaches $P^{\square}$ in AED distribution w.r.t. $(id_d \otimes \tau_N)$ and $(\tau_d \otimes \tau_N)$, reproducing, in some sense, the role that the partial sums $S_N = \frac{1}{\sqrt{N}}(X_1 + \cdots + X_N)$ of i.i.d. random variables play w.r.t. τ in the different central limit theorems (the corresponding Gaussian distribution being $P^{\square}$).

In addition, we will be able to set a fixed point equation whose unique solution is $G_{P^{\square}}$, for any polynomial matrix model.

In [CD11] the authors present also three examples of deterministic equivalents for non-polynomial models. Although the treatment that these examples require is slightly different, we will see that the solution to the deterministic equivalent for these models is again the Cauchy transform of an object which appears quite naturally in free probability. We introduce some of these examples here. A more detailed analysis is given in the last chapter.

In some sense this indicates that the only meaningful way to get closed systems of equations when dealing with certain self-adjoint random linear models (such as the ones

built up by deterministic matrices, zero mean i.i.d. random matrices or independent Haar unitary matrices) is to replace the model by the limit of some related model, whose distribution can be realized by a free-probabilistic operator.

6.1 Deterministic equivalents for Cauchy-transforms

A Wishart matrix is a random matrix $W_N = X_N X_N^*$, where X is a $N \times n$ random matrix with centered (complex) i.i.d. entries with variance $1/n$. If $N/n \to \lambda$, the AED of W_N converges to the Marchenko-Pastur law ν which is given by

$$\nu = \begin{cases} (1-\lambda)\delta_0 + \tilde{\nu}, & \text{if } 0 \leq \lambda \leq 1, \\ \tilde{\nu}, & \text{if } 1 < \lambda. \end{cases}$$

where $\tilde{\nu}$ is the measure supported on the interval $[(1-\sqrt{\lambda})^2, (1+\sqrt{\lambda})^2]$, with density

$$d\tilde{\nu}(t) = \frac{1}{2\pi t}\sqrt{4\lambda - (t-1-\lambda)^2}dt.$$

The Marchenko-Pastur law ν is the free law of small numbers and therefore it is often referred to as free Poisson.

Let us assume that $\lambda \geq 1$ (the other case is similar). We can think of W as $\frac{N}{n} Y_N P_n Y_N^*$, where Y_N is a non-self-adjoint Wigner matrix and P_n is a projection of trace n. By Voiculescu's asymptotic results, we know that $(W_N) \to (\lambda c P_\lambda c^*)$, where c is a circular operator which is free from a projection P_λ with $\tau(P_\lambda) = \lambda^{-1}$.

If we use the formula for cumulants with products as arguments (Theorem 4.3.2), it is quite simple to see that $R_n(cP_\lambda c^*, \ldots, cP_\lambda c^*) = \tau(P_\lambda) = \lambda^{-1}$ and hence

$$\mathcal{R}_{\lambda c P_\lambda c^*}(z) = \sum_{n \geq 1} \lambda^{n-1} z^{n-1} = \frac{1}{1-\lambda z}. \tag{6.1}$$

Therefore, we get the functional equation

$$\begin{aligned} G_\nu(z) = G_{\lambda c P_\lambda c^*}(z) &= (z - R_{\lambda c P_\lambda c^*}(G_{\lambda c P_\lambda c^*}(z)))^{-1} && (6.2) \\ &= (z - \frac{1}{1-\lambda G_{\lambda c P_\lambda c^*}(z)})^{-1}. && (6.3) \end{aligned}$$

Let us now consider some generalizations of the previous model:

First let $\Phi_{1,N} := X_N T_n X_N^*$, where X_N is as above and T_n is a deterministic matrix such that $(\mu_{T_n}) \to \mu_t$. Therefore, $(\Phi_{1,N}) \to \lambda ctc$, where c and t are free. We may proceed as

in the previous case and note that $\kappa_n(ctc^*,\dots,ctc^*) = \tau(t^n) = \lambda^{-1}\int_{\mathbb{R}} x^n d\mu_t(x)$. Hence

$$\begin{aligned}
\mathcal{R}_{\lambda ctc^*}(z) = \sum_{n\geq 1}\lambda^{n-1}z^{n-1}\int_{\mathbb{R}} x^n d\mu_t(x) &= (z\lambda)^{-2}G_t((z\lambda)^{-1}) - (z\lambda)^{-1} && (6.4)\\
&= (z\lambda)^{-1}\Big(\int_{\mathbb{R}}\frac{(z\lambda)^{-1}d\mu_t(x)}{x-(z\lambda)^{-1}} - 1\Big) && (6.5)\\
&= (z\lambda)^{-1}\Big(-\int_{\mathbb{R}}\frac{d\mu_t(x)}{1-xz\lambda} - 1\Big) && (6.6)\\
&= (z\lambda)^{-1}\Big(\int_{\mathbb{R}}\frac{xz\lambda d\mu_t(x)}{1-xz\lambda}\Big) && (6.7)\\
&= \int_{\mathbb{R}}\frac{x d\mu_t(x)}{1-xz\lambda}. && (6.8)
\end{aligned}$$

Therefore, the Cauchy transform of the limiting distribution $\mu_{\lambda ctc}$ satisfies the equation

$$\begin{aligned}
G_{\lambda ctc^*}(z) &= (z - R_{\lambda ctc^*}(G_{\lambda ctc^*}(z)))^{-1} && (6.9)\\
&= \Big(z - \int_{\mathbb{R}}\frac{x d\mu_t(x)}{1-x\lambda G_{\lambda cP_{\lambda c^*}}(z)}\Big)^{-1}. && (6.10)
\end{aligned}$$

For finite N we may consider the equation

$$G_N(z) = \Big(z - \int_{\mathbb{R}}\frac{x d\mu_{T_n}(x)}{1-xNn^{-1}G_N(z)}\Big)^{-1}. \tag{6.11}$$

From here it is quite intuitive that, as $N\to\infty$, the probability distribution μ_N with Cauchy transform G_N will approach that of G_{Φ_N} (since both coincide in the limit). In fact, we actually do not require T_n to have a limit in distribution (as long as some boundedness condition is imposed). The cumulative distribution functions of μ_N and $\mu_{\Phi_{1,N}}$ will satisfy that $F^{\mu_N} - F^{\mu_{\Phi_{1,N}}} \to 0$ even if the limit of $F^{\mu_{T_n}}$, and hence those of F^{μ_N} and $F^{\mu_{\Phi_{1,N}}}$ do not exist.

This possibility of considering non-converging sequences of deterministic matrices has given the impression that we are dealing with objects which will not satisfy asymptotic freeness and hence fall outside the realm of free probability.

We should note, however, that from the model $P = X_N T_n X_N^*$, we may define the ensemble $P_m = X_N^{(m)}(T_n\otimes I_m)(X_N^{(m)*})$, where $X_N^{(m)}$ is now an $Nm\times nm$ matrix with complex i.i.d. centered variables with variance $(nm)^{-1}$. We observe that the deterministic equivalent yields exactly the limiting distribution of the ensemble.

The model $P = \sum_{i=1}^k R_i X_i T_i X_i^* R_i^*$ (see Section 7.2) gives a further generalization of the previous one. In [CDS11], it was shown that a deterministic equivalent for the model is given as the solution of the equation

$$m_N(z) = \frac{1}{N}tr\Big(zI_N - \sum_{j=1}^k\int_{\mathbb{R}}\frac{x_j d\mu_{T_j}(x_j)}{1-x_j N n_j^{-1}e_j(z)}R_j R_j^*\Big)^{-1},$$

where

$$e_i = \frac{1}{N} tr R_i R_i^* (zI_N - \sum_{j=1}^{k} \int_{\mathbb{R}} \frac{x_j d\mu_{T_j}(x_j)}{1 - x_j N n_j^{-1} e_j(z)} R_j R_j^*)^{-1}.$$

The equations are now more involved, but it remains true that the deterministic equivalent G_N yields again the same distribution as the limit of the ensemble $P_m = \sum_{i=1}^{k} (R_i \otimes I_m) X_{i,m} (T_i \otimes I_m) X_{i,m}^* (R_i^* \otimes I_m)$.

This seems to be a general phenomenon: the deterministic equivalents for the 4 polynomial matrix models presented in [CD11] are actually limiting distributions of very concrete matrix ensembles which can be seen in the limit as polynomials on free operators.

Now let $W_N = X_N X_N^*$ where the $X_N = (x_{ij})_{i \leq N j \leq n}$ has complex, centered, independent entries with different variances σ_{ij}/n. Then

$$m_N(z) = \frac{1}{z} \frac{1}{N} \sum_{k=1}^{N} (z - \frac{1}{n} \sum_{i=1}^{n} \frac{\sigma_{ki}^2}{1 - e_{N,i}(z)})^{-1}, \tag{6.12}$$

where

$$e_{N,j}(z) = \frac{1}{n} \sum_{k=1}^{N} \sigma_{kj}^2 (z - \frac{1}{n} \sum_{i=1}^{n} \frac{\sigma_{ki}^2}{1 - e_{N,i}(z)})^{-1} \tag{6.13}$$

One can show that the deterministic equivalent for the first case corresponds to the Cauchy transform of operator-valued circulars element with variance profile, in the sense of Shlyakthenko [Shl96]. This means essentially replacing the complex i.i.d. entries by free circular entries.

