

Rectangular random matrices. Related convolution

Florent Benaych-Georges

▶ **To cite this version:**

Florent Benaych-Georges. Rectangular random matrices. Related convolution. 2008. hal-00007541v6

HAL Id: hal-00007541
https://hal.science/hal-00007541v6

Preprint submitted on 4 Mar 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RECTANGULAR RANDOM MATRICES, RELATED CONVOLUTION

FLORENT BENAYCH-GEORGES

Abstract. We characterize asymptotic collective behavior of rectangular random matrices, the sizes of which tend to infinity at different rates. It appears that one can compute the limits of all non commutative moments (thus all spectral properties) of the random matrices we consider because, when embedded in a space of larger square matrices, independent rectangular random matrices are asymptotically free with amalgamation over a subalgebra. Therefore, we can define a "rectangular free convolution", which allows to deduce the singular values of the sum of two large independent rectangular random matrices from the individual singular values. This convolution is linearized by cumulants and by an analytic integral transform, that we called the "rectangular R-transform".

hal-00007541, version 6 - 4 Mar 2008

Introduction

The first problem we are going to deal with in this paper is the modeling of asymptotic collective behavior of independent rectangular random matrices. In order to explain to the reader the way we will treat this problem, let us recall him the work already done by Wigner, Pastur, Marchenko, Girko, Bai, Voiculescu,... for square random matrices. First of all, in the 50's Wigner considered self-adjoint random matrices with gaussian entries (GUE) and proved that the spectral law (i.e. the uniform distribution on the set of eigenvalues) of these random matrices converges to the so-called *semicircle law.* This result was improved, and other results giving the asymptotic spectral law of random matrices were proved (see, among many other sources, [PL02]). In the same time, people studied the local structure of the spectrum of random matrices (see, e.g. [M67]), but this is not the kind of problem we are going to study here. A new point of view was adopted in the early 90's by Voiculescu, who proposed a way to compute the asymptotic normalized trace (when the dimension n of the matrices goes to infinity) of products

$$M(s_1, n)^{\varepsilon_1} \cdots M(s_k, n)^{\varepsilon_k}$$

of random matrices taken among a family $M(1,n), M(2,n), \ldots$ of independent random $n \times n$ matrices and their adjoints, the only hypothesis being a bound on the norms of the matrices, the fact that the matrices $M(1,n), M(2,n), \ldots$ have all a limit singular law (the *singular law* of a rectangular matrix is the uniform distribution on the set of its singular values, i.e. of the eigenvalues of the absolute value of the matrix), and a property of invariance of the distributions under an action of the unitary group. These results can be deduced from the fundamental article [V91], but are presented under this form in [HP00]. The advantage of being able to compute the limit of such normalized traces is that it gives us the asymptotic normalized trace of any

Date: March 4, 2008.
MSC 2000 subject classifications. 15A52, 46L54
Key words. random matrices, free probability, free convolution

noncommutative polynomial in our independent random matrices. Hence, since the normalized trace of the k-th power of a matrix is the k-th moment of its spectral law, we are able to give the asymptotic singular law of any polynomial of our random matrices. This is why this work is said to *model the asymptotic collective behavior of independent square random matrices.* For example, it can be proved (combine results of [HP00] and [HL00]) that the asymptotic singular law of the sum of two independent random matrices whose distributions are invariant under left and right actions of the unitary group and whose singular laws converge weakly to probability measures μ_1, μ_2, only depends on μ_1 and μ_2, and can be expressed easily from μ_1 and μ_2: it is the probability measure on $[0,\infty)$, the symmetrization of which is the free convolution of the symmetrizations of μ_1, μ_2. We often present the similar result for hermitian matrices, for which we work with spectral law in the place of singular law, and for which no symmetrization is necessary, but here, we shall work with rectangular matrices, which cannot be hermitian, so the square analogue of our work will be found in non hermitian matrices. In this text, as in the work of Voiculescu presented above, we will propose a way to compute the asymptotic normalized trace of products of random matrices taken among a family of independent rectangular random matrices, whose sizes tend to infinity, but with different rates. The notion involved, similarly to freeness in Voiculescu's modeling of asymptotics of square random matrices, is freeness with amalgamation over a finite dimensional subalgebra. This notion arises from operator-valued free probability theory, but we chose to use the point of view of operator-valued free probability only when necessary, because this point of view is not satisfying in all cases: if one uses this point of view, then he has to consider the case where the ratio of sizes of our random matrices tends to zero separately. Moreover, the point of view of operator-valued free probability in the asymptotics of rectangular random matrices is developed in another paper, where we analyze the related free entropy and Fischer information ([B-G2]).

This modeling of asymptotics of rectangular random matrices will allow us to define, for $\lambda \in [0,1]$, a binary operation $\boxplus_\lambda$ on the set of symmetric probability measures, called free convolution with ratio λ, and denoted by $\boxplus_\lambda$. For μ_1, μ_2 symmetric probability measures, $\mu_1 \boxplus_\lambda \mu_2$ is defined to be the limit of the singular law of a sum of two independent rectangular random matrices, whose dimensions tend to infinity in a ratio λ, one of them being bi-unitarily invariant, and whose singular laws tend to μ_1, μ_2. In the second part of this paper, after having analyzed related cumulants, we construct an analytic integral transform which linearizes $\boxplus_\lambda$ (like the R-transform does for free convolution) and give examples. This part of this paper is the base of other researches. Firstly, in [B-G1], we study the related infinite divisibility: it is proved that the set of $\boxplus_\lambda$-infinitely divisible distributions is in a deep correspondence with the set of symmetric classical infinitely divisible distributions. Secondly, in [BBG07], we study some questions related to the support and the regularity of $\mu_1 \boxplus_\lambda \mu_2$.

Acknowledgments. We would like to thank Philippe Biane, Dan Voiculescu, and Piotr Śniady for useful discussions, as well as Thierry Cabanal-Duvillard, who organized the workshop "Journée Probabilités Libres" at MAP5 in June 2004, where the author had the opportunity to have some of these discussions. Also, we would like to thank Cécile Martineau for her contribution to the english version of this paper.

1. Asymptotic behavior of rectangular random matrices

1.1. Notes on reduction of rectangular random matrices.

In this subsection, we recall some simple facts about polar decomposition of rectangular complex matrices, which can be found in [HJ91]. Consider a $p \times q$ matrix M. When $p \leq q$, we will denote by $|M|$ the only $p \times p$

positive hermitian matrix H such that we can write $M = HT$, with T a $p\times q$ matrix such that $TT^* = I_p$. In this case, the matrix $|M|$ is the square root of MM^*. When $p > q$, we will denote by $|M|$ the only $q\times q$ positive hermitian matrix H such that we can write $M = TH$, with T a $p\times q$ matrix such that $T^*T = I_q$. In this case, the matrix $|M|$ is the square root of M^*M. In both cases, the spectrum of $|M|$ is the only (up to a permutation) family $(h_1, h_2 \ldots)$ of nonnegative real numbers such that one can write

$$M = U[\delta_i^j h_i]_{\substack{1\leq i\leq p\\ 1\leq j\leq q}} V,$$

with U, V respectively $p\times p$, $q\times q$ unitary matrices. These numbers are called *singular values* of M. The uniform distribution on $h_1, h_2 ...$ will be called the *singular law* of M (whereas the uniform distribution on the spectrum of an hermitian matrix is called its *spectral law*). Note that in both cases, for any $\alpha > 0$,

$$\frac{1}{\min(p,q)} \operatorname{Tr} |M|^\alpha = \frac{1}{\min(p,q)} \operatorname{Tr}(MM^*)^{\alpha/2} = \frac{1}{\min(p,q)} \operatorname{Tr}(M^*M)^{\alpha/2}$$

is the α-th moment of the singular law of M.

A random matrix is said to be *bi-unitarily invariant* if its distribution is invariant under the left and right actions of the unitary group.

1.2. **Definitions.** For definitions of *singular values, singular law, spectral law, (bi-)unitarily invariant and uniform random unitary matrices*, we refer to subsection 1.1. For all positive integer d, we denote the set $\{1, \ldots, d\}$ by $[d]$. We denote the normalized trace of a square matrix X by $\operatorname{tr} X$ (whatever the size of X is).

Consider a positive integer d, and sequences $q_1(n), \ldots, q_d(n)$ of pairwise distinct positive integers which all tend to infinity as n tends to infinity, and such that for all $k \in [d]$,

$$\frac{q_k(n)}{n} \underset{n\to\infty}{\longrightarrow} \rho_k \geq 0, \tag{1}$$

at most one of the ρ_k's being zero. We are going to work with rectangular random matrices of sizes $q_k(n)\times q_l(n)$. Free probability's modeling of the asymptotic behavior of square random matrices relies on a comparison between random matrices and elements of an algebra arising from operator algebra theory. To product analogous results for rectangular random matrices, we have embedded our matrices of sizes $q_k(n)\times q_l(n)$ $(k, l \in [d])$ in an algebra. Let us assume that for all n, $q_1(n) + \cdots + q_d(n) = n$ (since what we will prove would obviously also work when n is replaced by a subsequence, it is not a real restriction). Now, $n\times n$ matrices will be represented as d^2 block matrices, such that for all $k, l \in [d]$, the (k, l)-th block is a $q_k(n)\times q_l(n)$ matrix.

For all $k, l \in [d]$, for all $q_k(n)\times q_l(n)$ matrix M, let us denote by $\widetilde{M}$ the "$n\times n$ extension of M", that is the $n\times n$ matrix with (i, j)-th block M if $(i, j) = (k, l)$, and zero in the other case. Note we have preservation of adjoints ($\widetilde{M^*} = \widetilde{M}^*$), and of products:

$$\forall k, l, k', l' \in [d], \forall M \text{ of size } q_k(n)\times q_l(n), \forall N \text{ of size } q_{k'}(n)\times q_{l'}(n), \widetilde{M}\widetilde{N} = \begin{cases} \widetilde{MN} & \text{if } l = k', \\ 0 & \text{in the other case.} \end{cases}$$

At last, let us define, for all $k \in [d]$, the projector

$$p_k(n) := \widetilde{I_{q_k(n)}}.$$

Let us formalize the structure we inherit. Consider a $*$-algebra $\mathcal{A}$ endowed with a family $(p_1, \ldots, p_d)$ of non zero self-adjoint projectors (i.e. $\forall i, p_i^2 = p_i$) which are pairwise orthogonal

(i.e. $\forall i \neq j, p_i p_j = 0$), and such that $p_1 + \cdots + p_d = 1$. Any element x of $\mathcal{A}$ can then be represented

$$x = \begin{bmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & & \vdots \\ x_{d1} & \cdots & x_{dd} \end{bmatrix},$$

where $\forall i, j, x_{ij} = p_i x p_j$. This notation is compatible with the product and the involution. Assume each subalgebra $p_k \mathcal{A} p_k$ to be endowed with a tracial state φ_k (i.e. $\varphi_k(p_k) = 1$ and for all $x, y \in p_k \mathcal{A} p_k, \varphi_k(xy - yx) = 0$) such that for all $k, l \in [d], x \in p_k \mathcal{A} p_l, y \in p_l \mathcal{A} p_k$,

$$\rho_k \varphi_k(xy) = \rho_l \varphi_l(yx), \tag{2}$$

where $(\rho_1, \ldots, \rho_d)$ is still the sequence of nonnegative real numbers defined by (1). Recall that at most one of the ρ_k's is zero.

Definition 1.1. *Such a family $(\mathcal{A}, p_1, \ldots, p_d, \varphi_1, \ldots, \varphi_d)$ will be called a* $(\rho_1, \ldots, \rho_d)$-rectangular probability space. *Elements of the union of the $p_k \mathcal{A} p_l$'s ($k, l \in [d]$) will be called* simple elements.

Example 1.2. *$\mathcal{A}$ is the $*$-algebra of $n \times n$ complex matrices, $p_1, \ldots, p_d$ are $p_1(n), \ldots, p_d(n)$ previously defined, and each φ_k is $\frac{1}{q_k(n)} \operatorname{Tr}$. Then (2) is satisfied when each ρ_k is replaced by $q_k(n)/n$.*

Consider a $(\rho_1, \ldots, \rho_d)$-rectangular probability space $(\mathcal{A}, p_1, \ldots, p_d, \varphi_1, \ldots, \varphi_d)$. Denote by $\mathcal{D}$ the linear span of the p_k's, then $\mathcal{D}$ is an algebra and the function $\mathrm{E} : \mathcal{A} \to \mathbb{C}^d$, which maps $x \in \mathcal{A}$ to $\mathrm{E}(x) = \sum_{k=1}^d \varphi_k(x_{kk}) p_k$ is a *conditional expectation* from $\mathcal{A}$ to $\mathcal{D}$, i.e. $\mathrm{E}(1) = 1$ and $\forall (d, a, d') \in \mathcal{D} \times \mathcal{A} \times \mathcal{D}, \mathrm{E}(dad') = d\,\mathrm{E}(a)d'$.

The following definition, first appeared in [V95], gives the right notion to describe asymptotics of independent random matrices.

Definition 1.3. *In an $*$-algebra $\mathcal{B}$ endowed with a $*$-subalgebra $\mathcal{C}$ and a conditional expectation $\mathrm{E}_{\mathcal{C}}$ from $\mathcal{B}$ to $\mathcal{C}$, a family $(\mathcal{C} \subset \mathcal{B}_\alpha)_{\alpha \in A}$ of $*$-subalgebras is said to be* free with amalgamation over $\mathcal{C}$ *if for all $m \geq 1$, for all $\alpha_1 \neq \cdots \neq \alpha_m \in A$, for all $x_1, \ldots, x_m$ elements of respectively $\mathcal{B}_{\alpha_1}$, …, $\mathcal{B}_{\alpha_m}$ one has*

$$\mathrm{E}_{\mathcal{C}}(x_1) = \cdots = \mathrm{E}_{\mathcal{C}}(x_m) = 0 \quad \Longrightarrow \quad \mathrm{E}_{\mathcal{C}}(x_1 \cdots x_m) = 0. \tag{3}$$

A family of subsets of $\mathcal{B}$ is said to be free with amalgamation over $\mathcal{C}$ *if the subsets are contained in $*$-subalgebras which are free with amalgamation over $\mathcal{C}$.*

Definition 1.4. *In a $(\rho_1, \ldots, \rho_d)$-rectangular probability space, the* $\mathcal{D}$-distribution *of a family $(a_j)_{j \in J}$ of simple elements is the function which maps any polynomial P in the noncommutative variables $(X_j, X_j^*)_{j \in J}$, with coefficients in $\mathcal{D}$, to $\mathrm{E}(P(a_j, a_j^*)_{j \in J})$.*

The last notion we have to introduce is the *convergence in $\mathcal{D}$-distribution.* Note that the $\mathcal{D}$'s of different spaces can all be identified with $\mathbb{C}^d$.

Definition 1.5. • *If for all n, $(\mathcal{A}_n, p_{1,n}, \ldots, p_{d,n}, \varphi_{1,n}, \ldots, \varphi_{d,n})$ is a $(\rho_{1,n}, \ldots, \rho_{d,n})$-probability space such that*

$$(\rho_{1,n}, \ldots, \rho_{d,n}) \underset{n \to \infty}{\longrightarrow} (\rho_1, \ldots, \rho_d),$$

a family $(a_j(n))_{j \in J}$ of simple elements of $\mathcal{A}_n$ is said to converge in $\mathcal{D}$-distribution*, when n goes to infinity, to a family $(a_j)_{j \in J}$ of elements of $\mathcal{A}$ if the $\mathcal{D}$-distributions converge pointwise.*

• *In the context of example 1.2, if for all n, $a_j(n)$'s are $n \times n$ random matrices,* convergence

in $\mathcal{D}$-distribution in probability *of the family* $(a_j(n))_{j\in J}$ *to* $(a_j)_{j\in J}$ *is the convergence in probability, when* $n \to \infty$, *of the random variable* $\mathrm{E}(P(a_j(n), a_j(n)^*)_{j\in J})$ *to* $\mathrm{E}(P(a_j, a_j^*)_{j\in J})$ *for all polynomial* P *in the noncommutative variables* $(X_j, X_j^*)_{j\in J}$.

1.3. **Statement of the theorems about random matrices.** Let, for $s \in \mathbb{N}, k, l \in [d]$, $n \geq 1$, $R(s,k,l,n)$ be a $q_k(n)\times q_l(n)$ random matrix such that for all n, the family $(R(s,k,l,n))_{s,k,l}$ is independent. We suppose that for all $s \in \mathbb{N}, k \neq l \in [d]$, $R(s,k,l,n)$ is bi-unitarily invariant and that its singular law converge in probability to a deterministic probability measure. We also suppose that for all $s \in \mathbb{N}$ odd, for all $k \in [d]$, $R(s,k,k,n)$ is unitary uniform and that for all $s \in \mathbb{N}$ even, for all $k \in [d]$, $R(s,k,k,n)$ is hermitian, unitarily invariant and that its spectral law converge in probability to a deterministic probability measure. Consider also, for $n \geq 1$, a family $(C(i,n))_{i\in I}$ of diagonal deterministic matrices, each $C(i,n)$ having the size $q_{k_i}(n)\times q_{l_i}(n)$, for a certain $(k_i, l_i) \in [d]\times[d]$, which contains $p_1(n), \ldots, p_d(n)$ and is stable under product and taking adjoints. We suppose that for all $i \in I$ such that $k_i = l_i$, the normalized trace of $C(i,n)$ has a finite limit when n goes to infinity.

We suppose that for the operator norm $||\cdot||$ associated to canonical hermitian norms these matrices are uniformly bounded (in n, s, k, l, i).

The hypothesis of our main theorems are not the same if one supposes all ρ_k's to be positive or not. In the following theorem, all ρ_k's are supposed to be positive.

Theorem 1.6 (Case where all ρ_k's are positive)**.** *Under these hypothesis, the family* $(\widetilde{R}(s,k,l,n)\,;\, s \in \mathbb{N}, k, l \in [d]) \cup (\widetilde{C}(i,n)\,;\, i \in I)$ *converges in* $\mathcal{D}$*-distribution in probability to a family which is free with amalgamation over* $\mathcal{D}$.

In the following theorem, we suppose one of the ρ_k's, say ρ_1, to be zero. We supposed that at most one of the ρ_k's is zero, so $\rho_2, \ldots, \rho_d$ are positive.

Theorem 1.7 (Case where $\rho_1 = 0$, other ρ_k's are positive)**.** *Assume moreover that for all* $s, k \neq l, n$, $R(s,k,l,n)$ *has deterministic singular values and that for all* $s \in \mathbb{N}$ *even,* $k \in [d]$, $R(s,k,k,n)$ *has a deterministic spectrum. Then the family* $(\widetilde{R}(s,k,l,n)\,;\, s \in \mathbb{N}, k, l \in [d]) \cup (\widetilde{C}(i,n)\,;\, i \in I)$ *converges in* $\mathcal{D}$*-distribution in probability to a family which is free with amalgamation over* $\mathcal{D}$.

Corollary 1.8 (Product of rectangular matrices with asymptotic null ratio)**.** *Let* $q_1(n), q_2(n)$ *be sequences which tend to infinity in such a way that* $q_1(n)/q_2(n)$ *tends to zero. If, for all* n, $M_1(n), M_2(n)$ *are independent* $q_1(n)\times q_2(n)$ *bi-unitarily invariant rectangular random matrices with deterministic singular values and limit singular laws, then for all* $p \geq 1$, $i_1, \ldots, i_p \in \{1, 2\}$,

- $M_{i_1}(n)^* M_{i_2}(n) \cdots M_{i_{2p-1}}(n)^* M_{i_{2p}}(n)$ *has null limit normalized trace,*
- $M_{i_1}(n)^* M_{i_2}(n) \cdots M_{i_{2p-1}}(n)^* M_{i_{2p}}(n)$ *has null limit normalized trace, except if any matrix* $M_1(n)$ *or* $M_2(n)$ *in the product is followed by respectively* $M_1(n)^*$ *or* $M_2(n)^*$, *in which case the limit can be computed using the well known asymptotic freeness of square random matrices.*

1.4. **Comparison with already existing results.** There already exists some results which allow to compute asymptotics of normalized traces of products of rectangular random matrices: on one hand the results about Wishart matrices and on the other hand the theorem 4.1 of [Sh96].

