

# Outlier eigenvalues for non-Hermitian polynomials in independent i.i.d. matrices and deterministic matrices\*

Serban Belinschi<sup>†</sup>    Charles Bordenave<sup>‡</sup>    Mireille Capitaine<sup>†</sup>  
Guillaume Cébron<sup>†</sup>

## Abstract

We consider a square random matrix of size  $N$  of the form  $P(Y, A)$  where  $P$  is a noncommutative polynomial,  $A$  is a tuple of deterministic matrices converging in  $*$ -distribution, when  $N$  goes to infinity, towards a tuple  $a$  in some  $C^*$ -probability space and  $Y$  is a tuple of independent matrices with i.i.d. centered entries with variance  $1/N$ . We investigate the eigenvalues of  $P(Y, A)$  outside the spectrum of  $P(c, a)$  where  $c$  is a circular system which is free from  $a$ . We provide a sufficient condition to guarantee that these eigenvalues coincide asymptotically with those of  $P(0, A)$ .

**Keywords:** random matrices; free probability.

**MSC2020 subject classifications:** Primary 15B52; 60B20; 46L54, Secondary 15A18; 60F05.

Submitted to EJP on June 25, 2020, final version accepted on June 13, 2021.

## 1 Introduction

### 1.1 Previous results

Ginibre (1965) introduced the basic non-Hermitian ensemble of random matrix theory. A so-called Ginibre matrix is a  $N \times N$  matrix comprised of independent complex Gaussian entries. More generally, an *i.i.d. random matrix* is a  $N \times N$  random matrix  $X_N = (X_{ij})_{1 \leq i, j \leq N}$  whose entries are independent identically distributed complex-valued random variables with mean 0 and variance 1.

For any  $N \times N$  matrix  $B$ , denote by  $\lambda_1(B), \dots, \lambda_N(B)$  the eigenvalues of  $B$  and by  $\mu_B$  the empirical spectral measure of  $B$ :

$$\mu_B := \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(B)}.$$

---

\*G. C. was partly supported by the Project MESA (ANR-18-CE40-006) of the French National Research Agency (ANR).

<sup>†</sup>Institut de Mathématiques de Toulouse; UMR5219; Université de Toulouse; CNRS; UPS, F-31062 Toulouse, France. E-mail: serban.belinschi@math.univ-toulouse.fr, mireille.capitaine@math.univ-toulouse.fr, guillaume.cebron@math.univ-toulouse.fr

<sup>‡</sup>Institut de Mathématiques de Marseille; UMR7373; Aix-Marseille Université; CNRS; 13288 Marseille, France. E-mail: charles.bordenave@univ-amu.fr

The following theorem is the culmination of the work of many authors [2, 3, 19, 20, 25, 28, 35, 37].

**Theorem 1.1.** *Let  $X_N$  be an i.i.d. random matrix. Then the empirical spectral measure of  $\frac{X_N}{\sqrt{N}}$  converges almost surely to the circular measure  $\mu_c$  where  $d\mu_c = \frac{1}{\pi} \mathbf{1}_{|z| \leq 1} dz$ .*

One can prove that when the fourth moment is finite, almost surely for large  $N$ , there are no significant outliers to the circular law.

**Theorem 1.2.** *(Theorem 1.1 in [12] and Theorem 1.4 in [38]) Let  $X_N$  be an i.i.d. random matrix. Then the spectral radius  $\rho(\frac{X_N}{\sqrt{N}}) = \sup_{1 \leq j \leq N} \left| \lambda_j \left( \frac{X_N}{\sqrt{N}} \right) \right|$  converges to 1 in probability as  $N$  goes to infinity. If moreover the entries have finite fourth moment  $\mathbb{E}(|X_{11}|^4) < +\infty$ , then the spectral radius converges to 1 almost surely as  $N$  goes to infinity.*

An additive low rank perturbation  $A_N$  can create outliers outside the unit disk. Actually, when  $A_N$  has bounded rank and bounded operator norm and the entries of the i.i.d. matrix have finite fourth moment, Tao proved that outliers outside the unit disk are stable in the sense that outliers of  $N^{-1/2}X_N + A_N$  and  $A_N$  coincide asymptotically.

**Theorem 1.3.** *([38]) Let  $X_N$  be an i.i.d. random matrix whose entries have finite fourth moment. Let  $A_N$  be a deterministic matrix with rank  $O(1)$  and operator norm  $O(1)$ . Let  $\epsilon > 0$ , and suppose that for all sufficiently large  $N$ , there are:*

- no eigenvalues of  $A_N$  in  $\{z \in \mathbb{C} : 1 + \epsilon < |z| < 1 + 3\epsilon\}$ ,
- $j = O(1)$  eigenvalues  $\lambda_1(A_N), \dots, \lambda_j(A_N)$  in  $\{z \in \mathbb{C} : |z| \geq 1 + 3\epsilon\}$ .