However, since we cannot express the linear map which scales the variances $L(X_N) = \sum \sigma_{ij} P_i X P_j$ as a finite polynomial on X_N and deterministic matrices, we are not allowed in principle to use our asymptotic freeness toolbox. Hence, although the deterministic equivalent is a very free probabilistic object, the convergence does not follow directly from the results that we have been using.

However, by adjusting carefully the combinatorial arguments of the matrix-valued moment computations (as we already did when we moved from the scalar to the rectangular setting), one can usually justify (in some still quite ad-hoc way) the convergence results for this non-polynomial model (and, for example, the models described in Thm. 6.12 and 6.14 in [CD11]). We will no longer discuss the non-polynomial situation in this work.

6.2 Free deterministic equivalents

Definition 6.2.1. *Let*

$$P(x_1, \ldots, x_{n_1}, u_1, \ldots, u_{n_2}, d_1, \ldots, d_{n_3})$$

be a self-adjoint polynomial in non-commutative indeterminates $x_1, \ldots, x_{n_1}$ *,*$u_1, \ldots u_{n_2}$*,* $d_1, \ldots, d_{n_3}$ *and their adjoints.*

For some $N \geq 1$*, let* $P_0, \ldots, P_k \in M_N(\mathbb{C})$ *be pairwise orthogonal projections, such that* $P_0 + \cdots + P_k = I_N$*. Let* $X_1, \ldots, X_{n_1}$ *and* $U_1, \ldots, U_{n_2}$ *be independent random matrices, such that*

- *for each* $i \leq n_1$*,* $X_i = P_{j(i)} X_i P_{j(i)}$ *is a Wigner Matrix for some* $j(i) \leq k$ *(in the compressed space).*
- *for each* $i \leq n_2$*,* $U_i = P_{h(i)} U_i P_{h(i)}$ *is a Haar-unitary random matrix for some* $h(i) \leq k$ *(in the compressed space).*
- $D_1, \ldots, D_{n_3}$ *are deterministic matrices, with* $D_i = P_{h_1(i)} D_i P_{h_2(i)}$ *for some* $h_1(i), h_2(i) \leq n_3$
- $P(X_1, \ldots, X_{n_1}, U_1, \ldots, U_{n_2}, D_1, \ldots, D_{n_3}) =: P = P_0 P P_0$

The ***free deterministic equivalent*** $P^{\square}$ *of* P *is defined as*

$$P^{\square} = P(s_1, \ldots, s_{n_1}, u_1, \ldots, u_{n_2}, D_1 \ldots D_{n_3}),$$

where $s_1, \ldots, s_{n_1}, u_1, \ldots, u_{n_2}$ *are elements in some rectangular probability space* $(\mathcal{A}, \tau)$ *(with the same orthogonal projections* $P_0, \ldots, P_k \in M_N(\mathbb{C}) \subset \mathcal{A}$*) whose joint* $\mathcal{D}$*-distribution is determined by the following conditions:*

- *For all* $i \leq n_1$*,* $s_i = P_{j(i)} s_i P_{j(i)}$ *is a semicircular in the compressed space* $(\mathcal{A}^{j(i)}, \tau^{j(i)})$*.*
- *For all* $i \leq q$*,* $u_i = P_{j(i)} u_i P_{j(i)}$ *is a Haar-unitary in the compressed space* $(\mathcal{A}^{h(i)}, \tau^{j(i)})$*.*
- $s_1, \ldots, s_{n_1}, u_1, \ldots, u_{n_2}, \langle D_1, \ldots, D_{n_3} \rangle$ *are free with amalgamation over* $\langle P_0, \ldots, P_k \rangle$*.*

Observe that the joint distribution of $(s_1, \ldots, s_{n_1}, u_1, \ldots, u_{n_2}, D_1 \ldots D_{n_3})$ is exactly the asymptotic joint distribution of $(X_1^{(m)}, \ldots, X_{n_1}^{(m)}, U_1^{(m)}, \ldots, U_{n_2}^{(m)}, D_1 \otimes I_m, \ldots D_{n_3} \otimes I_m)$, where, for $m \geq 1$, $X_1^{(m)}, \ldots, X_{n_1}^{(m)}, U_1^{(m)}, \ldots, U_{n_2}^{(m)}$ are the corresponding independent blown-up Wigner matrices and Haar-distributed random unitary matrices. In particular, if $P_m = P(X_1^{(m)}, \ldots, X_{n_1}^{(m)}, U_1^{(m)}, \ldots, U_{n_2}^{(m)}, D_1 \otimes I_m \ldots D_{n_3} \otimes I_m)$, then $P_m \to P^{\square}$ as $m \to \infty$.

6.3 Simplified FDE's and correspondence to DE's

In Section 3.3 we saw that the distribution of $a + ubu^*$ where $\{a, b\}, \{u, u^*\}$ are free is the same as the distribution of $\tilde{a} + \tilde{b}$, where $a \sim \tilde{a}$ and $b \sim \tilde{b}$ and $\tilde{a}, \tilde{b}$ are free.

Hence we may, for example, reduce the problem of investigating the distribution of $a + ubu^*$ to that of $\tilde{a} + \tilde{b}$. In Chapter 7 we will be able to obtain the Cauchy transform of $P^{\square}$ from the matrix-valued transform of its linearizion, $L_{P^{\square}}$ which depends linearly on the variables of $P^{\square}$. The size of the linearizion and the number of free convolutions that we have to compute will depend on the complexity of the polynomial (in particular, on

the number of variables). It will be convenient then, to find a way to represent $P^{\square}$ so that we reduce the computational complexity of our algorithm to get its distribution.

We will present here an example of such simplifications. We show that the equations which determine the DE can be very easily derived as equations which determine the Cauchy transform of the corresponding FDE. These equations are derived, however, in an ad-hoc way. In Chapter 7 we present a general algorithm.

Consider again the model in Example 5.1.5. In this case we have that

$$P^{\square} = \sum_{i=1}^{k} R_i u_i T_i u_i^* R_i^*,$$

where $\langle R_1, \ldots, R_k, T_1, \ldots, T_k \rangle, u_1, \ldots, u_k$ are $\mathcal{D}$-free. Assume that the matrices are already embedded as in Example 5.1.5 and let $R = \sum_{i=1}^{k} R_i$, $U = \sum_{i=1}^{k} u_i$ and $T = \sum_{i=1}^{k} T_i$, then we have that $P^{\square} = RUTU^*R^*$. Moreover, the $\mathcal{D}$-distribution of (R, UTU^*) is the same as the $\mathcal{D}$-distribution of $(R, \tilde{T})$, where $R, \tilde{T}$ are $\mathcal{D}$-free and $(\tilde{T}) \overset{\mathcal{D}}{\sim} T$.

Proposition 6.3.1. *Let $(\mathcal{A}, \mathbf{F})$ be a $\langle p_1, \ldots, p_k \rangle$rectangular probability space. Let $A_1, A_2 \subset \mathcal{A}$ and $U = u_1 + u_2 + \cdots + u_k \in \mathcal{A}$ be such that $\langle D_1, D_2 \rangle, U$ are $\mathcal{D}$-free and $u_j = p_j u_j p_j$ is a Haar unitary in the compressed space $p_j \mathcal{A} p_j$. Then $D_1, U D_2 U^*$ are $\mathcal{D}$-free.*

Proof. Just replace τ by $\mathbf{F}$ in Prop. 3.3.1. □

Let $\mathbf{F}_2 : \mathcal{A} \to \langle R, \mathcal{D} \rangle := \mathcal{B}$ be the unique conditional expectation such that $\tau \circ \mathbf{F}_2 = \tau$. From the equations defining G, $\mathcal{R}$, (5.8) and in view of Prop. 5.2.2, we have that

$$\begin{aligned}
G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b) &= (b - \mathcal{R}^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b)))^{-1} && (6.14)\\
&= (b - R\mathcal{R}^{\langle R,\mathcal{D}\rangle}_{\tilde{T}}(R^* G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b) R) R^*)^{-1} && (6.15)\\
&= (b - R\mathcal{R}^{\mathcal{D}}_{\tilde{T}}(\mathbf{F}(R^* G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b) R)) R^*)^{-1} && (6.16)\\
&= (b - R\mathcal{R}^{\mathcal{D}}_{\tilde{T}}(\sum_{i=1}^{k} P_i \tau(P_i)^{-1} \tau(P_i R^* G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b) R)) R^*)^{-1} && (6.17)\\
&= (b - R\mathcal{R}^{\mathcal{D}}_{\tilde{T}}(\sum_{i=1}^{k} P_i \frac{M}{N_i} \frac{1}{M} \mathrm{Tr}(R_i R_i^* G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b))) R^*)^{-1}, && (6.18)
\end{aligned}$$

where we use at last that $\tau = \frac{1}{M}\mathrm{Tr}$ when restricted to $\langle R, \mathcal{D} \rangle$. Since $\tilde{t}_i = P_i u_i t_i u_i^*$ are orthogonal, we have that $\mathcal{R}^{\mathcal{D}}_{\tilde{T}}(b) = \sum_{i=1}^{k} P_i \mathcal{R}^{\mathcal{D}}_{\tilde{t}_i}(b) = \sum_{i=1}^{k} P_i \mathcal{R}^{\langle P_i \rangle}_{t_i}(b_i)$. This can be seen by functional calculus or, alternatively, by freeness over $\mathcal{D}$, which can be seen easily with cumulants, as the arguments are orthogonal. Hence

$$P_0 G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b)P_0 = P_0(b - R(\sum_{i=1}^{k} P_i \mathcal{R}^{\langle P_i\rangle}_{t_i}(\frac{1}{N_i}\mathrm{Tr}(R_i R_i^* G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b))))R^*)^{-1}P_0 \quad (6.19)$$

$$= P_0(b - \sum_{j=1}^{k} R_j R_j^* \mathcal{R}^{\langle P_i\rangle}_{t_j}(\frac{1}{N_j}\mathrm{Tr}(R_j R_j^* G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b))))^{-1}P_0 \quad (6.20)$$

$$= (b_0 - \sum_{j=1}^{k} R_j R_j^* \mathcal{R}^{\langle P_i\rangle}_{t_j}(\frac{1}{N_j}\mathrm{Tr}(R_j R_j^* G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b_0))))^{-1} \quad (6.21)$$