Theorems 1.6 and 1.7 allow us to compute the limit (for convergence in probability) of all normalized traces of matrices of the type

$$R(s_1,k_1,l_1,n)^{\varepsilon_1} C(i_1,n) R(s_2,k_2,l_2,n)^{\varepsilon_2} C(i_2,n) \cdots C(i_{p-1},n) R(s_p,k_p,l_p,n)^{\varepsilon_p} C(i_p,n), \tag{4}$$

with $p \geq 1$, $s_1, \ldots, s_p \in \mathbb{N}$, $k_1, l_1, \ldots, k_p, l_p \in [d]$, $\varepsilon_1, \ldots, \varepsilon_p \in \{1, *\}$, and $i_1, \ldots, i_{p-1} \in I$ such that the product is possible and is a square matrix.

The Wishart matrices are matrices of the type RR^*, where R is a rectangular gaussian matrix. In [HP00], [CC04] or [CDM05], the authors give results of asymptotic freeness for such random matrices. In the case we are interested in, where the random matrix R is bi-unitarily invariant, this asymptotic freeness is only a consequence of the asymptotic freeness of unitarily invariant hermitian random matrices. With our notations, it does not allow to suppose that $\rho_1 = 0$, and to consider products **where rectangular matrices are not immediately followed by their adjoint**, as the one of (4).

Theorem 4.1 of the paper [Sh96] by Shlyakhtenko allows, embedding rectangular matrices in larger self-adjoint matrices ($R \mapsto \begin{bmatrix} 0 & R \\ R^* & 0 \end{bmatrix}$), to consider products of the type of the one of (4), but **only when the ρ_i's are positive**, the matrices $R(s, k, l, n)$ are **gaussian**, and the deterministic matrices are **diagonal** and have, in a certain sense, **a limit in** $L^\infty[0, 1]$. Moreover, this theorem states the convergence of the **expectations** of the normalized traces, but not their convergence in probability (even though this could it is possible to expect to pass, using the concentration results of [GZ00], from convergence in expectation to convergence in probability in Shlyakhtenko's result).

As an example, no general result had been proved about the convergence, when $n, p \to \infty$ such that $\frac{n}{p} \to \lambda \geq 0$, of the singular law of $M(n, p)+N(n, p)$, where $M(n, p), N(n, p)$ are independent bi-unitarily invariant $n\times p$ random matrices with limit singular laws. We did not even know if this limit existed, even though some computations have been done in some particular cases ([DS04], [HLN04], [HLN05])). Indeed, the computation, by the moments method, of the singular law of $M(n, p) + N(n, p)$ involves the computation of the normalized trace of

$$[(M(n, p) + N(n, p))(M(n, p) + N(n, p))^*]^k,$$

the expansion of which contains products where some rectangular matrices are not followed by their adjoints.

1.5. **Proof of theorem 1.6 (case where all ρ_k's are positive).** Let us introduce, on the space of $n\times n$ matrices, the Schatten p-norms $||.||_p$ ($p \in [1, \infty]$), defined, for $p < \infty$, by

$$||X||_p = (\operatorname{tr} |X|^p)^{\frac{1}{p}},$$

and $||.||_\infty = ||.||$. They satisfy classical Hölder inequalities ([N74]).

Step I. Fix $k, l \in [d]$ and consider, for $n \geq 1$, a bi-unitarily invariant $q_k(n)\times q_l(n)$ random matrix $R(n)$ such that the sequence $(R(n))$ is uniformly bounded for $||.||$ by $C > 0$ and such that the singular law of $R(n)$ converge in probability to a deterministic probability measure. Note first that, by Carleman criterion, it implies that its singular law converges in probability to a probability measure μ. It is easy, transferring the proof of lemma 4.3.10 p. 160 of [HP00], to prove that the distribution of $R(n)$ can be realized as the distribution of $U(n)\Lambda(n)V(n)$ where $U(n)$, $\Lambda(n)$, $V(n)$ are independent, $U(n)$, $V(n)$ are respectively $q_k(n)\times q_k(n)$, $q_l(n)\times q_l(n)$ uniform random unitary matrices, and $\Lambda(n)$ is a rectangular diagonal matrix,

$$\Lambda(n) = [\delta_i^j \lambda(i, n)]_{\substack{1 \leq i \leq q_k(n) \\ 1 \leq l \leq q_l(n)}} \text{ with } 0 \leq \lambda(1, n) \leq \cdots \leq \lambda(\min(q_k(n), q_l(n)), n).$$

Choose deterministic $0 \leq \xi(1,n) \leq \cdots \leq \xi(\min(q_k(n), q_l(n)), n) \leq C$ such that the uniform distribution on the $\xi(1,n)$'s converges weakly to μ. Now set

$$\Xi(n) := [\delta_i^j \lambda(i,n)]_{\substack{1 \leq i \leq q_k(n) \\ 1 \leq j \leq q_l(n)}} \text{ and } N(n) := U(n)\Xi(n)V(n).$$

Let us prove that for all P polynomial in noncommutative random variables X, X^*, for all $p \in [1,\infty)$, $||P(\widetilde{R}(n)) - P(\widetilde{N}(n))||_p$ converges in probability to zero.

By lemma 4.3.4 p. 152 of [HP00], for all $p \geq 1$,

$$||\widetilde{R}(n)^* - \widetilde{N}(n)^*||_p = ||\widetilde{R}(n) - \widetilde{N}(n)||_p = ||\widetilde{\Lambda}(n) - \widetilde{\Xi}(n)||_p$$

converges in probability to zero. So, for all $m \geq 1$, $\varepsilon_1, \ldots, \varepsilon_m \in \{., *\}$,

$$\begin{aligned} & ||\widetilde{R}(n)^{\varepsilon_1} \cdots \widetilde{R}(n)^{\varepsilon_m} - \widetilde{N}(n)^{\varepsilon_1} \cdots \widetilde{N}(n)^{\varepsilon_m}||_p \\ \leq & \sum_{l=1}^m ||\widetilde{R}(n)^{\varepsilon_1} \cdots \widetilde{R}(n)^{\varepsilon_{l-1}} (\widetilde{R}(n)^{\varepsilon_l} - \widetilde{N}(n)^{\varepsilon_l}) \widetilde{N}(n)^{\varepsilon_{l+1}} \cdots \widetilde{N}(n)^{\varepsilon_m}||_p \\ \leq & \sum_{l=1}^m C^{m-1} ||\widetilde{R}(n)^{\varepsilon_l} - \widetilde{N}(n)^{\varepsilon_l})||_p, \end{aligned}$$

which proves the result.

Note that in the case where R is a uniformly bounded unitarily invariant $q_k(n) \times q_k(n)$ hermitian random matrix, the moments of the spectral law of which converge in probability, then the same work can be done, replacing $V(n)$ by $U(n)^*$ and singular values by eigenvalues, and the same conclusion holds: for all $P \in \mathbb{C}[X]$, for all $p \in [1,\infty)$, $||P(\widetilde{R}(n)) - P(\widetilde{N}(n))||_p$ converges in probability to zero.

Step II. By hypothesis, there exists $i_1, \ldots, i_d \in I$ such that for all n, k, $C(i_k, n) = p_k(n)$ and there exists a map β from I^2 to I and an involution $*$ from I such that for all $n \geq 0, i, j \in I$ such that $l_i = k_j$,

$$C(i,n)C(j,n) = C(\beta(i,j),n), \quad C(i,n)^* = C(*(i),n). \tag{5}$$

Consider, in a $(\rho_1, \ldots, \rho_d)$-rectangular probability space $(\mathcal{A}, p_1, \ldots, p_d, \varphi_1, \ldots, \varphi_d)$, a family $a(s,k,l)$ $(s \in \mathbb{N}, k, l \in [d])$, $c(i)$ $(i \in I)$ of elements of $\mathcal{A}$, with $c(i_1) = p_1$,..., $c(i_d) = p_d$, whose $\mathcal{D}$-distribution is defined by the following rules:

(i) For all $s \in \mathbb{N}, k, l \in [d]$, $a(s,k,l) \in p_k \mathcal{A} p_l$, and for all $i \in I$, $c(i) \in p_{k_i} \mathcal{A} p_{l_i}$.
(ii) For all $i, j \in I$ such that $l_i = k_j$, $c(i)c(j) = c(\beta(i,j)), c(i)^* = c(*(i))$.
(iii) For all $s \in \mathbb{N}, k \neq l \in [d]$, for all $r \geq 0$,

$$\varphi_k[(a(s,k,l)a(s,k,l)^*)^r] = \lim_{n\to\infty} \operatorname{tr}(R(s,k,l,n)R(s,k,l,n)^*)^r,$$

$$\varphi_l[(a(s,k,l)^*a(s,k,l))^r] = \lim_{n\to\infty} \operatorname{tr}(R(s,k,l,n)^*R(s,k,l,n))^r.$$

(iv) For all $s \in \mathbb{N}$ odd, $k \in [d]$, $a(s,k,k)$ is unitary in $p_k \mathcal{A} p_k$, and for all $r \in \mathbb{Z} - \{0\}$, $\varphi_k(a(s,k,k)^r) = 0$.
(v) For all $s \in \mathbb{N}$ even, $k \in [d]$, $a(s,k,k)$ is self-adjoint and for all $r \geq 0$,

$$\varphi_k[a(s,k,k)^r] = \lim_{n\to\infty} \operatorname{tr} R(s,k,k,n)^r,$$

(vi) For all $i \in I$ such that $k_i = l_i$, $\varphi_{k_i}(c(i))$ is the limit of the normalized trace of $C(i,n)$.
(vii) The family $\{a(s,k,l)\}$ $(s \in \mathbb{N}, k, l \in [d])$, $\{c(i)\,;\, i \in I\}$ is free with amalgamation over $\mathcal{D}$.

Note that such a space and such a family exist, they are given by the free product with amalgamation ([VDN91],[S98]).

Let us denote by $\mathcal{C}$ the linear span of the $c(i)$'s. By (5) and (ii), for all n, there is a $*$-algebra morphism ψ_n from $\mathcal{C}$ to the linear span $\mathcal{C}_n$ of the $\widetilde{C}(i,n)$'s. By (vi), for all $c \in \mathcal{C}$, for all $k \in [d]$, the normalized trace of the k-th diagonal block of $\psi_n(c)$ tends to $\varphi_k(c_{kk})$.

Since $1 \in \mathcal{C}$, it suffices to prove that for all $m \geq 0$, for all $(s_j, k_j, l_j)_{j\in[m]} \in (\mathbb{N}\times[d]\times[d])^m$, for all $P_1, \ldots, P_m$ polynomials in the noncommutative variables X, X^*, for all $d(0), \ldots, d(m) \in \mathcal{C}$, $\mathrm{E}[\prod_{j=1}^m P_j(\widetilde{R}(s_j,k_j,l_j,n))\psi_n(d(j))]$ converges in probability to $\mathrm{E}[d(0)\prod_{j=1}^m P_j(a(s_j,k_j,l_j))d(j)]$.

Step III. By the formula $\mathrm{Tr}\, XY = \mathrm{Tr}\, YX$ and since all ρ_k's are positive, it is easy to see that we can remove $d(0)$.

Step IV. Then, we claim that it suffices to prove it when the $R(s,k,l,n)$'s are replaced by the $N(s,k,l,n)$'s, where the $N(s,k,l,n)$'s, are defined from the $R(s,k,l,n)$'s like $N(n)$ from $R(n)$ in the step I. Indeed, the normalized trace of any diagonal block of

$$\prod_{j=1}^m P_j(\widetilde{R}(s_j,k_j,l_j,n))\psi_n(d(j)) - \prod_{j=1}^m P_j(\widetilde{N}(s_j,k_j,l_j,n))\psi_n(d(j))$$

has an absolute value less or equal than the $||\cdot||_1$ norm of their difference times one of the $n/q_k(n)$, which is less or equal, by a decomposition like in step I, than

$$\max_{k\in[d]} \frac{n}{q_k(n)} \sum_{j=1}^m C^m D^{m-1} ||P_j(\widetilde{R}(s_j,k_j,l_j,n)) - P_j(\widetilde{N}(s_j,k_j,l_j,n))||_1,$$

where $D = \max_{\substack{|z|,|z'|\leq C \\ j\in[m]}} |P_j(z,z')|$ which converges in probability to zero, by step I and because all ρ_k's are positive.

So let us prove that for all $k \in [d]$, the normalized trace of the k-th diagonal block of

$$\prod_{j=1}^m P_j(\widetilde{N}(s_j,k_j,l_j,n))\psi_n(d(j))$$

converges in probability to the k-th coordinate of $\mathrm{E}[\prod_{j=1}^m P_j(a(s_j,k_j,l_j))d(j)]$.

Step V. We will prove it by induction on m. If $m = 0$, the result is clear. In the other case, let us suppose the result to be proved to the ranks $0, \ldots, m-1$. Let us denote $x(j) = P_j(a(s_j,k_j,l_j))$ and $X(j,n) = P_j(\widetilde{N}(s_j,k_j,l_j,n))$. Since $\mathcal{C}$ is an algebra containing $p_1, \ldots, p_d$, if, for a certain $j \in [m]$, one would replace $X(j,n)$ by one of the $p_k(n)$'s and $x(j)$ by the corresponding p_k, then the result would follow from the induction hypothesis (and from step III if $j = 1$). Thus, by linearity, one can, for all $j \in [m]$, add a linear combination of the $p_k(n)$'s to $X(j,n)$ and the same linear combination of the p_k's to $x(j)$. Therefore we can assume that for all j, $E(x(j)) = 0$. By linearity, we can moreover suppose that for all j, $x(j)$ is a simple element. Note that these suppositions imply that for all n, j, $X(n,j)$ has only one non zero block, $\mathrm{E}(X(j,n))$ converges in probability to zero.

From now on, we use the cyclic order on $[m]$. This means that to put the index $m+1$ on an element amounts to put the index 1. As simple elements, the $x(j)$'s have the following property: the product of any of the $x(j)$'s by one of the p_k's is either zero or $x(j)$, and the same holds for $X(j,n)$'s with $p_k(n)$'s. Thus, if, for a certain $j \in [m]$ such that $(s_j,k_j,l_j) = (s_{j+1},k_{j+1},l_{j+1})$, one would replace $d(j)$ by a linear combination of the $p_k(n)$'s, then the result would follow from the induction hypothesis (and from step III if $j = m$). So one can suppose that for all $j \in [m-1]$ such that $(s_j,k_j,l_j) = (s_{j+1},k_{j+1},l_{j+1})$, $E(d(j)) = 0$.

At last, for $j \in [m]$ such that $(s_j, k_j, l_j) \neq (s_{j+1}, k_{j+1}, l_{j+1})$, one can write

$$d(j) = d'(j) + \sum_{k \in [d]} \lambda_k(j) p_k, \text{ with } \mathrm{E}(d'(j)) = 0.$$

Then, by linearity, one can suppose that for all such j, $E(d(j)) = 0$ or $\exists k \in [d], d(j) = p_k$.

To conclude, we have shown that we only have to prove that $\mathrm{E}[\prod_{j=1}^{m} X(j,n)\psi_n(d(j))]$ converges in probability to $\mathrm{E}[\prod_{j=1}^{m} x(j)d(j)]$ under the hypothesis that all $x(j)$'s are simple elements and satisfy $\mathrm{E}(x(j)) = 0$, and for all $j \in [m]$,

$$\mathrm{E}(d(j)) = 0 \text{ or } [d(j) \in \{p_1, \ldots, p_d\} \text{ and } (s_j, k_j, l_j) \neq (s_{j+1}, k_{j+1}, l_{j+1})].$$

Note that under this hypothesis, by freeness with amalgamation over $\mathcal{D}$, $\mathrm{E}[\prod_{j=1}^{m} x(j)d(j)] = 0$.

Step VI. For all s, k, l, n with $k \neq l$, the matrix $N(s,k,l,n)$ has been defined by the work of step I, so we can write

$$N(s,k,l,n) = U(s,k,l,n)\Xi(s,k,l,n)V(s,k,l,n),$$

with $U(s,k,l,n)$, $V(s,k,l,n)$ uniform random unitary matrices with respective sizes $q_k(n)$, $q_l(n)$, and $\Xi(s,k,l,n)$ a deterministic diagonal $q_k(n) \times q_l(n)$ matrix. In the same way, when s is even, $N(s,k,k,n)$ can be written

$$N(s,k,l,n) = U(s,k,k,n)\Xi(s,k,k,n)U^*(s,k,k,n),$$

and when s is odd, $N(s,k,k,n)$ is a uniform random matrix of $\mathbb{U}_k(n)$.

Consider $j \in [m]$. If $k_j \neq l_j$, then $X(j,n)$ is of one of the 4 following types:

(1) $\widetilde{U}(s_j,k_j,l_j,n)\widetilde{\Delta}_j\widetilde{V}(s_j,k_j,l_j,n)$, where Δ_j is a $q_{k_j}(n) \times q_{l_j}(n)$ diagonal matrix whose diagonal terms are the images of the diagonal terms of $\Xi(s_j,k_j,l_j,n)$ by an odd polynomial.

(2) $\widetilde{V}(s_j,k_j,l_j,n)^*\widetilde{\Delta}_j\widetilde{U}(s_j,k_j,l_j,n)^*$, where Δ_j is a $q_{l_j}(n) \times q_{k_j}(n)$ diagonal matrix whose diagonal terms are the images of the diagonal terms of $\Xi(s_j,k_j,l_j,n)$ by an odd polynomial.

(3) $\widetilde{U}(s_j,k_j,l_j,n)\widetilde{\Delta}_j\widetilde{U}(s_j,k_j,l_j,n)^*$, where Δ_j is a $q_{k_j}(n) \times q_{k_j}(n)$ diagonal matrix whose diagonal terms are the images of the diagonal terms of $\Xi(s_j,k_j,l_j,n)$ by an even polynomial.

(4) $\widetilde{V}(s_j,k_j,l_j,n)^*\widetilde{\Delta}_j\widetilde{V}(s_j,k_j,l_j,n)$, where Δ_j is a $q_{l_j}(n) \times q_{l_j}(n)$ diagonal matrix whose diagonal terms are the images of the diagonal terms of $\Xi(s_j,k_j,l_j,n)$ by an even polynomial.

And if $k_j = l_j$, when s_j is even, only the last type is possible, and when s_j is odd, $X(j,n)$ is a linear combination of positive or negative powers of $\widetilde{N}(s_j,k_j,l_j,n)$. By linearity again, this combination can be replaced by a non zero power of $\widetilde{N}(s_j,k_j,l_j,n)$.

Note that in all types previously presented, the constant diagonal matrix Δ_j satisfies $\mathrm{E}(\widetilde{\Delta}_j) \underset{n\to\infty}{\longrightarrow} 0$. Moreover, for all n, the family of random unitary matrices arising in these decompositions are independent, so the result follows from case 1°) of proposition 4.4 and remark 4.5.

1.6. **Proof of theorem 1.7 (case where $\rho_1 = 0$, all other ρ_k's are positive).** We shall follow closely the proof of theorem 1.6 (case where all ρ_k's are positive) given in the previous section, but because of the weaker hypothesis (singular values of non hermitian random matrices and spectrum of hermitian ones are now supposed to be deterministic), we shall skip steps I and IV, in which non hermitian (resp. hermitian) random matrices where approximated by

random matrices with deterministic singular values (resp. deterministic spectrum). However, the hypothesis $\rho_1 = 0$ will make things slightly harder.

Copying step II of the previous proof, we have to prove, for all $m \geq 0$, the following proposition: $H(m)$:= "for all $(s_j, k_j, l_j)_{j\in[m]} \in (\mathbb{N}\times[d]\times[d])^m$, for all $P_1, \ldots, P_m$ polynomials in the noncommutative variables X, X^*, for all $d(0), \ldots, d(m) \in \mathcal{C}$, for all $k \in [d]$,

$$\mathrm{E}[\psi_n(d(0)) \prod_{j=1}^{m} P_j(\widetilde{R}(s_j, k_j, l_j, n))\psi_n(d(j))]$$

converges in probability to $\mathrm{E}[d(0)\prod_{j=1}^m P_j(a(s_j, k_j, l_j))d(j)]$."

Note first that by linearity, it suffices to prove it when $d(0), d(m)$ are supposed to be simple, which implies that there exists $u, v \in [d]$ such that $p_u d(0) = d(0)$, $d(m)p_v = d(m)$. Let us denote by $H(m, u, v)$ the proposition H in the particular case where $p_u d(0) = d(0)$, $d(m)p_v = d(m)$. Let us also denote by $H'(m)$ the proposition $H(m)$ in the particular case where $d(0) = 1$.

Note first that $H(m, u, v)$ is obvious when $u \neq v$. Moreover, like in step III of the previous section, we see that for all $u \in [d] - \{1\}$, $H'(m)$ implies $H(m, u, u)$. Thus we only have to prove that for all $m \geq 0$, $H'(m)$ and $H(m, 1, 1)$ hold.

Let us prove it by induction on m. For $m = 0$, the result is immediate. Consider $m \geq 1$ such that $H'(l)$ and $H(l, 1, 1)$ hold for all $l < m$.

Let us first prove $H'(m)$. Consider $(s_j, k_j, l_j)_{j\in[m]} \in (\mathbb{N}\times[d]\times[d])^m$, $P_1, \ldots, P_m$ polynomials in the noncommutative variables X, X^*, $d(1), \ldots, d(m) \in \mathcal{C}$. Let us prove that

$$\mathrm{E}[\prod_{j=1}^{m} P_j(\widetilde{R}(s_j, k_j, l_j, n))\psi_n(d(j))]$$

converges in probability to $\mathrm{E}[\prod_{j=1}^m P_j(a(s_j, k_j, l_j))d(j)]$. Let us denote $x(j) = P_j(a(s_j, k_j, l_j))$ and $X(j, n) = P_j(\widetilde{N}(s_j, k_j, l_j, n))$. Similarly to step V of the previous section, by the induction hypothesis and the fact that $H'(m-1)$ implies $H(m-1, u, u)$, we can suppose that for all n, j, $X(n, j)$ has only one non zero block and $\mathrm{E}(X(j, n))$ converges in probability to zero. We can also suppose that for all $j \in [m]$ such that with the cyclic order, $(s_j, k_j, l_j) = (s_{j+1}, k_{j+1}, l_{j+1})$, $E(d(j)) = 0$ or there exists $k \in [d]$ such that $d(j) = p_k$.

Under these assumptions, we only have to prove that $\mathrm{E}[\prod_{j=1}^m X(j, n)\psi_n(d(j))]$ converges in probability to $\mathrm{E}[\prod_{j=1}^m x(j)d(j)] = 0$. Following step VI of the previous section, it appears to be an application of case 1°) of proposition 4.4 and remark 4.5.

Similarly, $H(m, 1, 1)$ is an application of case 2°) of proposition 4.4.

2. The rectangular free additive convolution

For ν probability measure on $\mathbb{R}$, denote by $\tilde{\nu}$ the *symmetrization* of ν, which is the probability measure defined by $\tilde{\nu}(B) = \frac{1}{2}(\nu(B) + \nu(-B))$ for all Borel set B.

One of the interests of modeling the asymptotic behavior of square random matrices with free probabilities is the possibility to compute the asymptotic spectral law of an hermitian matrix which is a function of several independent random matrices. In particular ([V91]), if $M(1, n)$, $M(2, n)$ are independent hermitian unitarily invariant random matrices whose spectral laws tend respectively to ν_1, ν_2 when their dimension n goes to infinity, then the spectral law of $M(1, n) + M(2, n)$ tends to $\nu_1 \boxplus \nu_2$ (free additive convolution of ν_1, ν_2), and the spectral law of

$\sqrt{M(1,n)}M(2,n)\sqrt{M(1,n)}$ (when the matrices are positive) tends to $\nu_1 \boxtimes \nu_2$ (free multiplicative convolution of ν_1, ν_2). In the same way (combine Theorem 4.3.11 of [HP00] and Propositions 3.5, 3.6 of [HL00]), if the matrices $M(1,n), M(2,n)$ are square, bi-unitarily invariant (and still independent), with asymptotic singular distributions ν_1, ν_2, then the symmetrization of the singular distribution of $M(1,n)+M(2,n)$ tends to $\tilde{\nu}_1 \boxplus \tilde{\nu}_2$, and the push-forward, by the function $x \to x^2$, of the singular distribution of $M(1,n)M(2,n)$ is the free multiplicative convolution of the push-forwards, by $x \to x^2$, of ν_1, ν_2.

Now, suppose that $M(1,n)$, $M(2,n)$, instead of being square, are rectangular (independent and bi-unitarily invariant) $q_1(n) \times q_2(n)$ random matrices whose singular laws tend to ν_1, ν_2. We keep the notations introduced in section 1.2, but suppose that $d = 2$:

$$q_1(n) + q_2(n) = n, \quad q_1(n), q_2(n) \underset{n\to\infty}{\longrightarrow} \infty, \quad \frac{q_1(n)}{n} \underset{n\to\infty}{\longrightarrow} \rho_1 \geq 0, \quad \frac{q_2(n)}{n} \underset{n\to\infty}{\longrightarrow} \rho_2 > 0.$$

Since for $1 \leq q \leq q'$, the singular law ν of a $q \times q'$ matrix M is related to the singular law ν' of M^* by $\nu' = (q'-q)/q'\delta_0 + q/q'\nu$, we can, without restriction, suppose that $q_1(n) \leq q_2(n)$.

First, the singular distribution of $M(1,n)M(2,n)^*$ tends to a distribution that can be computed by free probability theory. Indeed, for all $r \geq 1$ even, the r-th moment of the singular distribution of $M(1,n)M(2,n)^*$ is

$$\frac{\mathrm{Tr}(M(1,n)M(2,n)^*M(2,n)M(1,n)^*)^{r/2}}{q_1(n)} = \frac{q_2(n)\,\mathrm{Tr}(M(2,n)^*M(2,n)M(1,n)^*M(1,n))^{r/2}}{q_1(n)q_2(n)}, \tag{6}$$

which tends to ρ_2/ρ_1 (when $\rho_1 > 0$) times the $r/2$-th moment of $\tau_1 \boxtimes \tau_2$, where τ_1, τ_2 are the limit spectral distributions of $M(2,n)^*M(2,n), M(1,n)^*M(1,n)$. τ_1, τ_2 can easily be computed from ν_1, ν_2. Thus, if $\rho_1 > 0$, the limit of the spectral distribution of $M(1,n)M(2,n)^*$ can be easily computed using tools of free probability. If $\rho_1 = 0$, then the explicit computation of the expectation and of the variance of (6) for $r = 2$ gives us the convergence of the singular law of $M(1,n)M(2,n)^*$ to the Dirac measure in zero.

But until now, it wasn't possible to compute the singular law of $M(1,n) + M(2,n)$ with the tools of free probability theory. Indeed, for $r \geq 1$ even, its r-th moment is given by

$$\frac{1}{q_1(n)}\,\mathrm{Tr}((M(1,n) + M(2,n))(M(1,n)^* + M(1,n)^*))^{r/2},$$

which can be expanded into the sum, over all $i \in \{1,2\}^r$, of

$$\frac{1}{q_1(n)}\,\mathrm{Tr}\, M(i_1,n)M(i_2,n)^* \cdots M(i_{r-1},n)M(i_r,n)^*. \tag{7}$$

Asymptotics of normalized traces like in (7) cannot be computed with free probability theory, because our random matrices are not square, and we cannot reduce the problem, as for $M(1,n)M(2,n)^*$, to a problem which involves only independent square random matrices. Thus we have to use theorem 1.6 (or corollary 1.8 if $\rho_1 = 0$) to compute the asymptotic singular law of $M(1,n) + M(2,n)$.

Proposition-Definition 2.1. *Let μ_1, μ_2 be two compactly supported symmetric probability measures on the real line.*
(a) Let, for all $n \geq 1$, $M(1,n)$, $M(2,n)$ be independent bi-unitarily invariant $q_1(n) \times q_2(n)$ random matrices, with deterministic singular values if $\rho_1 = 0$, uniformly bounded for $||.||$, and such that for all $i = 1, 2$, the symmetrization of the singular law of $M(i,n)$ converges in probability to μ_i. Then the symmetrization of the singular law of $M(1,n)+M(2,n)$ converges in probability to a symmetric probability measure on the real line, denoted $\mu_1 \boxplus_\lambda \mu_2$, which depends only on μ_1, μ_2,

and $\lambda := \lim_{n\to+\infty} q_1(n)/q_2(n)$.
(b) $\mu_1 \boxplus_\lambda \mu_2$ *is the unique symmetric measure* μ *such that for all* $r \in \mathbb{N}$ *even, the* r*-th moment of* μ *is* $\varphi_1(((a_1+a_2)(a_1+a_2)^*)^{r/2})$*, where* a_1*,* a_2 *are free with amalgamation over* $\mathcal{D}$ *elements of* $p_1\mathcal{A}p_2$*,* $(\mathcal{A}, p_1, p_2, \varphi_1, \varphi_2)$ *is a* (ρ_1, ρ_2)*-probability space, and for all* $i = 1, 2$*, for all* $r \in \mathbb{N}$ *even,* $\varphi_1((a_i a_i^*)^{r/2})$ *is the* r*-th moment of* μ_i.
(c) The support of $\mu_1 \boxplus_\lambda \mu_2$ *is contained in the sum of the convex hulls of the supports of* μ_1 *and* μ_2.

The binary operation $\boxplus_\lambda$ on the set of compactly supported symmetric probability measures, called *rectangular free convolution with ratio* λ, will be extended in section 3.9 to the set of symmetric probability measures, and the same result will stay true without any hypothesis of boundedness.

Proof. By theorem 1.6 (or theorem 1.7 if $\rho_1 = 0$), if a_1, a_2 are elements of a (ρ_1, ρ_2)-probability space $(\mathcal{A}, p_1, p_2, \varphi_1, \varphi_2)$ as in (b), then $(\widetilde{M}(1,n), \widetilde{M}(2,n))$ converges in $\mathcal{D}$-distribution in probability to (a_1, a_2). It implies that for all $k \geq 1$, the normalized trace of $[(M(1,n) + M(2,n))(M(1,n) + M(2,n))^*]^k$ converges in probability to a finite and deterministic limit m_k as n goes to infinity.

Let us first prove that the m_k's are the moments of a probability measure with support on $[0, (s_1+s_2)^2]$. It suffices to prove that for any real polynomial $P = \sum_k c_k X^k$ which is nonnegative on this interval, $\sum_k c_k m^k \geq 0$. Note that since the m_k's only depend of the $\mathcal{D}$-distribution of (a_1, a_2) one can suppose that for all n, for all $i = 1, 2$,

$$M(i,n) = U(i,n)[\delta_k^l \lambda_i(k,n)]_{\substack{1\leq k\leq q_1(n)\\ 1\leq l\leq q_2(n)}} V(i,n),$$

where $\lambda_i(1,n), \ldots, \lambda_i(q_1(n), n)$ are real numbers of the support of μ_i such that the uniform distribution on the $\lambda_i(k,n)$'s $(k = 1, \ldots, q_1(n))$ converges weakly to μ_i and $U(i,n)$, $V(i,n)$, are independent respectively $q_1(n)\times q_1(n)$, $q_2(n)\times q_2(n)$ uniform unitary random matrices. In this case, $\sum_k c_k m^k$ is the limit of the normalized trace of

$$\sum_k c_k [(M(1,n) + M(2,n))(M(1,n) + M(2,n))^*]^k.$$

By hypothesis on P, this matrix is hermitian and nonnegative for all n, hence $\sum_k c_k m^k \geq 0$.

Since on the set of probability measures on a compact interval, convergence of moments is equivalent to weak convergence, (a) and (c) are proved. Moreover, (b) follows from the first paragraph of the proof. □

Remark 2.2. *When* $\lambda = 1$*, the rectangular free convolution with ratio* λ *is the well known additive free convolution. To see it, apply the previous proposition to sequences of square matrices. We know (Theorem 4.3.11 of* [HP00]*) that independent bi-unitarily invariant square matrices are asymptotically free R-diagonal elements and that the symmetrization of the distribution of the absolute value of the sum of two free R-diagonal elements is the free convolution of the symmetrizations of the distributions of their absolute values (Proposition 3.5 of* [HL00]*), which proves that* $\boxplus_\lambda = \boxplus$ *for* $\lambda = 1$*. We will see later that the free convolution with ratio* 0 *is also related to the free convolution.*

3. The rectangular R-transform

3.1. **Relation moments-cumulants in a (ρ_1, ρ_2) probability space.** For an introduction to the general theory of cumulants in a $\mathcal{D}$-probability space and the particular case of a (ρ_1, ρ_2))probability space, we refer to section 2 of [B-G2]. Basically, if $(\mathcal{A}, p_1, p_2, \varphi_1, \varphi_2)$ is a (ρ_1, ρ_2)-probability space, then for all n, there are two linear functions $c_n^{(1)}, c_n^{(2)}$ from the n-th tensor product with amalgamation over $\mathcal{D}$ of $\mathcal{A}$ to $\mathbb{C}$ such that:

(a) for all $a \in p_1\mathcal{A}p_2$, for all $n \geq 1$,

$$\varphi_1((aa^*)^n) = \sum_{\pi \in \mathrm{NC'}(2n)} \prod_{\substack{V \in \pi \\ \min V \text{ odd}}} c_{|V|}^{(1)}(a \otimes \cdots \otimes a^*) \prod_{\substack{V \in \pi \\ \min V \text{ even}}} c_{|V|}^{(2)}(a^* \otimes \cdots \otimes a), \tag{8}$$

$$\varphi_2((a^*a)^n) = \sum_{\pi \in \mathrm{NC'}(2n)} \prod_{\substack{V \in \pi \\ \min V \text{ odd}}} c_{|V|}^{(2)}(a^* \otimes \cdots \otimes a) \prod_{\substack{V \in \pi \\ \min V \text{ even}}} c_{|V|}^{(1)}(a \otimes \cdots \otimes a^*), \tag{9}$$

where NC'$(2n)$ is the set of noncrossing partitions of $[2n]$ where all blocks have an even cardinality. This formula can easily be deduced from equations (2.2), (2,6) and (b) of section 2.2 of [B-G2],

(b) $a \in p_1\mathcal{A}p_2$, for all n even, we have $\rho_1 c_n^{(1)}(a \otimes a^* \otimes \cdots \otimes a^*) = \rho_2 c_n^{(2)}(a^* \otimes a \otimes \cdots \otimes a)$,

(c) for all $a, b \in p_1\mathcal{A}p_2$ free with amalgamation over $\mathcal{D}$, for all for all n even, we have

$$c_n^{(1)}((a+b) \otimes (a+b)^* \otimes \cdots \otimes (a+b)^*) = c_n^{(1)}(a \otimes a^* \otimes \cdots \otimes a^*) + c_n^{(1)}(b \otimes b^* \otimes \cdots \otimes b^*). \tag{10}$$

From this, we deduce, for $\lambda = \rho_1/\rho_2$:

Proposition 3.1. *For $a \in p_1\mathcal{A}p_2$, n positive integer,*

$$\varphi_1((aa^*)^n) = \sum_{\pi \in \mathrm{NC'}(2n)} \lambda^{e(\pi)} \prod_{V \in \pi} c_{|V|}(a), \tag{11}$$

where for all r even, $c_r(a)$ denotes $c_r^{(1)}(a \otimes \cdots \otimes a^*)$, and for all π, $e(\pi)$ denotes the number of blocks of π with even minimum.

3.2. **The case $\lambda = 0$, characterization of $\boxplus_0$.**

Lemma 3.2. *If $\lambda = 0$, then for all $a \in p_1\mathcal{A}p_2$, for all $n \geq 1$, $c_{2n}(a)$ is the n-th free cumulant of the element aa^* of $(p_1\mathcal{A}p_1, \varphi_1)$.*

Proof. It suffices to prove that for all $n \geq 1$,

$$\varphi_1((aa^*)^n) = \sum_{\pi \in \mathrm{NC}(n)} \prod_{V \in \pi} c_{2|V|}(a). \tag{12}$$

By (11), the left hand term is equal to $\sum_{\pi \in \mathrm{NC'}(2n), e(\pi)=0} \prod_{V \in \pi} c_{|V|}(a)$. Let us introduce the set $\mathrm{NC_d}(n)$ of non-crossing partitions π of $[2n]$ such that for all k even in $[2n]$, $k-1 \overset{\pi}{\sim} k$. Note that $\mathrm{NC_d}(2n)$ is contained in the set of elements of NC'$(2n)$ in which no block has even minimum. We claim that the inverse inclusion is true. Indeed, suppose the existence of $\pi \in \mathrm{NC'}(2n) \setminus \mathrm{NC_d}(2n)$ in which no block has even minimum. Define $k = \min\{k \in [2n]\,; k \text{ even and } k-1 \overset{\pi}{\nsim} k\}$. k cannot be the minimum of its block in π, and by minimality of k, its preceding element j in this block has to be even. Then, as π is non-crossing, the set $\{j+1, \ldots, k-1\}$ is a union of classes of π, but its cardinality is odd, which is in contradiction with $\pi \in \mathrm{NC'}(2n)$.

Now, (12) follows easily from the fact that $\mathrm{NC_d}(2n)$ is in correspondence with $\mathrm{NC}(n)$ by the order-preserving bijection from $\mathrm{NC}(n)$ onto $\mathrm{NC_d}(2n)$, that maps any noncrossing partition π to the partitionthat links two elements i, j of $[2n]$ if and only if the upper integer parts of $i/2$ and $j/2$ are linked by π. $\square$

Now we are able to give the link between the free convolution with null ratio $\boxplus_0$ and the free convolution $\boxplus$:

Proposition 3.3. *The free convolution with null ratio of two compactly supported symmetric probability measures is the unique symmetric probability whose push-forward by the square function is the free convolution of the push-forwards by the square function of the two probabilities.*

Proof. Denote, for ν probability measure, $m_n(\nu)$ its n-th moment, and $\mathfrak{K}_n(\nu)$ its n-th free cumulant. Recall ([S94]) that free cumulants of probability measures are defined by the formula

$$\forall n \geq 1, m_n(\nu) = \sum_{\pi \in \mathrm{NC}(n)} \prod_{B \in \pi} \mathfrak{K}_{|B|}(\nu), \tag{13}$$

and that the free convolution of ν_1, ν_2 compactly supported probability measures is the only probability with free cumulants $\mathfrak{K}_n(\nu_1) + \mathfrak{K}_n(\nu_2)$ $(n \geq 1)$.
Consider two symmetric compactly supported probability measures μ_1, μ_2. Denote by ν_1, ν_2 their respective push-forwards by the square function. It suffices to prove that for all $n \geq 1$, $m_{2n}(\mu_1 \boxplus_0 \mu_2) = m_n(\nu_1 \boxplus \nu_2)$, i.e. that for all $n \geq 1$,

$$m_{2n}(\mu_1 \boxplus_0 \mu_2) = \sum_{\pi \in \mathrm{NC}(n)} \prod_{B \in \pi} (\mathfrak{K}_{|B|}(\nu_1) + \mathfrak{K}_{|B|}(\nu_2)).$$

Consider $a_1, a_2 \in p_1 \mathcal{A} p_2$ free with amalgamation over $\mathcal{D}$ such that

$$\forall i = 1, 2, \forall n \geq 1, \varphi_1((a_i a_i^*)^n) = m_{2n}(\mu_i).$$

Then one has, for all $n \geq 1$,

$$\begin{aligned} m_{2n}(\mu_1 \boxplus_0 \mu_2) &= \varphi_1[((a_1 + a_2)(a_1^* + a_2^*))^n] \\ &= \sum_{\pi \in \mathrm{NC_d}(2n)} \prod_{B \in \pi} \underbrace{c_{|B|}(a_1 + a_2)}_{= c_{|B|}(a_1) + c_{|B|}(a_2) \text{ by (10)}} \qquad \text{by (11)} \end{aligned}$$

But by lemma 3.2, for all $i = 1, 2$, for all m even, $c_m(a_i)$ is the $m/2$-th free cumulant of the element $a_i a_i^*$ of $(\mathcal{A}_1, \varphi_1)$, i.e. the $m/2$-th free cumulant of ν_i. So

$$m_{2n}(\mu_1 \boxplus_0 \mu_2) = \sum_{\sigma \in \mathrm{NC}(n)} \prod_{V \in \sigma} (\mathfrak{K}_{|V|}(\nu_1) + \mathfrak{K}_{|V|}(\nu_2)) = m_n(\nu_1 \boxplus \nu_2),$$

which proves the result. $\square$

3.3. **Generating series of the cumulants of an element of $p_1 \mathcal{A} p_2$.** We are going to derive from (11) a formula which links generating series of the sequence $\varphi_1((aa^*)^n)$ and $c_{2n}(a)$. Note that such a formula could be derived from the very general Theorem 2.2.3 of [S98], but the work needed to translate this result to our context and tu put it into the form we shall use it is as long as this section.