Note that $\tau^{(0)} = \frac{1}{N_0}\mathrm{Tr}$ when restricted to the compressed space $\langle R_1 R_1^*, \ldots, R_k R_k^*\rangle = P_0\langle R, \mathcal{D}\rangle P_0 \subset P_0 \mathcal{A} P_0$. Hence we get

$$G^{P_0\langle R,\mathcal{D}\rangle P_0}_{R\tilde{T}R^*}(zI_{N_0}) = (zI_{N_0} - \sum_{j=1}^{k} R_j R_j^* \mathcal{R}^{\langle P_j\rangle}_{t_j}(\frac{1}{N_j}\mathrm{Tr}(R_j R_j^* G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(zI_{N_0}))))^{-1} \quad (6.22)$$

If we define $f_j(z) := \frac{1}{N_i}\mathrm{Tr}(R_j R_j^* G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(zI_{N_0}))$, then our desired Cauchy-transform

$$G_{P^\square}(z) = \frac{1}{N_0}\mathrm{Tr}(G^{P_0\langle R,\mathcal{D}\rangle P_0}_{R\tilde{T}R^*}(zI_{N_0}))$$

satisfies the equation

$$G_{P^\square}(z) = \frac{1}{N_0}\mathrm{Tr}\left[\left(zI_{N_0} - \sum_{j=1}^{k} R_j R_j^* \mathcal{R}^{\langle P_j\rangle}_{t_j}(f_j(z))\right)^{-1}\right] \quad (6.23)$$

where

$$f_i(z) = \frac{1}{N_j}\mathrm{Tr}\left[R_i R_i^*\left(zI_{N_0} - \sum_{j=1}^{k} R_j R_j^* \mathcal{R}^{\langle P_j\rangle}_{t_j}(f_j(z))\right)^{-1}\right] \quad (6.24)$$

These equations are equivalent to the ones showing up in [CHD11] (since there they do not use the $\mathcal{R}$-transform of the matrices T_i, this information has to be encoded in another set of equations in their approach).

One should of course also consider the question whether those equations determine the $f_j(z)$ uniquely, within a suitably chosen class of functions. This questions is answered affirmatively for the present example in [CHD11]. In a general case, it is usually possible to find several solutions to the equations and the main problem is to single out the correct one. Sometimes this has been solved by putting constrains on the input matrices so that the convergence to right fixed point can be controlled (e.g. by asking some positivity condition). The ideal situation would be to avoid further constraining and simply find fixed point equations that leave some subset invariant and such that the unique solution on this subset is the right one.

In Chapter 7 we solve this general situation by extending the algorithm in [BMS13] to obtain equations which always determine $\mu_{P^\square}$ uniquely.

7 Linearization trick for FDE's

One of the main ingredients of our algorithm was already suggested by Voiculescu in his earlier papers on operator-valued free probability: the possibility to transfer questions about the distribution of a polynomial in non-commutative random variables to a question about the matrix-valued distribution of a related polynomial with matrix-valued coefficients but such that it is linear on the non-commutative variables.

The idea was formalized and put into practice by Haagerup and Thorbjornsen [HT05]. Some years later, Anderson [And11] found linearizations which preserve self-adjointness properties. In the next section we generalize Anderson's self-adjoint linearizion trick to be able to deal with operator-valued situations.

Our machinery to deal with matrix and rectangular distributions is very well behaved with respect to the different elements of the numerical algorithm, developed in [BMS13], to compute distributions of self-adjoint polynomials on free self-adjoint random variables. For this reason, we mainly point out those few steps where our situation differs.

In the last section, we give FDE's and suggest some numerically efficient linearizations for the models in [CHD11]. Before the examples, we first try to give a general overview on how the specific composition of the models influence the limiting models and the orders/types of convergence.

7.1 Linearization trick

Proposition 7.1.1. *Let $(\mathcal{A}, \mathcal{B})$ be a $\mathcal{B}$-probability space and let $x_1, \ldots, x_n \in \mathcal{A}$. Let $P = P(x_1, \ldots, x_n) \in \mathcal{B}\langle x_1, \ldots x_n, x_1^*, \ldots x_n^*\rangle$ be a self-adjoint $\mathcal{B}$-valued polynomial in $x_1, \ldots, x_n$ and their adjoints. There exist $m \geq 1$ and an element $L_P \in M_m(\mathbb{C}) \otimes \mathcal{A}$ such that:*

1. *$L_P = c_1 \otimes x_1 + c_1^* \otimes x_1^* + \ldots c_n \otimes x_n + c_n^* \otimes x_n^* + c \in M_m(\mathbb{C}) \otimes \mathcal{A}$, with $c \in M_m(\mathbb{C}) \otimes \mathcal{B}$ and, for $i \geq 1$ $c_i \in M_m(\mathbb{C})$.*

2. *If $\Lambda_\varepsilon(b)) = diag(b, i\varepsilon, i\varepsilon, \ldots, i\varepsilon) \in M_m(\mathbb{C}) \otimes \mathcal{B}$, then*

$$G_P^{\mathcal{B}}(b) = \lim_{\varepsilon \downarrow 0} (G_{L_P}^{M_m(\mathbb{C}) \otimes \mathcal{B})}(\Lambda_\varepsilon(b)))_{11}.$$

Proof. The main idea is to think of the polynomial $P \in \mathcal{B}\langle x_1, \ldots, x_n, x_1^*, \ldots, x_n^*\rangle$ as a polynomial $P \in \mathcal{B}\langle x_1, \ldots, x_n, x_{n+1}, \ldots, x_{2n}, b_1, \ldots, b_s\rangle$, where $x_{n+j} = x_j^*$ and the b_i's are the elements of $\mathcal{B}$ which appear as coefficients in P. With this, we are able to use [[BMS13], Prop. 3.2, Cor. 3.3 and Prop. 3.4].

Note that, by proceeding as in [[BMS13], Cor. 3.5], we will also get a self-adjoint linearizion

$$L_P = c_1 \otimes x_1 + \cdots + c_n \otimes x_n + d_1 \otimes x_1^* + \cdots + d_n \otimes x_n^* + e_1 \otimes b_1 + \cdots + e_s \otimes b_s + f.$$

The fact that $L_P = L_P^*$ will mean of course that $d_i = c_i^*$ and $c^* = c := e_1 \otimes b_1 + \cdots + e_s \otimes b_s + f$. So our linearizion has the desired form.

In view of [[BMS13], Cor. 3.3], one has again that $(b - P)^{-1} = [(\Lambda_0(b) - L_P)^{-1}]_{11}$ whenever $(b-P)$ (or, equivalently $\Lambda_0(b) - L_P$) is invertible. Hence, the linearizion works actually at the level of resolvents and the translation to Cauchy-transforms is obtained by applying $id_m \otimes \mathbf{F}$ to the resolvent of the right side (we must, however, consider $\Lambda_\varepsilon(b)$ as in [[BMS13], Cor. 3.6] so that the argument belongs to the operatorial upper-halfplane, which is the right domain of the Cauchy-transform for a later application of Theorem. 5.3.1). □

We include below the adaptations of [[BMS13], Prop 3.4 and Cor. 3.5] to our situation, which provide such linearizations.

Remark 7.1.2. *We recall one procedure to obtain a self-adjoint linearizion. A general monomial $p = b_0 x_{i_1} b_1 \cdots x_{i_k} b_k$ has a (possibly non-self-adjoint) linearizion*

$$L_P = \begin{bmatrix} & & & & b_0 \\ & & & x_{i_1} & -1 \\ & & b_1 & -1 & \\ & \cdots & \cdots & & \\ x_{i_k} & -1 & & & \\ b_k & -1 & & & \end{bmatrix}$$

If $p = p_1 + \cdots + p_k$ and each p_j has a linearizion

$$L_{p_j} = \begin{bmatrix} 0 & u_j \\ v_j & Q_j \end{bmatrix},$$

then a linearizion of p is given by

$$L_p = \begin{bmatrix} 0 & u_1 & \cdots & u_k \\ v_1 & & & \\ \vdots & & \ddots & \\ v_k & & & Q_k \end{bmatrix}.$$

Finally, if P is self-adjoint, we may view it as $P = q + q^*$ for $q = p/2$. If

$$L_q = \begin{bmatrix} 0 & u \\ v & Q \end{bmatrix}$$

is a linearizion of q then

$$L_P = \begin{bmatrix} 0 & u & v^* \\ u^* & 0 & Q^* \\ v & Q & 0 \end{bmatrix}$$

is a self-adjoint linearization of P.