In this section, we work in the field $\mathbb{C}((X))$ of fractions of the ring $\mathbb{C}[[X]]$ of formal power series in X with complex coefficients, endowed with the classical addition and product. Note that for any element S of $\mathbb{C}[[X]]$ with null constant term, $F \mapsto F \circ S$ is a well defined operation

on $\mathbb{C}((X))$. If moreover, the coefficient of X in S is non null, then S has a left and right inverse for $\circ$, denoted by $S^{<-1>}$.

Lemma 3.4. *Consider a sequence $(c_{2n})_{n\in\mathbb{N}^*}$ of complex numbers. Define the sequence $(m_{2n}^{(e)})_{n\in\mathbb{N}}$ by $m_0^{(e)}=1$ and for each $n\in\mathbb{N}^*$,*

$$m_{2n}^{(e)} = \sum_{\pi\in \mathrm{NC'}(2n)} \lambda^{e(\pi)} \prod_{B\in\pi} c_{|B|}.$$

Define the formal power series

$$C(X)=\sum_{n\geq 1} c_{2n}X^n, \qquad M^{(e)}(X)=\sum_{n\geq 1} m_{2n}^{(e)}X^n.$$

Then we have $C[X(\lambda M^{(e)}(X)+1)(M^{(e)}(X)+1)]=M^{(e)}(X)$.

Proof. *Step I.* Define, for n positive integer and $\pi\in \mathrm{NC'}(2n)$, $o(\pi)$ to be the number of blocks of π with odd minimum. Then we have

$$\lambda m_{2n}^{(e)} = \sum_{\pi\in \mathrm{NC'}(2n)} \lambda^{o(\pi)} \prod_{B\in\pi} c_{|B|}.$$

Indeed, if c denotes the cycle $(2n\to 2n-1\to\cdots\to 2\to 1)$ of $[2n]$, then c induces a permutation of $\mathrm{NC'}(2n)$ by

$$R:\pi\in \mathrm{NC'}(2n)\mapsto R(\pi):=\{c(B)\,;\,B\in\pi\}.$$

For example, for $\pi=\{\{1,4,5,6\},\{2,3\}\}$, $R(\pi)=\{\{3,4,5,6\},\{1,2\}\}$. For $\pi\in\mathrm{NC'}(2n)$, for B block of π, $\min c(B)$ is $(\min B)-1$ if $\min B>1$, and the second element of B minus 1 in the other case. Note that in the first case, $\min B$ and $\min c(B)$ have different parities, whereas in the second one, they have the same parities, by definition of $\mathrm{NC'}(2n)$. Hence we have $o(R(\pi))=e(\pi)+1$. Thus

$$\lambda m_{2n}^{(e)} = \sum_{\pi\in \mathrm{NC'}(2n)} \lambda^{e(\pi)+1} \prod_{B\in\pi} c_{|B|} = \sum_{\pi\in \mathrm{NC'}(2n)} \lambda^{o(R(\pi))} \prod_{B\in\pi} c_{|B|} = \sum_{\pi\in \mathrm{NC'}(2n)} \lambda^{o(\pi)} \prod_{B\in\pi} c_{|B|}.$$

Step II. Thus we have to prove that $C[X(M^{(o)}(X)+1)(M^{(e)}(X)+1)]=M^{(e)}(X)$, where $M^{(o)}(X)=\sum_{n\geq 1} m_{2n}^{(o)}X^n$, and $m_{2n}^{(o)} = \sum_{\pi\in \mathrm{NC'}(2n)} \lambda^{o(\pi)} \prod_{B\in\pi} c_{|B|}$.

Define, for each positive n and each $j\in[n]$, the "decomposition map" D_j, from the set of partitions of $\mathrm{NC'}(2n)$ in which the class of 1 has cardinality $2j$ to $\bigcup_{\substack{l\in\mathbb{N}^{2j}\\ l_1+\cdots+l_{2j}=n-j}} \mathrm{NC'}(2l_1)\times\cdots\times \mathrm{NC'}(2l_{2j})$ (the set $\mathrm{NC'}(0)$ is considered as a singleton, on which e and o have value 0), which maps $\pi\in NC'(n)$ to $(\pi_1,\ldots,\pi_{2j})$ defined in the following way: if $\{1=k_1<\ldots<k_{2j}\}$ is the class of 1 in π, then for each r, let π_r be the restriction of π to the interval $]k_r,k_{r+1}[$, with $k_{2j+1}=2n+1$. It is easy to see that the map D_j is well defined (recall that for all r, k_r and k_{r+1} have opposite parity), and that it is a bijection. Moreover, we have

$$\begin{aligned} e(\pi) &= o(\pi_1)+e(\pi_2)+\cdots+o(\pi_{2j-1})+e(\pi_{2j}),\\ o(\pi) &= 1+e(\pi_1)+o(\pi_2)+\cdots+e(\pi_{2j-1})+o(\pi_{2j}). \end{aligned}$$

We denote by $\langle X^n\rangle P$ the coefficient of X^n in a formal power series P in X. We have to show that for each $n\geq 1$,

$$\langle X^n\rangle C[X(M^{(o)}(X)+1)(M^{(e)}(X)+1)]=m_{2n}^{(e)}.$$

We have

$$\begin{aligned}
&\langle X^n\rangle C[X(M^{(o)}(X)+1)(M^{(e)}(X)+1)]\\
=\;& \sum_{j=1}^{n} c_{2j}\langle X^n\rangle[X^j(M^{(o)}(X)+1)^j(M^{(e)}(X)+1)^j]\\
=\;& \sum_{j=1}^{n} c_{2j}\langle X^{n-j}\rangle[(M^{(o)}(X)+1)^j(M^{(e)}(X)+1)^j]\\
=\;& \sum_{j=1}^{n} c_{2j} \sum_{\substack{l\in\mathbb{N}^{2j}\\ l_1+\cdots+l_{2j}=n-j}} m^{(o)}_{2l_1}m^{(e)}_{2l_2}\cdots m^{(o)}_{2l_{2j-1}}m^{(e)}_{2l_{2j}}\\
=\;& \sum_{j=1}^{n} c_{2j} \sum_{\substack{l\in\mathbb{N}^{2j}\\ l_1+\cdots+l_{2j}=n-j}} \sum_{\substack{\pi_1\in \mathrm{NC'}(2l_1)\\ \vdots\\ \pi_{2j}\in \mathrm{NC'}(2l_{2j})}} \lambda^{o(\pi_1)+e(\pi_2)+\cdots+e(\pi_{2j})} \prod_{r=1}^{2j}\prod_{B\in\pi_r} c_{|B|}
\end{aligned}$$

Hence the preliminaries about the bijections D_j tell us that

$$\langle X^n\rangle C[X(M^{(o)}(X)+1)(M^{(e)}(X)+1)] = \sum_{\pi\in \mathrm{NC'}(2n)} \lambda^{e(\pi)} \prod_{B\in\pi} c_{|B|} = m^{(e)}_{2n}.$$

□

Let us introduce two power series that will play a role in the computation of the generating series of the cumulants:

$$\begin{aligned}
T(X) &= (\lambda X+1)(X+1),\\
U(X) = (T-1)^{<-1>} &= \sum_{n\geq 1} \frac{(-1)^{n-1}\lambda^{n-1}\binom{2n}{n}}{2(\lambda+1)^{2n-1}(2n-1)}X^n.
\end{aligned}$$

$U(z)$ is the power expansion of $\frac{-\lambda-1+\left[(\lambda+1)^2+4\lambda z\right]^{1/2}}{2\lambda}$ (:= z if $\lambda=0$), where $z\mapsto z^{1/2}$ is the analytic version of the square root on the complement of the real non positive half line such that $1^{1/2}=1$.

Theorem 3.5. *Consider $a\in p_1\mathcal{A}p_2$ and define the formal power series $M(X)=\sum_{n\geq 1} m_{2n}(a)X^n$ and $C(X)=\sum_{n\geq 1} c_{2n}(a)X^n$. Then we have*

$$C = U\left(\frac{X}{(X\times(T\circ M))^{<-1>}}-1\right).$$

Proof. By proposition 3.1 and the previous lemma, we have $C\circ(X\times(T\circ M))=M$. Hence $T\circ C\circ(X\times(T\circ M))=T\circ M$. Dividing by $T\circ M$ on both sides, one gets $\frac{T\circ C}{X}\circ(X\times(T\circ M))=\frac{1}{X}$. Then by inversion, $\frac{X}{T\circ C}\circ(X\times(T\circ M))=X$. At last, $\frac{X}{T\circ C}=(X\times(T\circ M))^{<-1>}$, which implies $T\circ C=\frac{X}{(X\times(T\circ M))^{<-1>}}$, and then $C=U\left(\frac{X}{(X\times(T\circ M))^{<-1>}}-1\right)$. □

3.4. **Rectangular R-transform of compactly supported symmetric measures.** From now on, we will denote by $(\cdot)^{1/2}$ (resp. $\sqrt{\cdot}$) the analytic version of the square root on the complementary of the real non positive (resp. nonnegative) half line such that $1^{1/2}=1$ (resp. $\sqrt{-1}=i$).

Consider a symmetric compactly supported probability measure μ. For $n \geq 0$, denote by $m_n(\mu)$ its n-th moment and by $c_{2n}(\mu)$ the number $c_{2n}(a)$ with $a \in p_1 \mathcal{A} p_2$, where $\mathcal{A}$ is a (ρ_1, ρ_2)-probability space, such that for all $k \geq 1$, $\varphi_1((aa^*)^k) = m_{2k}(a)$.

Note that the generating series of the moments $M_\mu(X) = \sum_{n\geq 1} m_{2n}(\mu) X^n$ of μ is the power expansion of

$$M_\mu(z) = \frac{1}{\sqrt{z}} G_\mu(\frac{1}{\sqrt{z}}) - 1 \qquad (z \in \mathbb{C}\backslash[0, +\infty)),$$

where G_μ is the *Cauchy transform* of μ on the upper half plane: $G_\mu(z) = \int_{t\in\mathbb{R}} \frac{\mathrm{d}\mu(t)}{z-t}$. Hence $X\times(T \circ M)$ is the power expansion of

$$H_\mu(z) := zT \circ M_\mu(z) = \lambda G_\mu\left(\frac{1}{\sqrt{z}}\right)^2 + (1-\lambda)\sqrt{z} G_\mu\left(\frac{1}{\sqrt{z}}\right) \qquad (z \in \mathbb{C}\backslash[0, +\infty)),$$

that we shall call the *rectangular Cauchy transform with ratio* λ of μ. Since μ is compactly supported, from the power expansion of the Cauchy transform G_μ, we know that H_μ is analytic in a neighborhood of zero, that $H_\mu(0) = 0$ and that $H'_\mu(0) = 1$. It implies that we can invert the function H_μ in a neighborhood of zero and that the inverse function H_μ^{-1} has a power series expansion given by $(X\times(T \circ M_\mu))^{<-1>}$. Note that $\frac{z}{H_\mu^{-1}(z)} - 1$ is analytic in a neighborhood of zero, with value zero in zero. At last, by theorem 3.5, $C_\mu(X) := \sum_{n\geq 1} c_{2n}(a) X^n$ is the power expansion of

$$C_\mu(z) := U\left(\frac{z}{H_\mu^{-1}(z)} - 1\right) \quad \text{(well defined for } z \text{ small enough)},$$

where $U(\cdot)$ is the analytic function on a neighborhood of zero $U(z) = \frac{-\lambda-1+\left[(\lambda+1)^2+4\lambda z\right]^{1/2}}{2\lambda}$ ($:= z$ if $\lambda = 0$).

Moreover, by (10), we know that $\mu \mapsto C_\mu$ linearizes the rectangular convolution with ratio λ: for all μ, ν symmetric probability measures with compact support,

$$C_{\mu \boxplus_\lambda \nu}(z) = C_\mu(z) + C_\nu(z) \quad (z \text{ small enough}). \tag{14}$$

3.5. **Rectangular R-transform of unbounded measures: definition.** The *Cauchy transform* and the *rectangular Cauchy transform with ratio* λ of symmetric probability measures with unbounded support are defined with the same formulas as in the case of a compactly supported measures (see the previous paragraph).

Proposition 3.6. *Let A be a set of symmetric probability measures on the real line. Then the following assertions are equivalent*

(i) *A is tight,*
(ii) *for every $0 < \theta < \pi$, $\lim\limits_{\substack{z\to 0 \\ |\arg z - \pi| < \theta}} \frac{1}{z} H_\mu(z) = 1$ uniformly in $\mu \in A$,*
(iii) $\lim\limits_{\substack{x\to 0 \\ x\in(-\infty,0)}} \frac{1}{x} H_\mu(x) = 1$ *uniformly in $\mu \in A$.*

Proof. $(i) \Rightarrow (ii)$ follows from the well known fact ([BV93]) that if A is tight,

$$\forall 0 < \theta < \pi, \quad \lim_{\substack{z\to 0 \\ |\arg z - \frac{\pi}{2}| < \frac{\theta}{2}}} \frac{1}{z} G_\mu(\frac{1}{z}) = 1 \quad \text{uniformly in } \mu \in A.$$

$(ii) \Rightarrow (iii)$ is clear. Suppose (iii). Consider $\varepsilon > 0$. Take $\eta > 0$ such that $\forall u \in [0,1], |\lambda u^2 + (1-\lambda)u - 1| < \eta \Rightarrow |u-1| < \varepsilon/2$. Take $x \in (-\infty, 0)$ such that for all $\mu \in A$, $|\frac{1}{x} H_\mu(x) - 1| < \eta$. Put $\sqrt{x} = iy$. Then for all $\mu \in A$,

$$|\frac{1}{iy} G_\mu(\frac{1}{iy}) - 1| < \varepsilon/2, \text{ i.e. } \int_t \frac{t^2 y^2}{t^2 y^2 + 1} \mathrm{d}\mu(t) < \varepsilon/2.$$

Let M be such that $\forall t > M$, $\frac{t^2 y^2}{t^2 y^2 + 1} \geq \frac{1}{2}$. Then for all $\mu \in A$, $\mu([-M, M]^c) \leq 2 \int_t \frac{t^2 y^2}{t^2 y^2 + 1} \mathrm{d}\mu(t) < \varepsilon$. □

Define, for $\alpha \in (0, \pi)$, $\beta > 0$, $\Delta_{\alpha,\beta}$ to be the set of complex numbers z such that $|\arg z - \pi| < \alpha$ and $|z| < \beta$.

Let $\mathcal{H}$ be the set of functions f which are analytic in a domain $\mathcal{D}_f$ such that for all $\alpha \in (0, \pi)$, there exists β positive such that

$$\Delta_{\alpha,\beta} \subset \mathcal{D}_f.$$

A family $(f_a)_{a \in A}$ of functions of $\mathcal{H}$ is said to be *uniform* if for all $\alpha \in (0, \pi)$, there exists β positive such that

$$\forall a \in A, \quad \Delta_{\alpha,\beta} \subset \mathcal{D}_{f_a}.$$

The following theorem has already been used, in other forms, to define Voiculescu's R- and S-transforms (see paragraph 5 of [BV93]). Its proof relies on Rouché theorem.

Theorem 3.7. *Let $(H_a)_{a \in A}$ be a uniform family of functions of $\mathcal{H}$ such that for every $\alpha \in (0, \pi)$,*

$$\lim_{\substack{z \to 0 \\ |\arg z - \pi| < \alpha}} \frac{H_a(z)}{z} = 1 \text{ uniformly in } a \in A.$$

Then there exists a uniform family $(F_a)_{a \in A}$ of functions of $\mathcal{H}$ such that for every $\alpha \in (0, \pi)$,

$$\lim_{\substack{z \to 0 \\ |\arg z - \pi| < \alpha}} \frac{F_a(z)}{z} = 1 \text{ uniformly in } a \in A,$$

and there exists β positive such that

$$\forall a \in A, \quad H_a \circ F_a = F_a \circ H_a = I_d \text{ on } \Delta_{\alpha,\beta}.$$

Moreover, the family $(F_a)_{a \in A}$ is unique in the following sense: if a family $(\tilde{F}_a)_{a \in A}$ of functions of $\mathcal{H}$ satisfies the same conditions, then for all $\alpha \in (0, \pi)$, there exists β positive such that

$$\forall a \in A, \quad F_a = \tilde{F}_a \text{ on } \Delta_{\alpha,\beta}.$$

For every symmetric probability measure on the real line μ, let us define the *rectangular R-transform R_μ with ratio λ* of μ by

$$C_\mu(z) = U\left(\frac{z}{H_\mu^{-1}(z)} - 1\right),$$

where H_μ^{-1} is defined by the previous theorem and the function U is the one defined at the end of section 3.4.

One can summarize the different steps of the construction of the rectangular R-transform with ratio λ in the following chain

$$\underset{\substack{\text{sym. prob.}\\ \text{measure}}}{\mu} \longrightarrow \underset{\substack{\text{Cauchy}\\ \text{transf.}}}{G_\mu} \longrightarrow H_\mu(z) = \lambda G_\mu\left(\frac{1}{\sqrt{z}}\right)^2 + (1-\lambda)\sqrt{z}G_\mu\left(\frac{1}{\sqrt{z}}\right) \longrightarrow$$
$$\underset{\text{rect. } R\text{-transf. with ratio } \lambda}{C_\mu(z) = U\left(\frac{z}{H_\mu^{-1}(z)} - 1\right)}.$$

3.6. **The special cases $\lambda = 0$ and $\lambda = 1$.** Note that the rectangular R-transform with ratio 1 (resp. 0), for a symmetric distribution μ, is linked to the Voiculescu transform φ_μ of μ by the relation $C_\mu(z) = \sqrt{z}\varphi_\mu(1/\sqrt{z})$ (resp. $C_\mu(z) = z\varphi_\rho(1/z)$, where ρ is the push-forward of μ by the function $t \to t^2$) (see paragraph 5 of [BV93] for the construction of the Voiculescu transform).

3.7. **Remark about the characterization of R-transforms.** The definition of the rectangular R-transform of a probability measure μ presents many analogies with the definition of its Voiculescu transform φ_μ (see [BV93]). So it seems natural to state, as the authors of [BV93] did for the Voiculescu transform (Proposition 5.6), a characterization of the functions that are the rectangular R-transform of a symmetric probability measure. The characterization of the Voiculescu transform is based on the fact that for every probability measure μ, the inverse of $z + \varphi_\mu(z)$ extends to a Pick function (i.e. an analytic function defined on the upper half plane, whose imaginary part does not take negative values), and on the fact that any Pick function equivalent to z at infinity is of the type $1/G_\mu$, so its inverse is $z + \varphi_\mu(z)$. But unless $\lambda = 0$ or 1, Pick functions do not appear in an analogous place in the definition of the rectangular R-transform, so we cannot proceed similarly to characterize rectangular R-transforms.

Furthermore, the Lévy Kinchine formula for $\boxplus_\lambda$ (see [B-G1]), compared with theorem 5.10 of [BV93]) will state that when μ is $\boxplus_\lambda$-infinitely divisible, there exists a unique symmetric $\boxplus$-infinitely divisible distribution ν such that $C_\mu(z) = \sqrt{z}\varphi_\nu(1/\sqrt{z})$. So the question of the characterization of rectangular R-transforms joins another question: can we extend the correspondence $\mu \leftrightarrow \nu$ to a bijective correspondence from the set of symmetric distributions into itself such that for all μ, $C_\mu(z) = \sqrt{z}\varphi_\nu(1/\sqrt{z})$. The analytic functions f on $\Delta_{\alpha,\beta}$ of the type $f(z) = \sqrt{z}\varphi_\nu(1/\sqrt{z})$, with ν symmetric probability measure, are characterized by:

(i) $f(\bar{z}) = \overline{f(z)}$,
(ii) $\lim_{z\to 0} f(z) = 0$,
(iii) for all n and all $z_1, \ldots, z_n$, the matrix $\left[\frac{\overline{\sqrt{z_j}} - \sqrt{z_k}}{\overline{\sqrt{z_j}}(1+f(z_k)) - \sqrt{z_k}\overline{(1+f(z_j))}}\right]_{k,j=1}^n$ is positive.

But nothing allows us to claim that the rectangular R-transform of any symmetric distribution satisfies (iii), and that every function that satisfies (i), (ii), and (iii) is the rectangular R-transform of a symmetric distribution.

3.8. **Properties of the rectangular R-transform.**

Theorem 3.8 (Injectivity of the rectangular R-transform)**.** *If the rectangular R-transforms with ratio λ of two symmetric probability measures coincide on a neighborhood of 0 in $(-\infty, 0)$, then the measures are equal.*

Proof. If the rectangular R-transforms with ratio λ of two symmetric probability measures μ, ν coincide on a neighborhood of 0 in $(-\infty, 0)$, then by uniqueness of analytic continuation, they coincide on a $\Delta_{\alpha,\beta}$, and so $H_\mu = H_\nu$ on this set, and $zH_\mu(z) = zH_\nu(z)$ on this set. So there exists $M > 0$ such that for all $y > M$,

$$\lambda(iy2G_\mu(iy))^2 + (1-\lambda)iyG_\mu(iy) = \lambda(iy2G_\nu(iy))^2 + (1-\lambda)iyG_\nu(iy).$$

But if ρ is a symmetric probability measure, $iyG_\mu(iy) = \int_{t\in\mathbb{R}} \frac{y^2}{y^2+t^2} \mathrm{d}\rho(t) \in (0,1]$. So, by injectivity of $u \mapsto \lambda u^2 + (1-\lambda)u$ on $(0,1]$, $G_\mu(iy) = G_\nu(iy)$ for $y > 1$, and then, by analycity of the Cauchy transform, $G_\mu = G_\nu$, and by injectivity of the Cauchy transform (see [A61] or section 3.1 of [HP00]) $\mu = \nu$. □

The following remark gives a practical way to derive any symmetric probability measure on the real line μ from C_μ.

Remark 3.9 (How to compute μ when we know C_μ ?)**.** *First, we have $z/H_\mu^{-1}(z) = T(C_\mu(z))$, for $z \in \mathbb{C}\backslash\mathbb{R}^+$ small enough. From this, we can compute $H_\mu(z)$ for $z \in \mathbb{C}\backslash\mathbb{R}^+$ small enough. Then we can use the equation, for $z \in \mathbb{C}\backslash\mathbb{R}^+$,*

$$\frac{1}{z}H_\mu(z) = \lambda\left(\frac{1}{\sqrt{z}}G_\mu(\frac{1}{\sqrt{z}})\right)^2 + (1-\lambda)\frac{1}{\sqrt{z}}G_\mu(\frac{1}{\sqrt{z}}).$$

Moreover, when $z \in \mathbb{C}\backslash\mathbb{R}^+$ is small enough, $1/\sqrt{z}$ is large and in $\mathbb{C}^-$, so $\frac{1}{\sqrt{z}}G_\mu\left(\frac{1}{\sqrt{z}}\right)$ is closed to 1. $\frac{1}{z}H_\mu(z)$ is also closed to 1, and for h, g complex numbers closed to 1,

$$h = \lambda g^2 + (1-\lambda)g \Leftrightarrow g = V(h), \text{ with } V(z) = \frac{\lambda - 1 + ((\lambda-1)^2 + 4\lambda z)^{\frac{1}{2}}}{2\lambda} = U(z-1) + 1.$$

So one has, for $z \in \mathbb{C}\backslash\mathbb{R}^+$ small enough, $\frac{1}{\sqrt{z}}G_\mu(\frac{1}{\sqrt{z}}) = V(\frac{H_\mu(z)}{z})$, which allows to recover μ

The following lemma is an easy consequence of proposition 3.6 and theorem 3.7.

Lemma 3.10 (Tightness and rectangular R-transform)**.** *Let A be a set of symmetric probability measures. Then we have equivalence between*

(i) *A is tight,*
(ii) *for any $0 < \alpha < \pi$, $\lim\limits_{\substack{z\to 0 \\ |\arg z - \pi| < \alpha}} C_\mu(z) = 0$ uniformly in $\mu \in A$,*
(iii) $\lim\limits_{\substack{x\to 0 \\ x\in(-\infty,0)}} R_\mu(x) = 0$ *uniformly in $\mu \in A$.*

Theorem 3.11 (Paul Lévy's theorem for rectangular R-transform)**.** *Let (μ_n) be a sequence of symmetric probability measures. Then we have equivalence between:*

(i) *(μ_n) converges weakly to a symmetric probability measure;*
(ii) *there exists α, β such that*
  (a) $\lim\limits_{\substack{z\to 0 \\ |\arg z - \pi| < \alpha}} C_{\mu_n}(z) = 0$ *uniformly in n,*
  (b) *the sequence (C_{μ_n}) converges uniformly on every compact set of $\Delta_{\alpha,\beta}$;*
(iii) (a) $\lim\limits_{\substack{x\to 0 \\ x\in(-\infty,0)}} C_{\mu_n}(x) = 0$ *uniformly in n,*
  (b) *there exists $\beta > 0$ such that the sequence (C_{μ_n}) converges pointwise on $[-\beta, 0)$.*

Moreover, in this case, denoting by μ the weak limit of (μ_n), for every α, there exists β such that the sequence (C_{μ_n}) converges uniformly to C_μ on every compact set of $\Delta_{\alpha,\beta}$.

Proof. $(i) \Rightarrow (ii)$: suppose that (μ_n) converges weakly to μ. Then by the previous lemma, we have (a) of (ii). So there exists $\beta > 0$ such that on $\Delta_{\alpha,\beta}$, for all n, $|C_{\mu_n}| \leq 1$. So, by Montel's theorem, it suffices to show that the only possible limit, for uniform convergence on every compact, of any subsequence of (C_{μ_n}) is C_μ. Let C be an analytic function on $\Delta_{\alpha,\beta}$ such that a subsequence $\left(C_{\mu_{k_n}}\right)$ of (C_{μ_n}) converges uniformly to C on every compact of $\Delta_{\alpha,\beta}$. We know (see [A61] or section 3.1 of [HP00]) that the sequence (G_{μ_n}) converges uniformly on every compact of the upper half plane to G_μ. So the sequence (H_{μ_n}) converges uniformly on every compact of the complement of $[0, +\infty)$ to H_μ. Since $\lim_{z\to 0} C(z) = 0$, to prove $C = C_\mu$, it suffices to prove that

$$\begin{aligned} (\lambda C + 1)(C + 1) &= (\lambda C_\mu + 1)(C_\mu + 1). \\ \text{But } (\lambda C_\mu(z) + 1)(C_\mu(z) + 1) &= \frac{z}{H_\mu^{-1}(z)} \end{aligned}$$

So it suffices to prove that

$$\frac{z}{(\lambda C(z) + 1)(C(z) + 1)} = H_\mu^{-1}(z),$$

so by theorem 3.7, it suffices to prove that

$$H_\mu\left(\frac{z}{(\lambda C(z) + 1)(C(z) + 1)}\right) = z.$$

We have

$$\begin{aligned} & \left|H_\mu\left[\tfrac{z}{(\lambda C(z)+1)(C(z)+1)}\right] - z\right| \\ = & \left|H_\mu\left[\tfrac{z}{(\lambda C(z)+1)(C(z)+1)}\right] - H_{\mu_{k_n}}\left[\frac{z}{\left(\lambda C_{\mu_{k_n}}(z)+1\right)\left(C_{\mu_{k_n}}(z)+1\right)}\right]\right| \\ \leq & \underbrace{\left|H_\mu\left[\frac{z}{(\lambda C(z) + 1)(C(z) + 1)}\right] - H_\mu\left[\frac{z}{\left(\lambda C_{\mu_{k_n}}(z) + 1\right)\left(C_{\mu_{k_n}}(z) + 1\right)}\right]\right|}_{(1)} \\ & + \underbrace{\left|H_\mu\left[\frac{z}{\left(\lambda C_{\mu_{k_n}}(z) + 1\right)\left(C_{\mu_{k_n}}(z) + 1\right)}\right] - H_{\mu_{k_n}}\left[\frac{z}{\left(\lambda C_{\mu_{k_n}}(z) + 1\right)\left(C_{\mu_{k_n}}(z) + 1\right)}\right]\right|}_{(2)} \end{aligned}$$

By continuity of H_μ, (1) tends to zero when n tends to infinity, and, since $H_{\mu_{k_n}}$ converges uniformly to H_μ on every compact, (2) tends to zero. So $H_\mu\left(\frac{z}{(\lambda C(z)+1)(C(z)+1)}\right) = z$.

$(ii) \Rightarrow (iii)$ is clear.

$(iii) \Rightarrow (i)$: suppose (iii). Then by $(i) \Rightarrow (ii)$, every limit of a subsequence of (μ_n) has a rectangular R-transform equal to the pointwise limit of (C_{μ_n}) on $[\beta, 0)$. By uniqueness of analytic continuation, all the limits of subsequences of (μ_n) have the same rectangular R-transform, so, by injectivity of this transform (previous theorem), there cannot be more than one limit of subsequence of (μ_n). As by (a) of (iii), the set $\{\mu_n\,;\, n \in \mathbb{N}\}$ is tight, (μ_n) converges weakly to a symmetric probability measure. □

3.9. **Rectangular convolution of measures with unbounded support.**

Theorem 3.12. *The binary operation $\boxplus_\lambda$ defined on the set of compactly supported symmetric probability measures in section 2, extends in a unique way to a commutative, associative, and continuous (with respect to the weak convergence) binary operation on the set of symmetric probability measures on the real line. This operation is called the rectangular free convolution with ratio λ. Moreover, for all μ, ν symmetric probability measures one has $C_{\mu \boxplus_\lambda \nu} = C_\mu + C_\nu$.*

Proof. Let μ, ν be symmetric probability measures. If (μ_n) (resp. (ν_n)) is a sequence of compactly supported symmetric probability measures that converges weakly to μ (resp. ν), then by theorem 3.11 and (14), the sequence $(\mu_n \boxplus_\lambda \nu_n)$ converges to a measure whose rectangular R-transform is $C_\mu + C_\nu$ (thus by injectivity of the rectangular R-transform, this measure does not depend on the choice of the sequences (μ_n) and (ν_n) and is equal to $\mu \boxplus_\lambda \nu$ when μ and ν are compactly supported). Note that $C_{\mu \boxplus_\lambda \nu} = C_\mu + C_\nu$ stays true for all symmetric probability measures μ and ν. Moreover, this equation shows that $\boxplus_\lambda$ is a commutative, associative, and continuous binary operation on the set of symmetric probability measures. □

Now we can extend the proposition 2.1 to the case where μ_1, μ_2 are not compactly supported:

Theorem 3.13. *Consider sequences $q_1(n), q_2(n)$ such that*

$$q_1(n) \leq q_2(n), \quad q_1(n) \underset{n\to\infty}{\longrightarrow} +\infty, \quad q_1(n)/q_2(n) \underset{n\to\infty}{\longrightarrow} \lambda \in [0,1].$$

Let, for all n, A_n, B_n be independent $q_1(n) \times q_2(n)$ random matrices, one of them being bi-unitarily invariant, such that the symmetrizations of the singular laws of A_n, B_n converge in probability to probability measures μ, ν. Then the symmetrization of the singular law of $A_n + B_n$ converges in probability to $\mu \boxplus_\lambda \nu$.

To prove the theorem, we will need the following lemma, for which we need to introduce an extended functional calculus. When F is a real Borel function on the real line, for any $q_1(n) \times q_2(n)$ matrix C, we define $F(C)$ to be the matrix

$$U [\delta_i^j F(h_i)]_{\substack{1 \leq i \leq q_1(n) \\ 1 \leq j \leq q_2(n)}} V,$$

where $h_1, \ldots, h_{q_1(n)} \geq 0$, and U, V are respectively $q_1(n) \times q_1(n)$, $q_2(n) \times q_2(n)$ unitary matrices such that

$$C = U [\delta_i^j h_i]_{\substack{1 \leq i \leq q_1(n) \\ 1 \leq j \leq q_2(n)}} V.$$

Note that $(F+G)(C) = F(C) + G(C)$ and $F(C)G(C)^* = 0$ when $FG = 0$. For any probability measure σ, we will denote by $F(\sigma)$ (for example σ^2, $\sigma^{1/2}$, $|\sigma| \ldots$) the push-forward of σ by F. We denote by μ_H the spectral measure of an hermitian matrix H. Recall that we denote by $\tilde{\sigma}$ the symmetrization of any probability measure σ. For example, the symmetrization of the singular law of a rectangular matrix A will be denoted by $\tilde{\mu}_{|A|}$.

Lemma 3.14. *Let (C_n) be a sequence of $q_1(n) \times q_2(n)$ random matrices such that $\tilde{\mu}_{|C_n|}$ converges in probability to a probability measure σ. Then, for any odd real function F continuous at σ-almost every point of $\mathbb{R}$ and such that $F(\mathbb{R}^+) \subset \mathbb{R}^+$, $\tilde{\mu}_{|F(C_n)|}$ converges in probability to $F(\sigma)$.*

Proof. $F(\sigma)$ is symmetric and the symmetrization is a continuous operation, so it suffices to prove that $\mu_{|F(C_n)|}$ converges in probability to $|F(\sigma)| = F(|\sigma|)$. But $\mu_{|F(C_n)|} = F(\mu_{|C_n|})$, $\mu_{|C_n|}$ converges in probability to $|\sigma|$, so the conclusion follows from the fact that the function $\sigma \mapsto F(\sigma)$ on the set of probability measures on the real line is weakly continuous at $|\sigma|$ because F is continuous at $|\sigma|$-almost every point of the real line ([B68]). □

Proof of the theorem. First of all, for U_n, V_n independent $q_1(n)\times q_1(n)$, $q_2(n)\times q_2(n)$ uniform random unitary matrices which are independent from $\{A_n, B_n\}$, the random matrices $U_nA_nV_n, U_nB_nV_n$ are independent (it is a property of the Haar measure on compact groups), have the same singular values as respectively A_n, B_n, both of them are bi-unitarily invariant and the singular values of $U_nA_nV_n + U_nB_nV_n = U_n(A_n + B_n)V_n$ are the same ones as the ones of $A_n + B_n$. Hence one can suppose that both A_n, B_n are bi-unitarily invariant.

The function that maps a probability measure σ on $\mathbb{R}^+$ to the symmetrization $\widetilde{\sigma^{1/2}}$ of its push-forward by $x \to x^{1/2}$ is continuous. So it suffices to prove that the push-forward of the spectral law of $|A_n + B_n|$ by $x \to x^2$, i.e. the spectral law of $M_n := (A_n + B_n)(A_n + B_n)^*$ converges in probability to $(\mu \boxplus_\lambda \nu)^2$. We can define a distance on the set of probability measures on the real line with the Cauchy transform by $(\sigma_1, \sigma_2) \mapsto \sup\{|G_{\sigma_1}(z) - G_{\sigma_2}(z)| \; ; \Im z \geq 1\}$. This distance defines the topology of weak convergence ([A61], [PL02]). The Cauchy transform of the spectral distribution of an hermitian matrix M is the normalized trace of its resolvant $\mathfrak{R}_z(M) = (z - M)^{-1}$. So the spectral distribution of a sequence (X_n) of $q_1(n)\times q_1(n)$ hermitian random matrices converges in probability to a probability measure σ on the real line if and only if for each $\varepsilon > 0$, the probability of the event

$$\sup_{\Im z \geq 1} |\operatorname{tr} \mathfrak{R}_z (X_n) - G_\sigma(z)| > \varepsilon$$

tends to zero as n tends to infinity. Choose $\varepsilon > 0$. We will show that

$$\mathrm{P}\{ \sup_{\Im z \geq 1} |\operatorname{tr}(\mathfrak{R}_z(M_n)) - G_{(\mu \boxplus_\lambda \nu)^2}(z)| \; > \; \varepsilon\} \stackrel{n\to\infty}{\longrightarrow} 0, \tag{15}$$

where P designs the probability measure of the probability space where the random matrices are defined. Let us define, for $t > 0$, the function F_t on the real line by

$$F_t : x \mapsto \begin{cases} x & \text{if } |x| \leq t, \\ 0 & \text{else.} \end{cases}$$

For every probability measure σ, $F_t(\sigma)$ converges weakly to σ when t tends to infinity, so, by continuity of $\boxplus_\lambda$, there exists $t \in (0, +\infty)$ such that t and $-t$ are not atoms of the measures μ and ν, and such that

$$\sup_{\Im z \geq 1} |G_{(F_t(\mu) \boxplus_\lambda F_t(\nu))^2}(z) - G_{(\mu \boxplus_\lambda \nu)^2}(z)| \; < \; \frac{\varepsilon}{3}, \tag{16}$$

$$\mu(\mathbb{R}\backslash[-t-1, t+1]) + \nu(\mathbb{R}\backslash[-t-1, t+1]) \; < \; \frac{\varepsilon}{18}. \tag{17}$$

We will now use the notations $F_t = F$, $G(x) = x - F(x)$ and $M_{n,t} = (F(A_n) + F(B_n))(F(A_n) + F(B_n))^*$. By triangular inequality and (16), we have

$$\begin{aligned} & \mathrm{P}\{ \sup_{\Im z \geq 1} |\operatorname{tr}(\mathfrak{R}_z(M_n)) - G_{(\mu \boxplus_\lambda \nu)^2}(z)| \; > \; \varepsilon\} \\ \leq \; & \mathrm{P}\{ \sup_{\Im z \geq 1} |\operatorname{tr}[\mathfrak{R}_z(M_n) - \mathfrak{R}_z(M_{n,t})]| \; > \; \tfrac{\varepsilon}{3}\} \\ + \; & \mathrm{P}\{ \sup_{\Im z \geq 1} |\operatorname{tr} \mathfrak{R}_z(M_{n,t}) - G_{(F(\mu) \boxplus_\lambda F(\nu))^2}(z)| \; > \; \tfrac{\varepsilon}{3}\} \end{aligned} \tag{18}$$

To treat the first term of the right hand side of (18), recall that for any $q_1(n)\times q_1(n)$ matrix T, $|\operatorname{tr} T| \leq \frac{1}{q_1(n)} ||T|| \operatorname{rg} T$. The operator norm $||\mathfrak{R}_z(M_n) - \mathfrak{R}_z(M_{n,t})||$ is not less than 2, because $\Im z \geq 1$. Moreover, due to the equation

$$\mathfrak{R}_z(M_n) - \mathfrak{R}_z(M_{n,t}) = -\mathfrak{R}_z(M_n)(M_n - M_{n,t})\mathfrak{R}_z(M_{n,t}),$$

the rank of $\mathfrak{R}_z(M_n) - \mathfrak{R}_z(M_{n,t})$ is not more than the one of $M_n - M_{n,t}$. One has (omitting the indices n in A_n, B_n)

$$M_n - M_{n,t} = G(A)G(A)^* + G(B)G(B)^* + G(A)B^* + F(A)G(B)^* + G(B)A^* + F(B)G(A)^*,$$

so its rank is not more that $3\operatorname{rg} G(A_n) + 3\operatorname{rg} G(B_n)$. Hence

$$\begin{aligned} & \mathrm{P}\{\sup_{\Im z \geq 1} |\mathrm{tr}\,[\mathfrak{R}_z(M_n) - \mathfrak{R}_z(M_{n,t})]| \ > \ \tfrac{\varepsilon}{3}\} \\ \leq \quad & \mathrm{P}\left\{\tfrac{6}{q_1(n)}(\operatorname{rg} G(A_n) + \operatorname{rg} G(B_n)) > \tfrac{\varepsilon}{3}\right\} \\ = \quad & \mathrm{P}\left\{\mu_{A_n}(\mathbb{R}\backslash[-t,t]) + \mu_{B_n}(\mathbb{R}\backslash[-t,t]) > \tfrac{\varepsilon}{18}\right\}, \end{aligned}$$

which tends to zero when n goes to infinity, by (17). On the other side, the second term of right hand side of (18) goes to zero when n goes to infinity by definition of $\boxplus_\lambda$ and by the previous lemma. So (15) is proved. □

3.10. **Examples.** In this section, we give examples of rectangular R-transforms of symmetric probability measures, and examples of computations of rectangular free convolutions.