Remark 7.1.3. *Since we are able to compute operator-valued Cauchy transforms of arbitrary deterministic matrices (as these are just partial traces), the products of deterministic matrices do not really bother us. We should use the linearization trick only to transform the polynomial into a polynomial with matrix coefficients which is linear in the variables which correspond to random matrices but needs not necessarily to be linear on the variables corresponding to deterministic matrices.*

7.2 Examples: random Matrix Models for wireless communication

In this Section we consider some matrix models and their deterministic equivalents from [CD11]. Understanding the nature of such matrix models was one of the main motivations of our work.

Before proceeding to the examples, we give a general idea of the weak and strong points of our method with respect to the original works.

In order to contrast both situations, we present later the models and results as they appear in [CD11] and then we study the models by our means.

7.2.1 General comparison to previous methods and results

Our free probabilistic machinery allowed us to define and understand deterministic equivalents for quite general polynomial matrix models. The models presented in this chapter will not seem very diverse in terms of the polynomial which presents the ensemble.

We must point out, however, that some of our results concerning these particular models may be slightly weaker in two possible senses: we may obtain weaker forms/rates of convergence, or we may have to assume more restrictive conditions on the matrices of the model.

At first glance, one of the most notorious weak points of our methods is that the matrices R_i's and T_i's that we assume to be deterministic may actually be random (although independent from the Wigner/Haar unitary matrices and asked to satisfy certain conditions, such as tightness or uniform boundedness) in the original works.

Randomness of the matrices R_i, T_i

We have seen that the FDE is an object which is well defined even if the deterministic matrices R_i, T_i (and hence the model itself) have no limit in distribution. In order to show that the model approaches its FDE one proceeds in the same way as when showing asymptotic freeness by computing mixed moments of Gaussian, Wigner, Haar matrices random matrices and deterministic matrices: The important thing to show is that the contributions to the moments which are indexed (in some way) by crossing partitions vanishes as N grows. The precise nature of the partitions that we will use depends on the type of matrices that build our models. In particular, the order of deviation from the average as $N \to \infty$ depends quite a lot on the specific shape of the polynomial.

For example, the quite recurrent shape $RXTX^*R^*$ of our polynomials forces only alternating moments to appear. This, and other similar considerations, such as inputing a specific i.i.d. distribution in the Wigner matrices, requiring R-diagonality conditions, Gaussianity, etc, accelerates the order of convergence of the considered models and hence has an influence on the properties of the analytical transforms that we may associate to the model (or to its limit).

There are generalizations of the asymptotic freeness results where one may replace the deterministic matrices by non-deterministic matrices with relatively small fluctuations on their distributions. In order to do so, we need some bound on the moments (or the norms) of the deterministic matrices. These bounds are obtained if, for example, we ask tightness on the spectral distributions of R_i, T_i, or even better, uniform boundedness. These conditions allow us to mimic the case where the deterministic matrices do converge in distribution.

If R_i, T_i are allowed to be random, the FDE yields a distribution modulo the inputs R_i's and T_i's. To obtain AED we simply have to replace the choices of R_i, T_i in some standard fixed point equation. Hence, the resulting distribution will depend on the specific realization of the R_i's and T_i's. The main idea is that, under these tightness or boundedness assumptions, one is able to give such a general formula, based on the deterministic case.

More technically speaking, instead of simply amalgamating over $M_N(\mathbb{C})$ (which would already contain R_i, T_i in the deterministic case) we will need to amalgamate over the algebra generated by $M_N(\mathbb{C})$ and the random matrices $(R_1, T_1 \dots R_k, T_k)$. If $(R_1, T_1 \dots R_k, T_k)$ and the Wigner/Haar $(X_1, \dots, X_k)$ matrices are independent, the conditional expectation onto the algebra generated by $(R_1, T_1 \dots R_k, T_k)$ is obtained by integrating out the terms corresponding to the $(x_1, \dots, x_k)$.

From the deterministic situations we know how to approximate a model

$$P_N(R_1, T_1, X_1 \dots R_k, T_k, X_k),$$

where $X_1, \dots, X_k$ are Haar-unitary or Wigner random matrices, by a deterministic, free probabilistic operator, $P_N^{\square}(R_1, T_1, x_1 \dots R_k, T_k, x_k)$. If we allow $R_1, T_1 \dots R_k, T_k$ to be random we simply approximate the model

$$P_N(R_1(\omega), T_1(\omega), X_1 \dots R_k(\omega), T_k(\omega), X_k),$$

whose randomness is only on $X_1, \dots, X_k$ by the deterministic operator

$$P_N^{\square}(R_1(\omega), T_1(\omega), x_1 \dots R_k(\omega), T_k(\omega), x_k).$$

Since the case where the matrices $(R_1, T_1, \dots R_k, T_k)$ are deterministic is sufficiently rich and illustrative, we will restrict our analysis to this case in the rest of the chapter.

Fixed point algorithm

Moving to another aspect, once that we have identified the DE as the Cauchy-Stieltjes transform of a free probabilistic operator, our general algorithm to compute the distribution seems to be much more direct and cleaner than the fixed point equations obtained

in the original works. In particular, the h transform that we use in Theorem 5.3.1 preserves half-planes and hence we no longer have to deal with the usually cumbersome procedure of selecting the right solution. At least in the earliest works, this problematics sometimes even led to assumptions (e.g. positivity, diagonality) on the deterministic matrices which were not really necessary but helped while tracking the right solution.

Orders and rates of convergence

The rates and types of convergence obtained in the original works seem to be nicer than the ones we achieve with our general methods. We want to investigate this situation more carefully.

The fact that we consider matrices of different sizes does not affect rates or types of convergence. We saw that the moments and cumulants of rectangular situations are just scaled by weights which are quite stable with N (due to the assumption on the stability of the ratios of the sizes of the matrices). Hence we may reduce our study to the case where all the matrices are of size $N \times N$.

Our main free probabilistic tool is the asymptotic freeness of deterministic matrices, Wigner matrices and Haar-unitary matrices. This result is quite general: It allows us to compute any asymptotic mixed moment on such matrices.

For example, if we have certain $N \times N$ random and deterministic matrices $X_1, \ldots, X_m$, we know how to compute any asymptotic moments

$$\lim_{N\to\infty} \tau_N(X_{i_1}^{\varepsilon_1} X_{i_2}^{\varepsilon_2} \cdots X_{i_k}^{\varepsilon_k}) = \tau(x_{i_1}^{\varepsilon_1} x_{i_2}^{\varepsilon_2} \cdots x_{i_k}^{\varepsilon_k}),$$

for all $k \geq 1$, $i = (i_1, \ldots, i_k) \in [m]^k$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_k) \in \{1, *\}^k$, where $(x_1, \ldots, x_m)$ are certain non-commutative random variables.

If the random matrices are Gaussian and/or Haar-unitary, we know that

$$\tau_N(i, \varepsilon) := \tau_N(X_{i_1}^{\varepsilon_1} X_{i_2}^{\varepsilon_2} \cdots X_{i_k}^{\varepsilon_k}) \to \tau(x_{i_1}^{\varepsilon_1} x_{i_2}^{\varepsilon_2} \cdots x_{i_k}^{\varepsilon_k}) =: \tau(i, \varepsilon)$$

almost surely (see [AGZ10] [MS12]).

The usual way to show almost sure convergence is by investigating, for each pair of tuples i, ε, the rate of convergence to 0 of the variances

$$\mathrm{Var}_N(i, \varepsilon) := \tau_N(ii^*, \varepsilon\varepsilon^*) - \tau_N(i, \varepsilon)\tau_N(i^*, \varepsilon^*),$$

where i^* is just the tuple i with the inverse order, ε^* switches the 1's by $*$'s (and vice-versa) and ii^*, $\varepsilon\varepsilon^*$ are just concatenations. In particular, for showing almost sure convergence, it is enough to show that

$$\sum_{N\geq 1} \mathrm{Var}_N(i, \varepsilon) < \infty.$$

There is a general theory in free probability for studying such fluctuations of mixed moments. We may, however, not be interested in computing *all* mixed moments, but only some classes of mixed moments which exhibit a better behavior.

In most of the examples below, the crucial randomness of the model comes from random matrices with independent complex, zero-mean random variables. Such random matrices are not self-adjoint, and hence we would need in principle to study all their $*$-moments. However, the shape of some of the models forces the random matrices to appear in an alternating way.