The case where $\lambda = 1$ (see section 3.6), where the rectangular R-transform and the rectangular free convolution are essentially the Voiculescu transform and the free convolution defined by Voiculescu and Bercovici, doesn't present any new interest (one can find the little number of examples known in, e.g., [HP00]), so we will suppose that $\lambda \in [0, 1)$. Unfortunately, for $\lambda \in (0, 1)$, few computations can be done, still less than for the well known case $\lambda = 1$.

3.10.1. *Convolution of symmetric Bernouilli distributions.*

Proposition 3.15. *Suppose $\lambda > 0$. Then $\tau := \frac{\delta_{-1}+\delta_1}{2} \boxplus_\lambda \frac{\delta_{-1}+\delta_1}{2}$ has support*

$$[-(2+\kappa)^{1/2}, -(2-\kappa)^{1/2}] \cup [(2-\kappa)^{1/2}, (2+\kappa)^{1/2}], \tag{19}$$

with $\kappa = 2(\lambda(2-\lambda))^{1/2} \in (0, 2)$, and it admits a density with formula

$$\frac{\left[\kappa^2 - (x^2-2)^2\right]^{1/2}}{\pi\lambda|x|(4-x^2)} \tag{20}$$

on its support.

Remark 3.16. *By theorem 3.13, this proposition means concretely that for sequences $q_1(n) \leq q_2(n)$ which tend to infinity such that $q_1(n)/q_2(n)$ tends to $\lambda \in (0, 1)$, if one considers independent random matrices A_n, B_n, U_n, V_n such that for all n, A_n, B_n are $q_1(n) \times q_2(n)$ matrices such that $A_n A_n^*$, $B_n B_n^*$ are the identity matrix and U_n, V_n are uniformly distributed unitary random matrices with respective sizes $q_1(n), q_2(n)$, then the symmetrized singular law of $A_n + U_n B_n V_n$ tends to the probability measure τ with the density of (20) on the support of (19). Equivalently, one can say that the spectral law of*

$$(A_n + U_n B_n V_n)(A_n + U_n B_n V_n)^*$$

tends to the push forward of τ by the map $x \mapsto x^2$, i.e. to the probability measure with support $[2-\kappa, 2+\kappa]$ and with density

$$\frac{\left[\kappa^2 - (x-2)^2\right]^{1/2}}{\pi\lambda x(4-x)}.$$

Remark 3.17. *Note that the rectangular free convolution of the Dirac mass at zero with any other symmetric probability measure is the measure itself, and that for any $c > 0$, $\frac{\delta_{-c}+\delta_c}{2} \boxplus_\lambda \frac{\delta_{-c}+\delta_c}{2}$ can be deduced from this proposition using a dilation. Note also that if $\lambda = 0$, then $\tau = \frac{\delta_{-\sqrt{2}}+\delta_{\sqrt{2}}}{2}$.*

Proof. Define $\nu = \frac{\delta_{-1}+\delta_1}{2}$. We have $G_\nu(z) = \frac{z}{z^2-1}$, hence $G_\nu(1/\sqrt{z}) = \frac{\sqrt{z}}{1-z}$ and

$$H_\nu(z) = \frac{\lambda z}{(1-z)^2} + (1-\lambda)\frac{z(1-z)}{(1-z)^2} = \frac{(\lambda-1)z^2+z}{(1-z)^2}.$$

Hence for $x, y \in \mathbb{C}\backslash\mathbb{R}^+$ in a neighborhood of zero, the equation $H_\nu(y) = x$ is equivalent to $(\lambda-1)y^2 + y = x - 2xy + xy^2$, i.e. to $y^2(x+1-\lambda) - y(1+2x) + x = 0$. In this case, since $y = H_\nu^{-1}(x)$ must tend to zero as x tends to zero, one has

$$H_\nu^{-1}(x) = \frac{1+2x-(1+4\lambda x)^{\frac{1}{2}}}{2(x+1-\lambda)} \text{ and } \frac{x}{H_\nu^{-1}(x)} - 1 = \frac{(1+4\lambda x)^{\frac{1}{2}} - 1 + 2x}{2}.$$

Note that $(\lambda+1)^2 + 4\lambda\left(\frac{x}{H_\nu^{-1}(x)} - 1\right) = \left(\lambda + (1+4\lambda x)^{\frac{1}{2}}\right)^2$, hence

$$C_\nu(x) = \frac{-\lambda - 1 + \left[(\lambda+1)^2 + 4\lambda\left(\frac{x}{H_\mu^{-1}(x)} - 1\right)\right]^{\frac{1}{2}}}{2\lambda} = \frac{1}{2\lambda}\left[(1+4\lambda x)^{\frac{1}{2}} - 1\right].$$

Thus by (14) and remark 3.9, for $y \in \mathbb{C}\backslash\mathbb{R}^+$ in the neighborhood of zero, if one denotes $(1+4\lambda y)^{\frac{1}{2}}$ by w, one has

$$H_\tau^{-1}(y) = \frac{\lambda y}{1+4\lambda y + (\lambda-1)(1+4\lambda y)^{\frac{1}{2}}} = \frac{(w^2-1)/4}{w^2 + (\lambda-1)w}.$$

Hence for $x \in \mathbb{C}\backslash\mathbb{R}^+$ in the neighborhood of zero, if one denotes $H_\tau(x)$ by y and $(1+4\lambda y)^{\frac{1}{2}}$ by w, one has $H_\tau^{-1}(y) = x$, i.e. $w^2 - 1 = 4xw^2 + 4x(\lambda-1)w$, i.e., since w tends to 1 as x tends to zero,

$$w = \frac{1}{1-4x}[2x(\lambda-1) + [1-4x+4(\lambda-1)^2x^2]^{\frac{1}{2}}].$$

Then an easy computation leads to

$$(\lambda-1)^2 + \frac{4\lambda y}{x} = \left(\frac{\lambda - 1 + 2[1-4x+4(\lambda-1)^2x^2]^{\frac{1}{2}}}{1-4x}\right)^2.$$

Thus by remark 3.9, for $z \in \mathbb{C}\backslash\mathbb{R}^+$ in the neighborhood of zero,

$$\frac{1}{\sqrt{z}} G_\tau\left(\frac{1}{\sqrt{z}}\right) = V\left(\frac{H_\tau(z)}{z}\right) = \frac{1}{2\lambda}\left(\lambda - 1 + \frac{\lambda - 1 + 2[1-4z+4(\lambda-1)^2z^2]^{\frac{1}{2}}}{1-4z}\right),$$

hence $G_\tau\left(\frac{1}{\sqrt{z}}\right) = \frac{1}{2\lambda\frac{1}{\sqrt{z}}}\left(\lambda - 1 + \frac{\lambda-1+2[1-4/(1/\sqrt{z})^2+4(\lambda-1)^2/(1/\sqrt{z})^4]^{\frac{1}{2}}}{1-4\frac{1}{(1/\sqrt{z})^2}}\right)$, so for $z \in i\mathbb{R}^-$ large enough, $G_\tau(z) = \frac{(2z^2-4)(\lambda-1)+2\left[(z^2-2)^2-\kappa^2\right]^{1/2}}{2\lambda z(z^2-4)}$ and by analytic continuation, for any $z \in \mathbb{C}^-$, $G_\tau(z) = \frac{(4-2z^2)(\lambda-1)+2s(z)}{2\lambda z(4-z^2)}$, where $s(z)$ is analytic on $\mathbb{C}^-$ and satisfies $s(z)^2 = (z^2-2)^2 - \kappa^2$. In order to compute τ, we shall use the following well known result (lemma 2.17 of [B06]): for Lebesgue-almost all real number x, the limit, as $z = u + iv \in \mathbb{C}^-$ tends to x in such a way that $(u-x)/v$ stays bounded, of $\Im G_\tau(z)/\pi$ is the density of the absolutely continuous part of τ at x.

Let us define

$$O := (-(2+\kappa)^{1/2}, -(2-\kappa)^{1/2}) \cup ((2-\kappa)^{1/2}, (2+\kappa)^{1/2}).$$

By what precedes, for Lebesgue-almost all $x \in \mathbb{R}$, there is no sequence $w_n > 0$ tending to 0 such that $\lim_{n\to\infty} \Im G_\tau(x - iw_n) < 0$. Hence since for all $x \in O$, $(x^2-2)^2 - \kappa^2$ is negative, for Lebesgue-almost all $x \in O$, there is $\varepsilon > 0$ such that for all $w \in (0, \varepsilon)$, for $z = x - iw$,

$$G_\tau(z) = \begin{cases} \frac{(4-2z^2)(\lambda-1)+2\sqrt{(z^2-2)^2-\kappa^2}}{2\lambda z(4-z^2)} & \text{if } x > 0, \\ \frac{(4-2z^2)(\lambda-1)-2\sqrt{(z^2-2)^2-\kappa^2}}{2\lambda z(4-z^2)} & \text{if } x < 0. \end{cases}$$

It implies that for Lebesgue-almost all $x \in O$,

$$\lim_{y\to 0^-} \Im G_\tau(x+iy)/\pi = \frac{\Im(\sqrt{(x^2-2)^2-\kappa^2})}{\pi\lambda|x|(4-x^2)} = \frac{\left[\kappa^2 - (x^2-2)^2\right]^{1/2}}{\pi\lambda|x|(4-x^2)}.$$

Now, to complete the proof, it suffices to prove that τ is equal the absolutely continuous part of its restriction to O, i.e. that

$$\int_O \frac{\left[\kappa^2 - (x^2-2)^2\right]^{1/2}}{\pi\lambda|x|(4-x^2)} \mathrm{d}x = 1.$$

This can be proved using the changing of variable $y = \frac{x^2-2}{(\kappa^2-(x^2-2)^2)^{1/2}}$ and the well known formula of the Cauchy transform of the Cauchy distribution. □

3.10.2. *The rectangular analogues of the Gaussian distributions.* Here, $\lambda > 0$ (the case $\lambda = 0$ can be found in [B-G1]). It is proved in [B-G1] that for all $\sigma > 0$, the rectangular analogue of the symmetric Gaussian distribution with variance σ^2 is the law N_{σ^2} with density

$$\frac{\left[4\lambda - (\sigma^2x^2 - 1 - \lambda)^2\right]^{1/2}}{2\pi\lambda\sigma^2|x|} \chi(\sigma^2x^2),$$

where χ stands for the characteristic function of the interval $[(1-\lambda^{1/2})^2, (1+\lambda^{1/2})^2]$. Its support is $[-\sigma(1+\lambda^{1/2}), -\sigma(1-\lambda^{1/2})] \cup [\sigma(1-\lambda^{1/2}), \sigma(1+\lambda^{1/2})]$, its rectangular R-transform is $\sigma^2 z$. One has, for all $\sigma, \alpha > 0$, $N_{\sigma^2} \boxplus_\lambda N_{\alpha^2} = N_{\sigma^2+\alpha^2}$.

Remark 3.18 (Growth of the support in the related semigroup)**.** *Note that in the additive semigroup $(\{N_c\,;\, c > 0\}, \boxplus_\lambda)$ the size of the support of N_c in not linear in c but in $c^{\frac{1}{2}}$. This kind of phenomenon, which recalls us that we are in a gaussian context, but still surprising since the support (and not the variance) is concerned, had already been observed in free probability.*

3.10.3. *The rectangular analogues of the Cauchy distributions.* For all $t > 0$, the rectangular analogue of the symmetric Cauchy law with parameter t by the Bercovici-Pata bijection with ratio λ is

$$\nu_t := 1_{\mathbb{R}\setminus[-\frac{t(1-\lambda)}{2}, \frac{t(1-\lambda)}{2}]}(x) \frac{t}{\pi(\lambda t^2 + x^2)} \left[1 - \frac{t^2(\lambda-1)^2}{4x^2}\right]^{\frac{1}{2}} \mathrm{d}x.$$

Its support is $\mathbb{R}\setminus(-\frac{t(1-\lambda)}{2}, \frac{t(1-\lambda)}{2})$. It is proved in [B-G1] that the R-transform with ratio λ of ν_t admits an analytic extension to $\mathbb{C}\setminus\mathbb{R}^+$ given by the formula $it\sqrt{z}$. Hence one has, for all $t_1, t_2 > 0$, $\nu_{t_1} \boxplus_\lambda \nu_{t_2} = \nu_{t_1+t_2}$.

3.10.4. *The rectangular analogues of other symmetric stable laws.* It can also be proved that for all $\alpha \in (0,2)$, the rectangular analogue of the classical semigroup of stable laws with index α is the set of laws with rectangular R-transforms with ratio λ given by the formula $z \mapsto -C(-z)^{\alpha/2}$, with $C > 0$ and where the $\alpha/2$-th power is defined with respect to the canonical determination of the argument on $\mathbb{C}\backslash\mathbb{R}^-$. Unfortunately, though we give their rectangular R-transforms with ratio λ, except for $\alpha = 1$ (see the previous paragraph), we cannot give the densities of these laws. However, just like for free stable laws (see [BPB99]), results can be proved about their densities and supports (see [BBG07]).

3.10.5. *The rectangular analogues of the symmetric Poisson distributions.* It is proved, in [B-G1], that for all $c > 0$, the sequence

$$\left(\left(1 - \frac{c}{n}\right)\delta_0 + \frac{c}{2n}(\delta_{-1} + \delta_1)\right)^{\boxplus_\lambda n}$$

has a weak limit, that we shall denote by P_c, and that the rectangular R-transform of P_c is given by the formula $\frac{cz}{1-z}$. We cannot compute P_c itself, but we know that for all $c, c' > 0$, $P_c \boxplus_\lambda P_{c'} = P_{c+c'}$.

4. Appendix: a technical result

In this section, we do not make any particular hypothesis about ρ_k's: they are nonnegative, and at most one of them is zero. So we can suppose that $\rho_2, \ldots, \rho_d$ are positive.

For all integers n, k, for each partition $\mathcal{P}$ of $[k]$, $[n]^{\mathcal{P}}$ will design the set of elements $i = (i_1, \ldots, i_k)$ of $[n]^k$ such that $h \overset{\mathcal{P}}{\sim} h' \Leftrightarrow i_h = i_{h'}$. Moreover, $|\mathcal{P}|$ denotes the number of blocks of $\mathcal{P}$.

Definition 4.1. *An $n \times n$ matrix is said to be an* almost diagonal matrix *if it is of the type $\widetilde{M}$, with M diagonal (either square or rectangular) matrix. In this case, for all $j \in [n]$, A_j will denote the only nonzero term on the j-th column of A if it exists, and zero in the other case.*

Lemma 4.2. *Let K be a positive integer, and let for each $n \geq 1$, $A(1,n), \ldots, A(K,n)$ be a family of almost diagonal $n \times n$ matrices. Let $\mathcal{H}$ be a subset of $[K]$, and let $\mathcal{P}$ be a partition of $[K]$ such that for all $\{h\}$ singleton class of $\mathcal{P}$, $A(h,n)$ is a diagonal matrix and $\operatorname{Tr} A(h,n) = 0$. We suppose too that the family $(A(i,n))_{i,n}$ is uniformly bounded, and that for all $h \in \mathcal{H}$, for all $i \in \{q_1(n)+1, \ldots, n\}$, $A_i(h,n) = 0$. Then*

$$\sum_{i \in [n]^{\mathcal{P}}} \prod_{h=1}^{K} A_{i_h}(h,n) = O(q_1(n)^{|\mathcal{P}_\mathcal{H}| - p_\mathcal{H}/2} n^{|\overline{\mathcal{P}}_\mathcal{H}| - \overline{p}_\mathcal{H}/2}),$$

where $\mathcal{P}_\mathcal{H} = \{B \cap \mathcal{H}\,;\, B \in \mathcal{P}\}\backslash\{\emptyset\}$ and $\overline{\mathcal{P}}_\mathcal{H}\{B \cap (K\backslash\mathcal{H})\,;\, B \in \mathcal{P}\}\backslash\{\emptyset\})$ and $p_\mathcal{H}$ (resp. $\overline{p}_\mathcal{H}$) the number of singletons in $\mathcal{P}_\mathcal{H}$ (resp. of $\overline{\mathcal{P}}_\mathcal{H}$).