For our examples:

$$Q_N = \sum_{i \leq k} R_i X_i T_i T_i^* X_i^* R_i^*, \quad P_N := \sum_{i \leq k} R_i X_i T_i \sum_{j \leq k} T_j^* X_j^* R_j^*$$

we are actually interested, respectively, in the asymptotic joint $*$-distribution of

$$(R_1 X_1 T_1 T_1^* X_1^* R_1^*, \ldots, R_k X_k T_k T_k^* X_k^* R_k^*), \quad (R_1 X_1 T_1, T_1^* X_1^* R_1^*, \ldots, R_k X_k T_k, T_k^* X_k^* R_k^*),$$

and not in the much more general joint distribution $(X_1, R_1, T_1, \ldots, X_k, R_k, T_k)$.

Of course, our knowledge about the joint distribution of $(X_1, R_1, T_1, \ldots, X_k, R_k, T_k)$ is crucial and we use it to compute the joint distributions above, but we should restrict our estimates of the orders of contribution only to those partitions which will actually show up in our model.

For simplicity, let us first think about the case $k = 1$ (in this case, although P_N, Q_N are equal, the joint distributions afore mentioned are still not the same), where $X_1 = G$ is a complex Gaussian random matrix and R_N, T_N are deterministic matrices. We drop the N subindex from our notation.

We first look at the distributions of $(RGT, T^*G^*R^*)$ and its free deterministic equivalent $(RCT, T^*C^*R^*)$ where C is circular, free from $\{R, T\}$.

Since $NC(n) = \mathcal{P}(n)$ for $n \leq 3$, the first discrepancy between the $M_N(\mathbb{C})$-joint distributions of $(RGT, T^*G^*R^*)$ and $(RCT, T^*C^*R^*)$ happens when computing mixed moments of fourth order. Indeed, one can show that the $M_N(\mathbb{C})$-valued mixed moments

$$\mathbf{F}_N((RGT)(T^*G^*R^*)(T^*G^*R^*)(RGT)) - \mathbf{F}_N((RCT)(T^*C^*R^*)(T^*C^*R^*)(RCT)) \\ = N^{-2} RR^* RR^* T^* TT^* T,$$

and similarly

$$\mathbf{F}_N((T^*G^*R^*)(RGT)(RGT)(T^*G^*R^*)) - \mathbf{F}_N((T^*C^*R^*)(RCT)(RCT)(T^*C^*R^*)) \\ = N^{-2} T^* TT^* TRR^* RR^*.$$

At least for the Gaussian case (a condition which is also assumed in example 7.3.3 below), one can perform similar estimates in general and hence, under the assumption of uniform boundedness of $\|T_N\|$ and $\|R_N\|$, almost sure convergence can be achieved.

In fact the contribution of the partition $\{(1,3)(2,4)\}$ above calculated only shows up if we consider the model

$$\sum_{1 \leq i \leq k} R_i X_i T_i + T_i^* X_i^* R_i^*$$

which depends on more general moments than the model

$$P_N = \sum_{i \leq k} R_i X_i T_i \sum_{j \leq k} T_j^* X_j^* R_j^*.$$

Now let us move to the better behaved model

$$Q_N = RGTT^*G^*R^*, \quad Q_N^{\square} = RCTT^*C^*R^*.$$

The first non-vanishing fluctuation of Q_N w.r.t. $Q_N^{\square}$ appears only when computing the third order moment (which is of order six on G and C). Indeed, observe that the partition $\{(1,3),(2,4)\}$ no longer contributes in the second moment since G and G^* are forced to appear in an alternating order. The first crossing pairing which respects the alternating pattern $(1,*,1,*,\ldots,1,*)$ is the pairing $\{(1,4),(2,5),(3,6)\}$. We have that $\mathbb{E}_N(Q_N - Q_N^{\square})^k = 0$ for all $k \leq 2$ and

$$\mathbb{E}_N(Q_N - Q_N^{\square})^3 = N^{-3} RR^*RTT^*TT^*TT^*R^*RR^*.$$

This gives some hints on why almost sure convergence is achieved for the example 7.3.2, even while assuming weaker conditions on the matrices that build the model.

7.3 Examples from wireless communications

We now introduce the matrix models as they appear in [CD11]. After each model, we suggest embeddings of the matrices in rectangular spaces. We then present the FDE and we re-derive equations of the DE's to show that they correspond to the same distributions. Finally, we give the linearization which allows to plot the distribution.

We should point out that in free probability one usually works with the Cauchy-transform G_x and not with the Stieltjes-transform $-G_x$, as one usually does in random matrix theory.

7.3.1 Unitary precoded channels [CHD11]

Let $k \geq 1$ be fixed. For each N and $i \leq k$, consider a non-negative $n_i \times n_i$ Hermitian deterministic matrix T_i with uniformly bounded spectral norm along n_i. Let U_i be the $n_i \leq N_i$ columns of a Haar-distributed random matrix in $\mathcal{U}(N_i)$. Let R_i be an $N \times N_i$ random matrix with uniformly bounded spectral norm along N, almost surely. Define $c_i = \frac{n_i}{N_i}$, and $\bar{c}_i = \frac{N_i}{N}$ and denote

$$P_N = \sum_{i=1}^{k} R_i U_i T_i U_i^* R_i^*,$$

Then as $N, N_1, \ldots N_k, n_1, \ldots, n_k$ grow to infinity with ratios $\bar{c}_i$ satisfying $0 < \liminf \bar{c}_i \leq \limsup \bar{c}_i < \infty$ and $0 \leq c_i \leq 1$ for all i, the following limit holds true almost surely:

$$F_{P_N} - F_N \to 0,$$

where F_{P_N} and F_N are, respectively, the cumulative distribution functions of μ_{P_n} and the cumulative distribution function of the probability measure with Stieltjes transform $m_N(z)$ defined by

$$m_N(z) = \frac{1}{N}\mathrm{Tr}\left(\sum_{i\leq k} \bar{e}_i(z) R_i R_i^* - zI_N\right)^{-1},$$

where $(\bar{e}_1, \ldots \bar{e}_k)$ is the unique solution to the system of equations

$$\bar{e}_i(z) = \frac{1}{N}\mathrm{Tr}(T_i(e_i(z)T_i + [\bar{c}_i - e_i(z)\bar{e}_i(z)]I_{n_i})^{-1}),$$

where the

$$e_i(z) = \frac{1}{N}\mathrm{Tr}R_i R_i^* \left(\sum_{j\leq k} \bar{e}_j(z) R_j R_j^* - zI_N\right)^{-1}$$

are constrained to be Stieltjes transforms of non-negative probability measures and, for z real negative, $0 \leq e_i(z) < c_i\bar{c}_i/\bar{e}_i(z)$ for all i. Moreover, for each real negative z,

$$\bar{e}_i(z) = \lim_{t\to\infty} \bar{e}_i^{(t)}(z),$$

where $\bar{e}_i^{(t)}(z)$ is the unique solution of

$$\bar{e}_i^{(t)}(z) = \frac{1}{N}\mathrm{Tr}(T_i(e_i^{(t)}(z)T_i + [\bar{c}_i - e_i^{(t)}(z)\bar{e}_i^{(t)}(z)]I_{n_i})^{-1}),$$

within the interval $[0, c_i\bar{c}_i/e_i^{(t)}(z))$, $e_i^{(0)}(z)$ can take any positive value and $e_i^{(t)}$ is recursively defined by

$$e_i^{(t)}(z) = \frac{1}{N}\mathrm{Tr}R_i R_i^* \left(\sum_{j\leq k} \bar{e}_j(z)^{(t-1)} R_j R_j^* - zI_N\right)^{-1}.$$

FDE

Since we defined the FDE in principle only for cases with square Haar or Wigner Matrices, we should simply think that the U_i's are $N_i \times N_i$ and that we complete T_i with zeros to make an $N_i \times N_i$ matrix.

We embed the matrices as in Fig. 7.1.

The embedding is explained in detail in Example 5.1.5. The correspondence between the Cauchy transform of the FDE

$$P^{\square} = \sum_{i=1}^{k} R_i u_i T_i u_i^* R_i^*$$

and the DE was shown in Section 6.3.

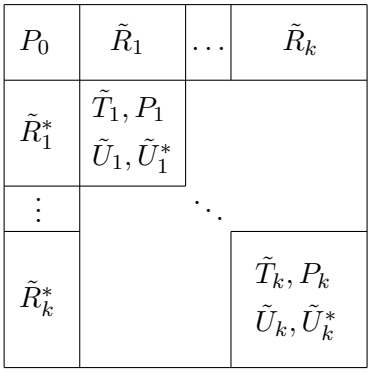

Fig. 7.1: Embedding of the matrices of P on a Rectangular Space

Linearization

For this case, the linearization is very simple. By Prop. 6.3.1, we may replace the standard FDE $P^{\square} = RuTu^*R^*$ by simply $P^{\square} = R\tilde{T}R^*$, where $\{R, R^*\}, T$ are $(P_0, \ldots, P_k) =: \mathcal{D}$-free. Hence we get

$$L_{P^{\square}} = \begin{bmatrix} 0 & 0 & R \\ 0 & T & -1 \\ R^* & -1 & 0 \end{bmatrix},$$

where each entry is really an $M \times M$ block, with $M = N_0 + \cdots + N_k$. The individual $M_3(\mathcal{D})$-valued Cauchy transforms of the self-adjoint elements

$$L_1 = \begin{bmatrix} 0 & 0 & R \\ 0 & 0 & -1 \\ R^* & -1 & 0 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & T & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

can be computed by performing partial traces, as explained in Section 5.4.

In Fig. 7.2 we compute the distribution of $P^{\square}$ for

$$P_N = Q_N U_N R_N U_N^* Q_N^* + S_N V_N T_N V_N^* S_N^*,$$

where Q_N, S_N, R_N, T_N are the blown-ups of some arbitrary deterministic matrices $Q_1, S_1, R_1 = R_1^*, T_1 = T_1^*$ of sizes 5×8, 5×4, 8×8 and 4×4, respectively, and $U_N \in \mathcal{U}(8N)$, $V_N \in \mathcal{U}(4N)$ are unitary matrices chosen independently with uniform distribution on the compact unitary groups $\mathcal{U}(8N)$ and $\mathcal{U}(4N)$.

7.3.2 Correlated MIMO multiple access channels [CDS11]

Let $k \geq 1$ be fixed and consider for each N a model

$$B_N = \sum_{i=1}^{k} R_i^{1/2} X_i T_i X_i^* R_i^{1/2} + A,$$

where

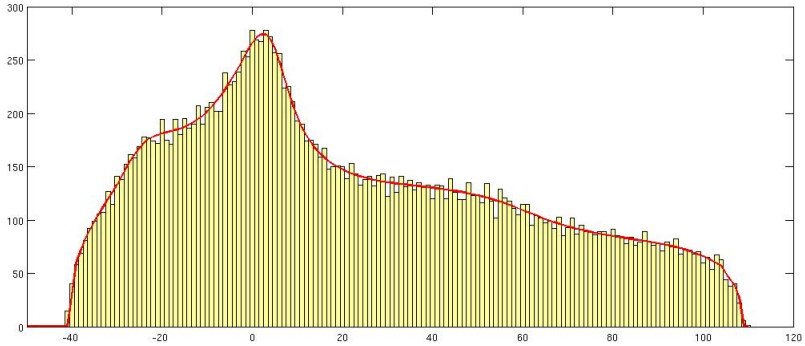

Fig. 7.2: 20000 eigenvalues (from 100 realizations) of P_{40} vs $\mu_{P^{\square}}$ (red).

- $X_i = (\frac{1}{\sqrt{n_i}} X_{i,rs})$ is an $N \times n_i$ random matrix with identically distributed entries $X_{i,rs}$, independent for each fixed n and such that $\mathbb{E}|X_{i,rs} - \mathbb{E}X_{i,rs}|^2 = 1$.
- $R_i^{1/2}$ is the Hermitian square root of the $N \times N$ matrix R_i,
- T_i is a non-negative $n_k \times n_k$ diagonal matrix
- The eigenvalue distributions of $R_1, \ldots, R_k$, $T_1, \ldots, T_k$ are tight: For each $\varepsilon > 0$ there exist $M > 0$ such that of $\mu_Y[M, \infty) < \varepsilon$ for all N and all $Y \in \{T_1, \ldots T_k, R_1, \ldots R_k\}$.
- A is $N \times N$ Hermitian, non-negative definite.
- If $c_i = N/n_i$, all $i \leq k$ there exist $0 < a < b < \infty$ such that

$$a \leq \liminf_N c_i \leq \limsup_N c_i \leq b$$

Then the Stieltjes transform $m_{B_N}(z)$ satisfies

$$m_{B_N}(z) - m_N(z) \to 0 \tag{7.1}$$

almost surely, where

$$m_N(z) = \frac{1}{N} \mathrm{Tr} \left(A + \sum_{i=1}^{k} \int \frac{t_i d\mu_{T_i}(t_i)}{1 + c_i t_i e_{N,i}(z)} R_i - z I_N \right)^{-1},$$

and $e_{N,1}(z), \ldots, e_{N,k}(z)$ are the solution to the system of equations

$$e_{N,j}(z) = \frac{1}{N} \mathrm{Tr} R_j \left(A + \sum_{i=1}^{k} \int \frac{t_i d\mu_{T_i}(t_i)}{1 + c_i t_i e_{N,i}(z)} R_i - z I_N \right)^{-1}.$$

Moreover, for any $\varepsilon > 0$, the convergence of equation (7.1) is uniform over any region of $\mathbb{C}$ bounded by a contour interior to

$$\mathbb{C} \setminus (\{z : |z| \leq \varepsilon\} \cup \{z = x + iv : x > 0, |v| \leq \epsilon\}).$$

For all N, the function m_N is the Stieltjes transform of a distribution function F_N, and

$$F^{B_N} - F^N \to 0$$

almost surely as $N \to \infty$.

FDE

In view of our discussion in the first section, we only consider here the FDE of the case where the X_i are centered non-self-adjoint Wigner matrices and the A, R_i, T_i are deterministic. We also complete the deterministic matrices with zeros in such a way that all the X_i can be thought as square matrices and we re-normalize (for example, by scaling the $R_i's$), so that all random matrices X_i have entries with the same variance.

From the free probabilistic point of view this model can be treated exactly as the previous model. We assume that $N \geq n_i$ (all other cases are similar). With the notation of the previous example, we make the embeddings in such a way that $P_0\tilde{R}_iP_i = \tilde{R}_i$, $P_i\tilde{X}_iP_i = \tilde{X}_i$, $P_i\tilde{T}_iP_i = \tilde{T}_i$, where X_i and T_i are also thought as $N \times N$ matrices, where X_i is a full matrix and only the upper $n_i \times n_i$ corner of T_i is nonzero. Hence, we are using the same rectangular space as in the previous case, but this time the projections $P_0, \ldots, P_k$ have the same size.

By the discussion in Section 6.1 the random matrix $P_i\tilde{X}_i\tilde{T}_i\tilde{X}_i^*P_i$ has, in the compressed space $\mathcal{A}^{(i)}$, the distribution of a free compound Poisson. Hence our FDE will be just

$$P^{\square} = \sum_{i=1}^{k} R_i\lambda_i c_i T_i c_i^* R_i^*.$$

The crucial fact, briefly discussed in Section 3.3, is that, similar to the case of conjugation by a free Haar-unitary, the conjugation of the T_i's by circular elements separates them from the R_i's. The only difference here is that, unlike $u_iT_iu_i^*$, the conjugation $c_iT_ic_i^*$ no longer preserves the distribution of T_i, but we still know how to compute it.

This means that all the arguments about operator-valued freeness that we used from

$$(\tilde{T}_1, \ldots \tilde{T}_k, R_1, \ldots, R_k),$$

to bring $G^{\langle R,\mathcal{D}\rangle}_{R\tilde{T}R^*}(b)$ to its final form (described by Equations (7.2) and (7.3)) hold also for

$$(\lambda_i c_1 T_1 c_1^*, \ldots \lambda_i c_k T_k c_k^*, R_1, \ldots, R_k).$$

Hence we have that

$$G_{P^{\square}}(z) = \frac{1}{N_0}\mathrm{Tr}\left[\left(zI_{N_0} - \sum_{j=1}^{k} R_j R_j^* \mathcal{R}^{\langle P_j\rangle}_{\lambda_j c_j T_j c_j^*}(f_j(z))\right)^{-1}\right] \tag{7.2}$$

where

$$f_i(z) = \frac{1}{N_j}\mathrm{Tr}\left[R_i R_i^*\left(zI_{N_0} - \sum_{j=1}^{k} R_j R_j^* \mathcal{R}^{\langle P_j\rangle}_{\lambda_j c_j T_j c_j}(f_j(z))\right)^{-1}\right]. \tag{7.3}$$

As we observed already in Section 6.1 the $\mathcal{R}$-transform of the free compound Poisson has an explicit integral representation

$$\mathcal{R}^{\langle P_j\rangle}_{\lambda_j c_j T_j c_j}(z) = \int_{\mathbb{R}} \frac{x d\mu_{T_i}(x)}{1 - xz\lambda},$$

which explains the shape of the equation.

Linearization

We know now that the Cauchy-transform of the FDE $P^{\square} = RcTc^*R^*$, where $\{R, R^*, T\}, c$ are free coincides with the solution of the equations of the DE. The linearization yields

$$L_{P^{\square}} = \begin{bmatrix} 0 & 0 & 0 & 0 & R \\ 0 & 0 & 0 & c & -1 \\ 0 & 0 & T & -1 & 0 \\ 0 & c^* & -1 & 0 & 0 \\ R^* & -1 & 0 & 0 & 0 \end{bmatrix},$$

where each entry is really $M \times M$, with $M = N_0 + \cdots + N_k$. The individual $M_5(\mathcal{D})$-valued Cauchy-transforms of the self-adjoint elements

$$L_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & R \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & T & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ R^* & -1 & 0 & 0 & 0 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & c^* & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

can be computed by, respectively, performing a partial trace and approximating by matrix-valued Riemann sums (or by using the method in [HFS07]), as explained in Section 5.4.

7.3.3 Frequency selective MIMO systems [DL07]

Let $k \geq 1$ be fixed and consider the model

$$P_N = \sum_{i=1}^{k}(R_i X_i T_i) \sum_{j=1}^{k}(T_j^* X_j^* R_j^*),$$

where the notations are the same as in the previous model, with the additional assumptions that $n = n_1 = \cdots = n_k$. The random matrix X_k has independent Gaussian entries and the spectral norms $\|T_i\|$ and $\|R_i\|$ are uniformly bounded with N. The assumption of T_i being diagonal can be removed.