Proof. We will prove this result by induction on $p = p_\mathcal{H} + \overline{p}_\mathcal{H}$. If $p = 0$, the result is clear. Now, let us suppose the result to be true to the ranks $0, \ldots, p-1$, with $p \geq 1$. Consider $\mathcal{P}$ such that $p_\mathcal{H} + \overline{p}_\mathcal{H} = p$, and a singleton class $\{h_0\}$ of $\mathcal{P}$. Denote by $\mathcal{P}'$ the partition $\mathcal{P} - \{\{h_0\}\}$ of $\{1, \ldots, \hat{h}_0, \ldots, K\}$ and let, for each class B of $\mathcal{P}$ with $B \neq \{h_0\}$, $\mathcal{P}(h_0 \to B)$ be the partition of

$[K]$ obtained from $\mathcal{P}$ by linking the classes $\{h_0\}$ and B. We have

$$\sum_{i\in[n]^{\mathcal{P}}}\prod_{h=1}^{K}A_{i_h}(h,n)=\sum_{(i_1,\ldots,\hat{i}_{h_0},\ldots,i_K)\in[n]^{\mathcal{P}'}}\left[\left(\sum_{i\in[n]-\{i_1,\ldots,\hat{i}_{h_0},\ldots,i_K\}}A_i(h_0,n)\right)\prod_{\substack{h=1\\h\neq h_0}}^{K}A_{i_h}(h,n)\right].$$

But, since $A(h_0,n)$ is diagonal and has null trace,

$$\sum_{i\in[n]-\{i_1,\ldots,\hat{i}_{h_0},\ldots,i_K\}}A_i(h_0,n)=-\sum_{i\in\{i_1,\ldots,\hat{i}_{h_0},\ldots,i_K\}}A_i(h_0,n).$$

So

$$\sum_{i\in[n]^{\mathcal{P}}}\prod_{h=1}^{K}A_{i_h}(h,n)=-\sum_{\substack{B\in\mathcal{P}\\B\neq\{h_0\}}}\sum_{i\in[n]^{\mathcal{P}(h_0\to B)}}\prod_{h=1}^{K}A_{i_h}(h,n).$$

For example, suppose $h_0\in\mathcal{H}$ (the other case is treated in the same way). With the induction hypothesis, and dividing the sum in $B\in\mathcal{P}_{\mathcal{H}}$ and $B\in\overline{\mathcal{P}}_{\mathcal{H}}$, one has (each ϵ_B being 1 or 0 according to whether B is a singleton or not)

$$\begin{aligned}\sum_{i\in[n]^{\mathcal{P}}}\prod_{h=1}^{K}A_{i_h}(h,n) &= \sum_{\substack{B\in\mathcal{P}_{\mathcal{H}}\\B\neq\{h_0\}}}O\left(q_1(n)^{|\mathcal{P}_{\mathcal{H}}|-1-(p_{\mathcal{H}}-1-\epsilon_B)/2}n^{|\overline{\mathcal{P}}_{\mathcal{H}}|-\overline{p}_{\mathcal{H}}/2}\right)\\ &\quad+\sum_{B\in\overline{\mathcal{P}}_{\mathcal{H}}}O\left(q_1(n)^{|\mathcal{P}_{\mathcal{H}}|-(p_{\mathcal{H}}-1)/2}n^{|\overline{\mathcal{P}}_{\mathcal{H}}|-1-(\overline{p}_{\mathcal{H}}-\epsilon_B)/2}\right)\\ &= q_1(n)^{\frac{\epsilon_B-1}{2}}\sum_{\substack{B\in\mathcal{P}_{\mathcal{H}}\\B\neq\{h_0\}}}O\left(q_1(n)^{|\mathcal{P}_{\mathcal{H}}|-p_{\mathcal{H}}/2}n^{|\overline{\mathcal{P}}_{\mathcal{H}}|-\overline{p}_{\mathcal{H}}/2}\right)\\ &\quad+q_1(n)^{\frac{1}{2}}n^{\frac{\epsilon_B-2}{2}}\sum_{B\in\overline{\mathcal{P}}_{\mathcal{H}}}O\left(q_1(n)^{|\mathcal{P}_{\mathcal{H}}|-p_{\mathcal{H}}/2}n^{|\overline{\mathcal{P}}_{\mathcal{H}}|-\overline{p}_{\mathcal{H}}/2}\right).\end{aligned}$$

Since $q_1(n)=O(n)$, the result is proved. □

The following proposition, main result of this subsection, is divided in two cases, which give very similar results with very similar proofs. The proposition involves special classes of matrices.

Definition 4.3. *For $k\in[d]$, we define $\mathbb{U}_k(n)$ to be the set of $n\times n$ matrices whose (k,k)-th block is a $q_k(n)\times q_k(n)$ unitary matrix, and with other blocks zero.*

Note that $(\mathbb{U}_k(n),\times)$ is a compact group, isomorphic to the one of $q_k(n)\times q_k(n)$ unitary matrices, and random matrices of $\mathbb{U}_k(n)$ distributed according to the Haar measure will be said to be *uniform.* In the following, we are going to use an integer N and the cyclic order on $[N]$. This means that to put the index $N+1$ on an element is equivalent to put the index 1.

Proposition 4.4. *Let, for $n\geq 1$, $V(s,k,n)$ ($s\in\mathbb{N},k\in[d]$), be a family of independent random matrices, such that for all s,k, $V(s,k,n)$ is uniform on $\mathbb{U}_k(n)$. Fix N positive integer and $R>0$.*

Case 1°) Let, for each $n\in\mathbb{N}$, $D(1,n),\ldots,D(N,n)$ be $n\times n$ constant matrices such that for all $r\in[N]$, $||D(r,n)||\leq R$, and there exists $u(r),v(r)\in[d]$ such that one of the two following conditions is realized:

(i) *$u(r)=v(r)$ and for all n, $D(r,n)=p_{u(r)}(n)$,*

(ii) *for all* n, $D(r,n)$ *is an almost diagonal matrix of the type* $\widetilde{M}$, *where* M *is a* $q_{u(r)}(n)\times q_{v(r)}(n)$ *diagonal matrix, and* $\mathrm{E}(D(r,n)) \underset{n\to\infty}{\longrightarrow} 0$.

Consider $(s_l,k_l)_{l\in[N]} \in (\mathbb{N}\times[d])^N$, *and suppose that for all* $r\in[N]$ *such that (i) above is satisfied,* $(s_r,k_r)\neq(s_{r+1},k_{r+1})$. *Consider also* $m_1,\dots,m_N\in\mathbb{Z}-\{0\}$ *and* $\eta>0$. *Then the probability of the event*

$$\{||\,\mathrm{E}(V(s_1,k_1,n)^{m_1}D(1,n)\cdots D(N-1,n)V(s_N,k_N,n)^{m_N}D(N,n))||\leq\eta\}$$

tends to 1 *as* n *goes to infinity.*

Case 2°) Let, for each $n\in\mathbb{N}$, $D(0,n),\dots,D(N,n)$ *be* $n\times n$ *constant matrices with norms also* $\leq R$. *We suppose that there exists* $v(0),u(N)\in[d]$ *such that* $D(0,n)$ *(resp.* $D(N,n)$*) is an almost diagonal matrix of the type* $\widetilde{M}$, *where* M *is a* $q_1(n)\times q_{v(0)}(n)$ *(resp.* $q_{u(N)}(n)\times q_1(n)$*), and that for all* $r\in[N-1]$, *there exists* $u(r),v(r)\in[d]$ *such that one of the two following conditions is realized:*

(i) $u(r)=v(r)$ *and for all* n, $D(r,n)=p_{u(r)}(n)$,

(ii) *for all* n, $D(r,n)$ *is an almost diagonal matrix of the type* $\widetilde{M}$, *where* M *is a* $q_{u(r)}(n)\times q_{v(r)}(n)$ *diagonal matrix, and* $\mathrm{E}(D(r,n)) \underset{n\to\infty}{\longrightarrow} 0$.

Consider $(s_l,k_l)_{l\in[N]} \in (\mathbb{N}\times[d])^N$, *and suppose that for all* $r\in[N-1]$ *such that (i) above is satisfied,* $(s_r,k_r)\neq(s_{r+1},k_{r+1})$. *Consider also* $m_1,\dots,m_N\in\mathbb{Z}-\{0\}$ *and* $\eta>0$. *Then the probability of the event*

$$\{||\,\mathrm{E}(D(0,n)V(s_1,k_1,n)^{m_1}D(1,n)\cdots D(N-1,n)V(s_N,k_N,n)^{m_N}D(N,n))||\leq\eta\}$$

tends to 1 *as* n *goes to infinity.*

Remark 4.5. *By linearity of* E, *the result stays true in case* 1°*), (resp. in case* 2°*)) if for all* r *in* $[N]$ *(resp. in* $[N-1]$*) such that (ii) is satisfied,* $B(r,n)$ *is replaced by a sum of constant almost diagonal matrices whose images by* E *also tend to zero.*

The proof is inspired from the one of Theorem 4.3.1 p. 147 in [HP00].

Proof. Both proofs will be made together, we will only have to separate them sometimes.

Step I. Cases 1°) and 2°) Let us denote, for all r which satisfies (ii) above and $u(r)=v(r)$,

$$D(r,n)=D'(r,n)+\lambda_r(n)p_{u(r)}(n),\ \text{with}\ \mathrm{E}(D'(r,n))=0.$$

Then by hypothesis, $\lambda_h(n) \underset{n\to\infty}{\longrightarrow} 0$. Hence, by linearity and boundedness hypothesis, it suffices to prove the result for $D(r,n)$ replaced by $D'(r,n)$. Thus we can from now on suppose that for all r such that (ii) is realized, we have $\mathrm{E}(D(r,n))=0$ for all n.

Step II. Case 1°) By definition of E, it suffices to prove that for all $k\in[d]$, the normalized trace of the k-th diagonal block of the product

$$V(s_1,k_1,n)^{m_1}D(1,n)\cdots D(N-1,n)V(s_N,k_N,n)^{m_N}D(N,n)$$

converges in probability to zero, i.e. that

$$\frac{1}{q_k(n)}\,\mathrm{Tr}\,p_k(n)V(s_1,k_1,n)^{m_1}D(1,n)\cdots D(N-1,n)V(s_N,k_N,n)^{m_N}D(N,n)p_k(n) \tag{21}$$

converges in probability to zero. So let us fix $k\in[d]$, and let us prove it.

If $k_1 \neq k$ or $v(N) \neq k$, then the k-th block of the matrix is always zero, so the result is clear. So let us suppose that $k_1 = v(N) = k$. We can then remove $p_k(n)$ in (21). By Markov inequality, it suffices to prove that (21) tends to zero in L^2. Since for all matrix M, $|\operatorname{Tr} M|^2 = \operatorname{Tr} M \operatorname{Tr} M^*$, we only have to prove that the expectation of

(22)
$$\operatorname{Tr} V(s_1,k_1,n)^{m_1} D(1,n)\cdots D(N,n) \operatorname{Tr} D(N,n)^* V(s_N,k_N,n)^{-m_N} D(N-1,n)^* \cdots V(s_1,k_1,n)^{-m_1}$$

is $o(q_k(n)^2)$.

Case 2°) We have to prove that

$$\frac{1}{q_1(n)} \operatorname{Tr} D(0,n) V(s_1,k_1,n)^{m_1} D(1,n)\cdots D(N-1,n) V(s_N,k_N,n)^{m_N} D(N,n)$$

converges in probability to zero. Since $\operatorname{Tr} XY = \operatorname{Tr} YX$, it suffices to prove that

$$\frac{1}{q_1(n)} \operatorname{Tr} V(s_1,k_1,n)^{m_1} D(1,n)\cdots D(N-1,n) V(s_N,k_N,n)^{m_N} D(N,n) D(0,n) \tag{23}$$

converges in probability to zero. From now on, we will denote $D(N,n)D(0,n)$ by $D(N,n)$, which is now an $n\times n$ matrix of the type $\widetilde{M}$ with M $q_{u(N)}(n)\times q_{v(0)}(n)$, but which has not more than $q_1(n)$ nonzero entries. For the same reason as above, we only have to prove that the expectation of

(24)
$$\operatorname{Tr} V(s_1,k_1,n)^{m_1} D(1,n)\cdots D(N,n) \operatorname{Tr} D(N,n)^* V(s_N,k_N,n)^{-m_N} D(N-1,n)^* \cdots V(s_1,k_1,n)^{-m_1}$$

is $o(q_1(n)^2)$.

Step III. Cases 1°) and 2°) Let us expand both traces in the previous product, and re index the sum on partitions.

Let $M = |m_1| + \cdots + |m_N|$, let $M(0) = 0$ and, for $r \in \{1,\ldots,N\}$,

$$M(r) = |m_1| + \cdots + |m_r|,$$

let $M(N+1) = 2M$, and let, for $r \in \{N+2,\ldots,2N\}$,

$$M(r) = M + |m_N| + |m_{N-1}| + \cdots + |m_{2N+2-r}|.$$

Let also, for $r \in \{N+1,\ldots,2N\}$, $D(r,n) = D(2N+1-r,n)^*$. We write, for $\kappa \in \mathbb{N}\times[d]$,

$$\begin{aligned} v_{i,j}(\kappa,1,n) &= V(\kappa,n)_{i,j}, \\ v_{i,j}(\kappa,-1,n) &= \bar{V}(\kappa,n)_{j,i}. \end{aligned}$$

With those notations, there exists two functions κ and ε such that the expectation of (22) is

$$\sum_{i\in[n]^{2M}} \prod_{r=1}^{2N} D(r,n)_{j_{M(r)}} \mathbb{E}\left[\prod_{h=1}^{2M} v_{i_h,j_h}(\kappa(h),\varepsilon(h),n)\right],$$

where $\mathbb{E}$ denotes expectation and for all $i = (i_1,\ldots,i_{2M}) \in [n]^{2M}$,

$$j_1 = i_2,\ j_2 = i_3, \ldots, j_M = i_1,\ \ j_{M+1} = i_{M+2},\ j_{M+2} = i_{M+3},\ \ldots, j_{2M-1} = i_{2M},\ j_{2M} = i_{M+1}.$$

Let us introduce the partition $\mathcal{Q}$ of $[2M]$ defined by $h \overset{\mathcal{Q}}{\sim} h'$ if and only if $i_h = i_{h'}$. Then we can rewrite the preceding sum

$$\sum_{\substack{\mathcal{Q} \text{ partition} \\ \text{of } [2M]}} \sum_{i \in [n]^{\mathcal{Q}}} \prod_{r=1}^{2N} D(r,n)_{j_{M(r)}} \mathbb{E}\left[\prod_{h=1}^{2M} v_{i_h,j_h}(\kappa(h),\varepsilon(h),n)\right].$$

Thus it suffices to prove that for all partition $\mathcal{Q}$ of $[2M]$, the sum

$$\sum_{i \in [n]^{\mathcal{Q}}} \prod_{r=1}^{2N} D(r,n)_{j_{M(r)}} \mathbb{E}\left[\prod_{h=1}^{2M} v_{i_h,j_h}(\kappa(h),\varepsilon(h),n)\right] \tag{25}$$

is $o(q_k(n)^2)$ in case 1°), and $o(q_1(n)^2)$ in case 2°), as n goes to infinity.

So we fix a partition $\mathcal{Q}$ of $[2M]$.

Step IV. Cases 1°) and 2°) Let us denote, for all $h \in [2M]$, $\kappa(h) = (s(h), k(h))$. Note that, for $i \in [n]^{\mathcal{Q}}$, for

$$\mathbb{E}\left[\prod_{h=1}^{2M} v_{i_h,j_h}(\kappa(h),\varepsilon(h),n)\right] \tag{26}$$

to be nonzero, i has to satisfy

$$\forall h \in [2M], i_h, j_h \in \{q_1(n) + \cdots + q_{k(h)-1}(n) + 1, \ldots, q_1(n) + \cdots + q_{k(h)-1}(n) + q_{k(h)}(n)\}. \tag{27}$$

Indeed, in the other case, one of the factors in the product is zero, by definition of $M \mapsto \widetilde{M}$. Note also that for all elements i of $[n]^{\mathcal{Q}}$, which satisfy (27), the expectation (26) is the same, by invariance of Haar measure on unitary groups under permutation of rows and columns. Let us denote this common expectation by $\mathbb{E}_{\mathcal{Q}}$. So, one can write the sum of (25)

$$\mathbb{E}_{\mathcal{Q}} \sum_{i} \prod_{r=1}^{2N} D(r,n)_{j_{M(r)}} \tag{28}$$

where the sum is taken on elements i of $[n]^{\mathcal{Q}}$ which satisfy (27).

It is well known (see [HP00], equation (4.2.11)) that if, for each m, $U(m)$ is a uniform unitary $m \times m$ random matrix, and $1 \leq i_m, j_m \leq m$, then the sequence $\mathbb{E}(|U(m)_{i_m,j_m}|^{2M})$ does not depend on the choices of i_m, j_m and

$$\mathbb{E}(|U(m)_{i_m,j_m}|^{2M})^{1/2M} = O(m^{-1/2}). \tag{29}$$

So, by Hölder inequality, for $i \in [n]^{\mathcal{Q}}$,

$$|\mathbb{E}_{\mathcal{Q}}| \leq \prod_{h=1}^{2M} \mathbb{E}(|v_{i_h,j_h}(\kappa(h),\varepsilon(h),n)|^{2M})^{1/2M} = O(q_1(n)^{-M'} n^{-M''}),$$

where $M' = \sum_{\substack{1 \leq r \leq N \\ k_r = 1}} |m_r|$ and $M'' = \sum_{\substack{1 \leq r \leq N \\ k_r \neq 1}} |m_r|$. Thus it suffices to prove that

$$\sum_{i} \prod_{r=1}^{2N} D(r,n)_{j_{M(r)}}, \tag{30}$$

is $o(q_k(n)^2 q_1(n)^{M'} n^{M''})$ in case 1°), and $o(q_1(n)^{M'+2} n^{M''})$ in case 2°), where the sum is taken on the elements i of $[n]^{\mathcal{Q}}$ which satisfy (27).

Step V. Cases 1°) and 2°) We are going to prove it as an application of lemma 4.2. The sum of (30) is the sum, on the same i's, of

$$\begin{aligned} & D(1,n)_{i_{M(1)+1}} D(2,n)_{i_{M(2)+1}} \cdots D(N-1,n)_{i_{M(N-1)+1}} D(N,n)_{i_1} \\ \times \quad & D(N+1,n)_{i_{M+1}} D(N+2,n)_{i_{M(N+2)}} D(N+3,n)_{i_{M(N+3)}} \cdots D(2N,n)_{i_{M(2N)}}. \end{aligned}$$

i.e. of

$$\begin{aligned} & D(N,n)_{i_1} D(1,n)_{i_{M(1)+1}} D(2,n)_{i_{M(2)+1}} \cdots D(N-1,n)_{i_{M(N-1)+1}} \\ \times \quad & D(N+1,n)_{i_{M+1}} D(N+2,n)_{i_{M(N+2)}} D(N+3,n)_{i_{M(N+3)}} \cdots D(2N,n)_{i_{M(2N)}}. \end{aligned}$$

Adding some matrices of the type $p_k(n)$ in the product, one sees that this sum is equal to

$$\sum_{i \in [n]^{\mathcal{Q}}} \prod_{h=1}^{2M} A_{i_h}(h,n), \tag{31}$$

where for $h \in [2M]$,

$$A(h,n) = \begin{cases} p_{k(h-1)}(n) D(r,n) p_{k(h)}(n) & \text{if } \exists r \in \{0,\dots,N-1\}, \\ & \quad h = M(r)+1 \\ p_{k(h)}(n) & \text{if } h \in [M] \text{ and } h \notin \\ & \quad \{M(0)+1,\dots,M(N-1)+1\} \\ p_{k(2M)}(n) D(N+1,n) p_{k(M+1)}(n) & \text{if } h = M+1 \\ p_{k(h)}(n) D(r,n) p_{k(h+1)}(n) & \text{if } \exists r \in \{N+2,\dots,2N\}, \\ & \quad h = M(r) \\ p_{k(h)}(n) & \text{if } h \in [2N] \text{ and } h \notin \\ & \quad [M+1] \cup \{M(N+2),\dots,M(2N)\} \end{cases}$$

Now, note that by lemma 4.2.2 of [HP00] (or because the uniform distribution on the unitary group is left and right invariant), if $i \in [n]^{\mathcal{Q}}$ corresponds to a nonzero term in equation (25), then for all $\kappa \in \mathbb{N}\times[d]$, for all $1 \leq \alpha, \beta \leq n$,

$$\begin{aligned} |\{h\,;\, i_h = \alpha, \kappa(h) = \kappa, \varepsilon(h) = 1\}| &= |\{h\,;\, j_h = \alpha, \kappa(h) = \kappa, \varepsilon(h) = -1\}|, \\ |\{h\,;\, j_h = \beta, \kappa(h) = \kappa, \varepsilon(h) = 1\}| &= |\{h\,;\, i_h = \beta, \kappa(h) = \kappa, \varepsilon(h) = -1\}|. \end{aligned}$$

Thus, if $\{h\}$ is a singleton class in $\mathcal{Q}$, and if $\mathbb{E}_{\mathcal{Q}}$ is non null, one has, using the cyclic orders on $[M]$ and on $\{M+1, M+2, \dots, 2M\}$,

$$\begin{aligned} \varepsilon(h) = 1 \quad &\Rightarrow \quad \kappa(h) = \kappa(h-1), \varepsilon(h-1) = -1, \\ \text{and} \qquad \varepsilon(h) = -1 \quad &\Rightarrow \quad \kappa(h) = \kappa(h-1), \varepsilon(h-1) = 1. \end{aligned}$$

It clearly follows that there exists $r \in \{0,\dots,2N\} - \{N+1\}$ such that $h = M(r)+1$. Moreover, by hypothesis, in case 1°), $A(h,n) = D(r,n)$ is a diagonal matrix with null trace, and, in case 2°), $A(h,n) = D(r,n)$ is a diagonal matrix with null trace whenever $h \notin \{M, M+1\}$.