Then we have

$$N(\tau(m_{B_N}(z)) - m_N(z)) = O(1/N),$$

with m_N defined, for $z \in \mathbb{C} \setminus \mathbb{R}^+$, as

$$m_N(z) = \frac{1}{N}\mathrm{Tr}\left(-z\left[I_N + \sum_{i\leq k} \bar{e}_{N,i}(z) R_i R_i^*\right]\right)^{-1}, \tag{7.4}$$

where $(\bar{e}_{N,1}, \ldots \bar{e}_{N,k})$ is the unique solution of

$$e_{N,i}(z) = \frac{1}{N} R_i R_i^* \mathrm{Tr}\left(-z\left[I_N + \sum_{j\leq k} \bar{e}_{N,j}(z) R_j R_j^*\right]\right)^{-1}, \tag{7.5}$$

$$\bar{e}_{N,i}(z) = \frac{1}{N} T_i T_i^* \mathrm{Tr}\left(-z\left[I_N + \sum_{j\leq k} e_{N,j}(z) T_j T_j^*\right]\right)^{-1}, \tag{7.6}$$

all with positive imaginary part if $z \in \mathbb{C}^+$, negative imaginary part if $z \in \mathbb{C}^-$ and positive if $z \in \mathbb{R}^-$

FDE

We make again $n = N$ by completing either the R_i's or the T_i's with zeros, so that we may consider square Gaussian matrices X_i. We embed again R_i, X_i in such a way that $P_i \tilde{R}_i P_0 = \tilde{R}_i$ and $P_i \tilde{X}_i P_i = \tilde{X}_i$ but this time we put $P_0 \tilde{T}_i P_i = \tilde{T}_i$.

If we put again $R = \sum \tilde{R}_i$, $T = \sum \tilde{T}_i$, $c = \sum c_i$, our FDE can be compactly written as

$$P_N^{\boxplus} = RcTT^*c^*R^*.$$

By proceeding as in the previous examples, it is not hard to bring the operator-valued Cauchy-transform of $P_N^{\boxplus}$ into implicit equations such as Eq. (7.4).

Linearization

The linearization will be very similar to the one in the previous case, the main difference is the way in which we have embedded the matrices T_i. We get

$$L_{P^{\boxplus}} = \begin{bmatrix} 0 & 0 & 0 & 0 & R \\ 0 & 0 & 0 & c & -1 \\ 0 & 0 & TT^* & -1 & 0 \\ 0 & c^* & -1 & 0 & 0 \\ R^* & -1 & 0 & 0 & 0 \end{bmatrix},$$

where each entry is $M \times M$, with $M = (k+1)\max\{N, n\}$. The individual $M_5(\mathcal{D})$-valued Cauchy-transforms of the self-adjoint elements

$$L_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & R \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & TT^* & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ R^* & -1 & 0 & 0 & 0 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & c^* & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

can be computed as explained in Section 5.4.

Bibliography

[AGZ10] Greg Anderson, Alice Guionnet, and Ofer Zeituni, *An Introduction to Random Matrices*, Cambridge University Press, Cambridge, 2010.

[AH13] Octavio Arizmendi and Takahiro Hasebe, *Semigroups related to Boolean and free convolutions.*, Studia Math. **215** (2013), no. 2, 157–185.

[AHLV14] Octavio Arizmendi, Takahiro Hasebe, Franz Lehner, and Carlos Vargas, *Relations between cumulants in noncommutative probability (pre-print)*, arXiv:1408.2977 (2014).

[AJ14] Octavio Arizmendi and Arturo Jaramillo, *Convergence of the fourth moment and infinite divisibility: quantitative estimates*, Elec. Comm. in Prob. **19** (2014), 1–12.

[And11] Greg Anderson, *Convergence of the largest singular value of a polynomial in independent Wigner matrices*, Preprint, arXiv:1103.4825v2 (2011).

[ANV] Octavio Arizmendi, Ion Nechita, and Carlos Vargas, *Block-linear transformations of random matrices (in preparation).*

[Ari12] Octavio Arizmendi, *K-divisible random variables in free probability (pre-print)*, arXiv: 1203.4780. (2012).

[Ari13] ———, *Convergence of the fourth moment and infinite divisibility*, Probab. Math. Statist. **33** (2013), no. 2, 201–212.

[Aub12] Guillaume Aubrun, *Partial transposition of random states and non-centered semicircular distributions.*, Random Matrices: Theory and Applications **01** (2012).

[AV12] Octavio Arizmendi and Carlos Vargas, *Products of free random variables and k-divisible partitions*, Elec. Comm. in Prob. **17** (2012), no. 11, 1–13.

[AV14] ———, *Norm convergence in non conmutative central limit theorems (to appear in memoirs of TVJMM).*

[BB07] Serban Belinschi and Hari Bercovici, *A new approach to subordination results in free probability*, Journal d'Analyse Mathematique **101** (2007), 357–365.

[BBLS11] Serban Belinschi, Marek Bożejko, Franz Lehner, and Roland Speicher, *The normal distribution is $\boxplus$-infinitely divisible*, Adv. Math. **226** (2011), no. 4, 3677–3698.

[BG09a] Florent Benaych-Georges, *Rectangular random matrices, related convolution*, Prob. Theory Rel. Field **144** (2009), 471–515.

[BG09b] ———, *Rectangular random matrices, related free entropy and free Fisher's information*, J. Operator Th. **62** (2009), no. 2, 371–419.

[BGS02] Anis Ben-Ghorbal and Michael Schurmann, *Non-commutative notions of stochastic independence*, Math. Proc. Camb. Phil. Soc. **133** (2002), 531–561.

[Bia98] Philippe Biane, *Processes with free increments*, Math. Z. **227** (1998), 143–174.

[BMS13] Serban Belinschi, Tobias Mai, and Roland Speicher, *Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem (pre-print)*, arXiv:1303.3196 (2013).

[BN08a] Serban Belinschi and Alexandru Nica, *Eta-series and a Boolean Bercovici-Pata bijection for bounded k-tuples*, Adv. in Math. **217** (2008), 1–41.

[BN08b] ———, *On a remarkable semigroup of homomorphisms with respect to free multiplicative convolution*, Indiana Univ. Math. Journal **57** (2008), no. 4, 1679–1713.

[BN12a] Teodor Banica and Ion Nechita, *Asymptotic eigenvalue distributions of block-transposed wishart matrices*, J. of Theoretical Probability **26** (2012), 855–869.

[BN12b] ———, *Block-modified Wishart matrices and free Poisson laws*, Houston J. Math. (2012).

[BP99] Hari Bercovici and Victorino Pata, *Stable laws and domains of attraction in free probability theory, with an appendix by Philippe Biane*, Ann. of Math. **149** (1999), no. 3, 1023–1060.

[BS06] Zhidong Bai and Jack Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Science Press (2006).

[BSTV14] Serban Belinschi, Roland Speicher, John Treilhard, and Carlos Vargas, *Operator-valued free multiplicative convolution: analytic subordination theory and applications to random matrix theory*, Internat. Math. Res. Notices (2014).

[BV93] Hari Bercovici and Dan Voiculescu, *Free convolution of measures with unbounded supports*, Indiana Univ. Math. J. **42** (1993), 733–773.

[BV95] ———, *Superconvergence to the central limit and failure of the Cramer theorem for free random variables*, Prob. Theory Rel. Field **102** (1995), 215–222.

[CD11] Romain Coulliet and Merouane Debbah, *Random Matrix Methods for Wireless Communications*, Cambridge University Press, Cambridge, 2011.

[CDS11] Romain Couillet, Merouane Debbah, and Jack Silverstein, *A deterministic equivalent for the analysis of correlated MIMO multiple access channels*, IEEE Trans. Inf. Theory **57** (2011), 3493–3514.

[CF69] Pierre Cartier and Dominique. Foata, *Problèmes combinatoires de commutation et réarrangements*, Lecture Notes in Mathematics, No. 85, Springer-Verlag, Berlin, 1969, Electronic reedition with three new appendices, Sem. Loth. Comb. 2006.

[CHD11] Romain Couillet, Jack Hoydis, and Merouane Debbah, *A deterministic equivalent approach to the performance analysis of isometric random precoded systems*, IEEE Trans. Inf. Theory (2011).

[Col03] Benoit Collins, *Moments and cumulants of polynomial random variables of unitary groups, the Itkinson-Zuber integral and free probability.*, Internat. Math. Res. Notices **227** (2003), 953–982.

[CS06] Benoit Collins and Piotr Sniady, *Integration with respect to the Haar measure on unitary, orthogonal and symplectic group.*, Comm. in Math. Phys. **264** (2006), no. 3, 773–795.

[DL07] Florian Dupuy and Philippe Loubaton, *Mutual information of frequency selective MIMO systems: an asymptotic approach*, Ann. Appl. Prob. **17** (2007), 875–930.

[DRS72] Peter Doubilet, Gian-Carlo Rota, and Richard Stanley, *On the foundations of combinatorial theory. VI. The idea of generating function*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory, Univ. California Press, Berkeley, Calif., 1972, pp. 267–318. MR 0403987 (53 #7796)

[Dyk93] Kenneth Dykema, *On certain free product factors via an extended matrix model*, J. Funct. Anal. **112** (1993), 31–60.