Case 1°) Let us apply the lemma with $\mathcal{P} = \mathcal{Q}$ and $\mathcal{H} = \{h \in [2M]\,;\, k(h) = 1\}$. By definition of the $A(h,n)$'s, for all $h \in \mathcal{H}$, for all $i \in \{q_1(n)+1, \dots, n\}$, $A_i(h,n) = 0$. By the lemma, the sum of (30), which is equal to the sum of (31), is $O(q_1(n)^{|\mathcal{P}_{\mathcal{H}}| - p_{\mathcal{H}}/2} n^{|\overline{\mathcal{P}}_{\mathcal{H}}| - \overline{p}_{\mathcal{H}}/2})$. But for any partition $\mathcal{X}$ with x singletons of a set S, one has $|\mathcal{X}| \leq x + (|S| - x)/2$, so $|\mathcal{X}| - x/2 \leq |S|/2$. So,

$$|\mathcal{P}_{\mathcal{H}}| + |\overline{\mathcal{P}}_{\mathcal{H}}| - \frac{p_{\mathcal{H}}}{2} - \frac{\overline{p}_{\mathcal{H}}}{2} \leq (2M)/2 = M' + M'',$$

Moreover, since for all $h \in [2M]$, $k(h) = 1$ implies $h \in \mathcal{H}$, $|\overline{\mathcal{P}}_{\mathcal{H}}| - \overline{p}_{\mathcal{H}}/2 \leq \frac{1}{2}|\{h\,;\, k(h) \neq 1\}| = M''$. Thus, since $q_1(n) \leq n$, $n^{|\overline{\mathcal{P}}_{\mathcal{H}}| - \overline{p}_{\mathcal{H}}/2} = n^{|\overline{\mathcal{P}}_{\mathcal{H}}| - \overline{p}_{\mathcal{H}}/2 - M''} n^{M''} = O(q_1(n)^{|\overline{\mathcal{P}}_{\mathcal{H}}| - \overline{p}_{\mathcal{H}}/2 - M''} n^{M''})$. Hence

$$q_1(n)^{|\mathcal{P}_{\mathcal{H}}| - p_{\mathcal{H}}/2} n^{|\overline{\mathcal{P}}_{\mathcal{H}}| - \overline{p}_{\mathcal{H}}/2} = O(q_1(n)^{|\mathcal{P}_{\mathcal{H}}| - \frac{p_{\mathcal{H}}}{2} + |\overline{\mathcal{P}}_{\mathcal{H}}| - \frac{\overline{p}_{\mathcal{H}}}{2} - M''} n^{M''}) = O(q_1(n)^{M'} n^{M''}).$$

The sum of (30) is $O(q_1(n)^{M'} n^{M''})$, and the proposition is proved.

Case 2°) Let us prove that the sum of (30) is $o(q_1(n)^{2+M'} n^{M''})$ by induction on the number $|\mathcal{Q}|$ of classes of $\mathcal{Q}$. Let us denote, for B, B' class of $\mathcal{Q}$, $\mathcal{Q}(M \to B)$ the partition of $[2M]$ obtained from $\mathcal{Q}$ by linking classes B and $\{M\}$, and $\mathcal{Q}(M \to B, M+1 \to B')$ the partition obtained from $\mathcal{Q}$ by linking classes B and $\{M\}$, and then the classes B' (which has become $B \cup \{M\}$ if $B' = B$) and $\{M+1\}$ Let $J(n)$ be the set of elements j of $[n]$ such that $D(N,n)_j \neq 0$. Note that, as noted to step II, $|J(n)| \leq q_1(n)$, and, by definition of $D(N+1,n)$, $J(n)$ is also the set of elements j of $[n]$ such that $D(N+1,n)_j \neq 0$.

• If $\mathcal{Q}$ has only one class, by boundedness hypothesis and since $|J(n)| \leq q_1(n)$, the sum of (30) is $O(q_1(n))$.

• Assume that the conclusion holds to ranks $1, \ldots, |\mathcal{Q}| - 1$.

- If neither $\{M\}$ nor $\{M+1\}$ are singletons of $\mathcal{Q}$, we apply the lemma, as for case 1°), with $\mathcal{P} = \mathcal{Q}$ and $\mathcal{H} = \{h \in [2M]\,;\, k(h) = 1\}$. It leads to the conclusion in the same way as in case 1°).

- If exactly one of $\{M\}, \{M+1\}$, say $\{M\}$, is a singleton of $\mathcal{Q}$, define $\mathcal{P} = \mathcal{Q} - \{\{M\}\}$, partition of the set $[2M] - \{M\}$. Note that the sum of (30), which is equal to the sum of (31), is

$$\sum_{j \in J(n)} \Big[A(M,n)_j \sum_{i \in ([n] - \{j\})^{\mathcal{P}}} \prod_{\substack{h=1 \\ h \neq M}}^{2N} A_{i_h}(h,n) \Big]$$

$$= \sum_{j \in J(n)} \Big[A(M,n)_j \Big(\sum_{i \in [n]^{\mathcal{P}}} \prod_{\substack{h=1 \\ h \neq M}}^{2N} A_{i_h}(h,n) - \sum_{\substack{i \in [n]^{\mathcal{P}} \\ \exists h, i_h = j}} \prod_{\substack{h=1 \\ h \neq M}}^{2N} A_{i_h}(h,n) \Big) \Big]$$

$$= \Big(\sum_{j \in J(n)} A(M,n)_j \Big) \Big(\sum_{i \in [n]^{\mathcal{P}}} \prod_{\substack{h=1 \\ h \neq M}}^{2N} A_{i_h}(h,n) \Big) - \Big(\sum_{\substack{B \in \mathcal{Q} \\ B \neq \{M\}}} \sum_{i \in [n]^{\mathcal{Q}(M \to B)}} \prod_{h=1}^{2N} A_{i_h}(h,n) \Big)$$

So by the induction hypothesis, for all $B \in \mathcal{P}, B \neq \{M\}$, we have

$$\sum_{i \in [n]^{\mathcal{Q}(M \to B)}} \prod_{h=1}^{2N} A_{i_h}(h,n) = o\left(q_1(n)^{2+M'} n^{M''} \right).$$

So it suffices to prove that we have

$$\Big(\sum_{j \in J(n)} A(M,n)_j \Big) \Big(\sum_{i \in [n]^{\mathcal{P}}} \prod_{\substack{h=1 \\ h \neq M}}^{2N} A_{i_h}(h,n) \Big) = o\left(q_1(n)^{2+M'} n^{M''} \right). \tag{32}$$

But by boundedness hypothesis, $\sum_{j \in J(n)} A(M,n)_j = 0(|J(n)|) = 0(q_1(n))$. So it suffices to prove that the second sum in (32) is $o(q_1(n)^{M'+1} n^{M''})$. Define $\mathcal{H} = \{h \in [2M] - \{M\}\,;\, k(h) = 1\}$. By

definition of the $A(h,n)$'s, for all $h \in \mathcal{H}$, for all $i \in \{q_1(n)+1,\dots,n\}$, $A_i(h,n)=0$. So by the lemma, the second sum in (32) is $O(q_1(n)^{|\mathcal{P}_\mathcal{H}|-p_\mathcal{H}/2}n^{|\overline{\mathcal{P}}_\mathcal{H}|-\overline{p}_\mathcal{H}/2})$. Since, as noted above, for any partition $\mathcal{X}$ with x singletons of a set S, one has $|\mathcal{X}|-x/2 \leq |S|/2$, we have

$$|\mathcal{P}_\mathcal{H}|+|\overline{\mathcal{P}}_\mathcal{H}|-\frac{p_\mathcal{H}}{2}-\frac{\overline{p}_\mathcal{H}}{2} \leq (2M-1)/2 = M'+M''-1/2,$$

Moreover, since for all $h \in [2M]-\{M\}$, $k(h)=1$ implies $h \in \mathcal{H}$,

$$|\overline{\mathcal{P}}_\mathcal{H}|-\overline{p}_\mathcal{H}/2 \leq \frac{1}{2}\,|\{h \in [2M]-\{M\}\,;\, k(h)\neq 1\}| \leq M''.$$

Thus, since $q_1(n) \leq n$, $n^{|\overline{\mathcal{P}}_\mathcal{H}|-\overline{p}_\mathcal{H}/2} = n^{|\overline{\mathcal{P}}_\mathcal{H}|-\overline{p}_\mathcal{H}/2-M''}n^{M''} = O(q_1(n)^{|\overline{\mathcal{P}}_\mathcal{H}|-\overline{p}_\mathcal{H}/2-M''}n^{M''})$. Hence $q_1(n)^{|\mathcal{P}_\mathcal{H}|-p_\mathcal{H}/2}n^{|\overline{\mathcal{P}}_\mathcal{H}|-\overline{p}_\mathcal{H}/2} = O(q_1(n)^{|\mathcal{P}_\mathcal{H}|-\frac{p_\mathcal{H}}{2}+|\overline{\mathcal{P}}_\mathcal{H}|-\frac{\overline{p}_\mathcal{H}}{2}-M''}n^{M''}) = O(q_1(n)^{M'-1/2}n^{M''})$, thus the conclusion holds.

- If $\{M\}$ and $\{M+1\}$ are singletons of $\mathcal{Q}$, define $\mathcal{P}=\mathcal{Q}-\{\{M\},\{M+1\}\}$, partition of the set $[2M]-\{M,M+1\}$. Note that the sum of (30), which is equal to the sum of (31), is

$$\sum_{\substack{j,j'\in J(n)\\ j\neq j'}}\Big[A(M,n)_jA(M+1,n)_{j'}\sum_{i\in([n]-\{j,j'\})^\mathcal{P}}\prod_{\substack{h=1\\ h\neq M\\ h\neq M+1}}^{2N}A_{i_h}(h,n)\Big].$$

Now, let us define $\mathcal{Q}'$ to be the partition of $[2M]$ obtained from $\mathcal{Q}$ by linking classes $\{M\}$ and $\{M+1\}$. The sum of (30) can be written

$$\Big(\sum_{j,j'\in J(n)}A(M,n)_jA(M+1,n)_{j'}\Big)\Big(\sum_{i\in[n]^\mathcal{P}}\prod_{\substack{h=1\\ h\neq M\\ h\neq M+1}}^{2N}A_{i_h}(h,n)\Big)-\sum_{i\in[n]^{\mathcal{Q}'}}\prod_{h=1}^{2N}A_{i_h}(h,n)-\sum_{\mathcal{R}}\sum_{i\in[n]^\mathcal{R}}\prod_{h=1}^{2N}A_{i_h}(h,n),$$

where in the last sum, $\mathcal{R}$ runs over the set $\{\mathcal{Q}(M\to B, M+1\to B')\,;\, B,B'\in\mathcal{Q}\}$. So by the induction hypothesis, it suffices to prove that

$$\Big(\sum_{j,j'\in J(n)}A(M,n)_jA(M+1,n)_{j'}\Big)\Big(\sum_{i\in[n]^\mathcal{P}}\prod_{\substack{h=1\\ h\neq M\\ h\neq M+1}}^{2N}A_{i_h}(h,n)\Big) = o\left(q_1(n)^{2+M'}n^{M''}\right). \tag{33}$$

But by boundedness hypothesis, $\sum_{j,j'\in J(n)}A(M,n)_jA(M+1,n)_{j'} = 0(|J(n)|^2) = 0(q_1(n)^2)$. So it suffices to prove that the second sum in (33) is $o(q_1(n)^{M'}n^{M''})$. Define $\mathcal{H} := \{h\in[2M]-\{M,M+1\}\,;\, k(h)=1\}$. By definition of the $A(h,n)$'s, for all $h\in\mathcal{H}$, for all $i\in\{q_1(n)+1,\dots,n\}$, $A_i(h,n)=0$. So by the lemma, the second sum in (33) is

$$O(q_1(n)^{|\mathcal{P}_\mathcal{H}|-p_\mathcal{H}/2}n^{|\overline{\mathcal{P}}_\mathcal{H}|-\overline{p}_\mathcal{H}/2}).$$

Since, as noted above, for any partition $\mathcal{X}$ with x singletons of a set S, one has $|\mathcal{X}|-x/2\leq|S|/2$, we have

$$|\mathcal{P}_\mathcal{H}|+|\overline{\mathcal{P}}_\mathcal{H}|-\frac{p_\mathcal{H}}{2}-\frac{\overline{p}_\mathcal{H}}{2} \leq (2M-2)/2 = M'+M''-1,$$

Moreover, since for all $h\in[2M]-\{M,M+1\}$, $k(h)=1$ implies $h\in\mathcal{H}$,

$$|\overline{\mathcal{P}}_\mathcal{H}|-\overline{p}_\mathcal{H}/2 \leq \frac{1}{2}\,|\{h\in[2M]-\{M,M+1\}\,;\,k(h)\neq 1\}| \leq M''.$$

Thus, since $q_1(n) \leq n$,

$$n^{|\overline{\mathcal{P}}_{\mathcal{H}}|-\overline{p}_{\mathcal{H}}/2} = n^{|\overline{\mathcal{P}}_{\mathcal{H}}|-\overline{p}_{\mathcal{H}}/2-M''} n^{M''} = O(q_1(n)^{|\overline{\mathcal{P}}_{\mathcal{H}}|-\overline{p}_{\mathcal{H}}/2-M''} n^{M''}).$$

Hence

$$q_1(n)^{|\mathcal{P}_{\mathcal{H}}|-p_{\mathcal{H}}/2} n^{|\overline{\mathcal{P}}_{\mathcal{H}}|-\overline{p}_{\mathcal{H}}/2} = O(q_1(n)^{|\mathcal{P}_{\mathcal{H}}|-\frac{p_{\mathcal{H}}}{2}+|\overline{\mathcal{P}}_{\mathcal{H}}|-\frac{\overline{p}_{\mathcal{H}}}{2}-M''} n^{M''}) = O(q_1(n)^{M'-1} n^{M''}),$$

thus the conclusion holds. □

References

[A61] Akhiezer, N.I. *The classical moment problem*, Moscou, 1961

[B06] S. T. Belinschi, *The Lebesgue decomposition of the free additive convolution of two probability distributions*, Preprint (2006), ArXiv math.OA/0603104.

[BBG07] S. T. Belinschi, F. Benaych-Georges, A. Guionnet, *Regularization by free additive convolution, square and rectangular cases*, Preprint (2007), ArXiv.

[B-G1] Benaych-Georges, F. *Infinitely divisible distributions for rectangular free convolution: classification and matricial interpretation.* Probability Theory and Related Fields Volume 139, Numbers 1-2 / september 2007, 143-189.

[B-G2] Benaych-Georges, F. *Rectangular random matrices, related free entropy and free Fisher's information.* Available on the web page of the author, to appear in Journal of Operator Theory.

[BPB99] Bercovici, H., Pata, V., with an appendix by Biane, P. *Stable laws and domains of attraction in free probability theory* Annals of Mathematics, **149** (1999) 1023-1060

[BV93] Bercovici, H., Voiculescu, D. *Free convolution of measures with unbounded supports* Indiana Univ. Math. J. **42** (1993) 733-773

[B68] Billingsley, P. *Convergence of probability measures* Wiley, 1968

[CC04] Capitaine, M., Casalis, M. *Asymptotic freeness by generalized moments for Gaussian and Wishart matrices. Application to beta random matrices* Indiana Univ. Math. J. 53 (2004), no. 2, 397–431

[CDM05] Capitaine, M., Donati-Martin, C. *Strong asymptotic freeness for Wigner and Wishart matrices* preprint, 2005

[D74] Donoghue, W. *Monotone matrix functions and analytic continuation*, Springer, New-York, 1974

[DS04] Dozier, B. Silverstein, J. *Analysis of the limiting distribution of large dimensional information-plus-noise-type matrices* preprint, 2004

[GZ00] Guionnet, A., Zeitouni, O. *Concentration of the spectral measure for large matrices* Electron. Comm. Probab. 5 (2000), 119–136

[HL00] Haagerup, U., Larsen, F. *Brown's spectral distribution measure for R-diagonal elements in finite von Neumann algebras* Journ. Functional Analysis 176, 331-367 (2000).

[HLN04] Hachem, W., Loubaton, P., Najim, J. *The empirical distribution of the eigenvalues of a Gram matrix with a given variance profile* preprint 2004

[HLN05] Hachem, W., Loubaton, P., Najim, J. *Deterministic Equivalents for certain functionals of large random matrices* preprint, july 2005.

[HP00] Hiai, F., Petz, D. *The semicircle law, free random variables, and entropy* Amer. Math. Soc., Mathematical Surveys and Monographs Volume **77**, 2000

[HJ91] Horn, R.A., Johnson, C.R. *Topics in matrix analysis* Cambridge University Press, Cambridge, 1991

[M67] Mehta, M.L. *Random matrices and the statistical theory of energy levels* New York Academic Press, 1967

[N74] Nelson, E. *Notes on non-commutative integration* J. Functional Analysis **15** (1974), 103–116

[NSS99] Nica, Alexandru; Shlyakhtenko, Dimitri; Speicher, Roland *Some minimization problems for the free analogue of the Fisher information* Adv. Math. 141 (1999), no. 2, 282–321.

[NSS02] Nica, Alexandru; Shlyakhtenko, Dimitri; Speicher, Roland *Operator-valued distributions. I. Characterizations of freeness* Int. Math. Res. Not. 2002, no. 29, 1509–1538.

[PL02] Pastur, L., Lejay, A. *Matrices aléatoires : statistique asymptotique des valeurs propres* Séminaire de Probabilités XXXVI, Lecutre notes in M. **1801**, Springer 2002

[R64] Rota, Gian-Carlo *On the foundations of combinatorial theory. I. Theory of Möbius functions* Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 2 340–368 (1964)

[Sh96] Shlyakhtenko, D. *Random Gaussian band matrices and freeness with amalgamation* Internat. Math. Res. Notices 1996, no. 20, 1013–1025.

[S94] Speicher, R. *Multiplicative functions on the lattice of non-crossing partitions and free convolution*, Math. Annalen **298** (1994) 611-628

[S98] Speicher, R. *Combinatorial theory of the free product with amalgamation and operator-valued free probability theory*, Mem. Amer. Math. Soc. 132 (1998), no. 627

[S99] Speicher, R. (1999). Notes of my lectures on Combinatorics of Free Probability (IHP, Paris, 1999). Available on http://www.mast.queensu.ca/∼speicher

[ŚS01] Śniady, Piotr, Speicher, Roland, *Continuous family of invariant subspaces for R-diagonal operators* Invent. Math. 146, 329-363 (2001)

[V91] Voiculescu, Dan *Limit laws for random matrices and free products* Invent. Math. 104 (1991), no. 1, 201–220

[V95] Voiculescu, Dan *Operations on certain non-commutative operator-valued random variables* Recent advances in operator algebras (Orlans, 1992). Astrisque No. 232 (1995), 243–275.

[V98] Voiculescu, Dan *A strengthened asymptotic freeness result for random matrices with applications to free entropy* Internat. Math. Res. Notices 1998, no. 1, 41–63.

[VDN91] Voiculescu, D.V., Dykema, K., Nica, A. *Free random variables* CRM Monograghs Series No.1, Amer. Math. Soc., Providence, RI, 1992

[W58] Wigner, E. *On the distribution of the roots of certain symmetric matrices* Ann. Math. 67 325-327 (1958)

LPMA, UPMC Univ Paris 6, Case courier 188, 4, Place Jussieu, 75252 Paris Cedex 05, France.

E-mail address: florent.benaych@gmail.com