[Dyk06] ———, *On the S-transform over a Banach algebra*, J. Funct. Anal. **231** (2006), no. 1, 90–110.

[Ede80] Paul Edelman, *Chain enumeration and non-crossing partitions*, Discrete Math. **31** (1980).

[Ess56] Carl-Gustav Esseen, *A moment inequality with an application to the central limit theorem*, Skand. Aktuarietidskr. **39** (1956), 160–170.

[FOBS08] Reza Rashidi Far, Tamer Oraby, Wlodzimierz Bryc, and Roland Speicher, *On slow-fading MIMO systems with nonseparable correlation*, IEEE Trans. Inf. Theory **54** (2008), 544–553.

[Gir01] Vyacheslav Girko, *Theory of Stochastic Canonical Equations*, Mathematics and Its Applications, vol. 1200, Kluwer Academic Publications, 2001.

[HFS07] William Helton, Reza Rashidi Far, and Roland Speicher, *Operator-valued semicircular elements: Solving a quadratic matrix equation with positivity constraints*, Internat. Math. Res. Notices (2007).

[HLN07] Walid Hachem, Philippe Loubaton, and Jamal Najim, *Deterministic equivalents for certain functionals of large random matrices*, Ann. Appl. Prob. **17** (2007), 875–930.

[HS11a] Takahiro Hasebe and Hayato Saigo, *Joint cumulants for natural independence*, Electron. Commun. Probab. **16** (2011), 491–506.

[HS11b] ———, *The monotone cumulants*, Ann. Inst. Henri Poincaré Probab. Stat. **47** (2011), no. 4, 1160–1170.

[HT05] Uffe Haagerup and Steen Thorbjørnsen, *A new application of random matrices:* $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ *is not a group*, Ann. of Math. **162** (2005).

[JV13] Matthieu Josuat-Vergès, *Cumulants of the q-semicircular Law, Tutte Polynomials, and Heaps*, Canad. J. Math. **65** (2013), no. 4, 863–878.

[Kar07] Vladislav Kargin, *The norm of products of free random variables*, Probab. Theory Relat. Fields **139** (2007).

[KNPS12] Todd Kemp, Ivan Nourdin, Giovanni Peccati, and Roland Speicher, *Wigner chaos and the fourth moment*, Ann. Probab. **40** (2012), no. 4, 1577–1635.

[Kre72] Germain Kreweras, *Sur les partitions non croisées d'un cycle*, Discrete Math. **1** (1972), no. 4, 333–350.

[KS00] Bernadette Krawczyk and Roland Speicher, *Combinatorics of free cumulants*, J. Combin. Theory Ser. A **90** (2000), no. 2, 267–292.

[KS07] Todd Kemp and Roland Speicher, *Strong Haagerup inequalities for free R-diagonal elements*, J. Funct. Anal. **251** (2007), 141–173.

[Leh02] Franz Lehner, *Free cumulants and enumeration of connected partitions*, European J. Combin. **23** (2002), no. 8, 1025–1031.

[Leh04] ———, *Cumulants in noncommutative probability theory. I. Noncommutative exchangeability systems*, Math. Z. **248** (2004), no. 1, 67–100.

[Len10] Romuald Lenczewski, *Matricially free random variables*, J. Funct. Anal. **258** (2010), no. 12, 4075–4121.

[Len12] ———, *Matricial R-transform*, J. Funct. Anal. **262** (2012), no. 4, 1802–1844.

[MP67] Vladimir Marchenko and Leonid Pastur, *Distribution of eigenvalues for some sets of random matrices*, Math. USSR-Sbornik **1** (1967), 457–483.

[MS05] Aris Moustakas and Steven Simon, *Random matrix theory of multi-antenna communications: the Rician channel*, J. Phys. A. **38** (2005), 10859–10872.

[MS12] James Mingo and Roland Speicher, *Sharp bounds for sums associated to graphs of matrices*, J. Funct. Anal. **262** (2012), no. 5, 2272–2288.

[Mue02] Ralf Mueller, *A random matrix model of communications via antena arrays*, IEEE Trans. Inform. Th. **48** (2002), 2495–2506.

[Mur01] Naofumi Muraki, *Monotonic independence, monotonic central limit theorem and monotonic law of small numbers*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **4** (2001), no. 1, 39–58.

[Mur02] ———, *The five independences as quasi-universal products*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **5** (2002), no. 1, 113–134.

[NP05] David Nualart and Giovanni Peccati, *Central limit theorems for sequences of multiple stochastic integrals*, Ann. Probab. **33** (2005), 177–193.

[NS95] Peter Neu and Roland Speicher, *Rigorous mean-field theory for Coherent-Potential Approximation: Anderson model with free random variables*, J. Stat. Phys. **80** (1995), 1279–1308.

[NS98] Alexandru Nica and Roland Speicher, *Commutators of free random variables*, Duke Math. J. **92** (1998), no. 3.

[NS06] ———, *Lectures on the combinatorics of free probability*, London Mathematical Society Lecture Note Series, vol. 335, Cambridge University Press, Cambridge, 2006.

[NSS02] Alexandru Nica, Dimitri Shlyakhtenko, and Roland Speicher, *Operator-valued distributions I: Characterizations of freeness*, Internat. Math. Res. Notices **29** (2002), 1509–1538.

[Rot64] Gian-Carlo Rota, *On the foundations of combinatorial theory. I. Theory of Möbius functions*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **2** (1964), 340–368 (1964).

[Sch47] Marcel-Paul Schutzenberger, *Sur certains paramètres caractéristiques des systèmes d'événements compatibles et dépendants et leur application au calcul des cumulants de la répétition*, C. R. Acad. Sci. Paris **225** (1947), 277–278.

[Shl96] Dimitri Shlyakhtenko, *Random Gaussian band matrices and freeness with amalgamation*, Internat. Math. Res. Notices **20** (1996), 1013–1025.

[Spe94] Roland Speicher, *Multiplicative functions on the lattice of noncrossing partitions and free convolution*, Math. Ann. **298** (1994), no. 4, 611–628.

[Spe97] Roland Speicher, *On universal products*, Free probability theory (Waterloo, ON, 1995), Fields Inst. Commun., vol. 12, Amer. Math. Soc., Providence, RI, 1997, pp. 257–266.

[Spe98] Roland Speicher, *Combinatorial theory of the free product with amalgamation and operator-valued free probability theory*, Memoirs of the American Math. Society, vol. 132, 1998.

[Sta12] Richard P. Stanley, *Enumerative combinatorics. Volume 1*, second ed., Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 2012.

[SV12] Roland Speicher and Carlos Vargas, *Free deterministic equivalents, rectangular random matrix models, and operator-valued free probability theory, with appendix by Tobias Mai*, Random Matrices: Theory Appl. **1** (2012), no. 2, 26.

[SW97] Roland Speicher and Reza Woroudi, *Boolean convolution*, Free probability theory (Waterloo, ON, 1995), Fields Inst. Commun., vol. 12, Amer. Math. Soc., Providence, RI, 1997, pp. 267–279.

[SY13] Noriyoshi Sakuma and Hiroaki Yoshida, *New limit theorems related to free multiplicative convolution*, Studia Math. **214** (2013), 251–264.

[TV04] Antonia Tulino and Sergio Verdu, *Random Matrix Theory and Wireless Communications*, Now Publishers, Amsterdam, The Netherlands, 2004.

[Vas03] Vladimir Vasilchuk, *On the asymptotic distribution of the commutator and anticommutator of random matrices*, J. Math. Phys. **44** (2003), no. 4.

[Voi85] Dan Voiculescu, *Symmetries of some reduced free product C^*-algebras*, Operator algebras and their connections with topology and ergodic theory (Buşteni, 1983), Lecture Notes in Math., vol. 1132, Springer, Berlin, 1985, pp. 556–588.

[Voi86] ———, *Addition of certain non-commutative random variables*, J. Funct. Anal. **66** (1986), 323–346.

[Voi87] ———, *Multiplication of certain noncommuting random variables*, J. Operator Theory **18** (1987), 223–235.

[Voi91] ———, *Limit laws for random matrices and free products*, Invent. Math. **104** (1991), 201–220.

[Voi95] ———, *Operations on certain non-commutative operator-valued random variables. Recent advances in operator algebras (orleans, 1992)*, Asterisque **232** (1995).

[Voi98] ———, *A strengthened asymptotic freeness result for random matrices with applications to free entropy*, Internat. Math. Res. Notices (1998), 41–73.

[Voi00] ———, *The coalgebra of the free difference quotient and free probability*, Internat. Math. Res. Notices (2000), no. 2, 79–106.

[Voi02] ———, *Analytic subordination consequences of free Markovianity*, Indiana Univ. Math. J. **51** (2002), 1161–1166.

[Wig58] Eugene Wigner, *On the distribution of the roots of certain symmetric matrices*, Ann. of Math. **67** (1958), 325–327.

[Xu97] Feng Xu, *A random matrix model from two-dimensional yang-mills theory*, Com. Math. Phys. **190** (1997), 287–307.