Then, a.s, for sufficiently large  $N$ , there are precisely  $j$  eigenvalues of  $\frac{X_N}{\sqrt{N}} + A_N$  in  $\{z \in \mathbb{C} : |z| \geq 1 + 2\epsilon\}$  and after labeling these eigenvalues properly, as  $N$  goes to infinity, for each  $1 \leq i \leq j$ ,

$$\lambda_i\left(\frac{X_N}{\sqrt{N}} + A_N\right) = \lambda_i(A_N) + o(1).$$

Two different ways of generalization of this result were subsequently considered.

Firstly, [10] investigated the same problem but dealing with full rank additive perturbations. The main terminology related to free probability theory which is used in the following is defined in Section 3 below. Consider the deformed model:

$$S_N = A_N + \frac{X_N}{\sqrt{N}}, \tag{1.1}$$

where  $A_N$  is an  $N \times N$  deterministic matrix with operator norm  $O(1)$  and such that  $A_N$ , as a noncommutative variable from the set of  $N \times N$  matrices with complex entries endowed with the normalized trace  $(\mathcal{M}_N(\mathbb{C}), \text{tr}_N)$ , converges in  $*$ -moments to some noncommutative random variable  $a$  in some  $C^*$ -probability space  $(\mathcal{A}, \varphi)$ . According to Dozier and Silverstein [17], for any  $z \in \mathbb{C}$ , almost surely the empirical spectral measure of  $(S_N - zI_N)(S_N - zI_N)^*$  converges weakly to a nonrandom distribution  $\mu_z$  which is the distribution of  $(c + a - z)(c + a - z)^*$  where  $c$  is a circular operator which is  $*$ -free from  $a$  in  $(\mathcal{A}, \varphi)$ .

**Remark 1.4.** Note that for any operator  $K$  in some  $C^*$ -probability space  $(\mathcal{B}, \tau)$ ,  $K$  is invertible if and only if  $KK^*$  and  $K^*K$  are invertible. If  $\tau$  is tracial, the distribution  $\mu_{KK^*}$  of  $KK^*$  coincides with the distribution of  $K^*K$ . Therefore, if  $\tau$  is faithful and tracial, we have that  $0 \notin \text{supp}(\mu_{KK^*})$  if and only if  $K$  is invertible.

Therefore, since we can assume that  $\varphi$  is faithful and tracial,  $\text{spect}(c + a) = \{z \in \mathbb{C} : 0 \in \text{supp}(\mu_z)\}$ , where  $\text{spect}$  denotes the spectrum. Actually, we will present some results of [10] only in terms of the spectrum of  $c + a$  so that we do not need the

assumption (A3) in [10] on the limiting empirical spectral measure of  $S_N$ . The authors in [10] gave a sufficient condition to guarantee that outliers of the deformed model (1.1) outside the spectrum of  $c + a$  are stable. For this purpose, they introduced the notion of well-conditioned matrix which is related to the phenomenon of lack of outlier and of well-conditioned decomposition of  $A_N$  which lead to the statement of a sufficient condition for the stability of the outliers. We will denote by  $s_1(B) \geq \dots \geq s_N(B)$  the singular values of any  $N \times N$  matrix  $B$ . For any set  $K \subset \mathbb{C}$  and any  $\epsilon > 0$ ,  $B(K, \epsilon)$  stands for the set  $\{z \in \mathbb{C} : d(z, K) \leq \epsilon\}$ ,  $d$  denoting the Euclidean distance.

**Definition 1.5.** Let  $\Gamma \subset \mathbb{C} \setminus \text{spect}(c + a)$  be a compact set.  $A_N$  is well-conditioned in  $\Gamma$  if for any  $z \in \Gamma$ , there exists  $\eta_z > 0$  such that for all  $N$  large enough,  $s_N(A_N - zI_N) > \eta_z$ .

**Theorem 1.6.** ([10]) Assume that  $A_N$  is well-conditioned in  $\Gamma$ . Then, a.s. for all  $N$  large enough,  $S_N$  has no eigenvalue in  $\Gamma$ .

**Corollary 1.7.** ([10]) If for any  $z \in \mathbb{C} \setminus \text{spect}(c + a)$ , there exists  $\eta_z > 0$  such that for all  $N$  large enough,  $s_N(A_N - zI_N) > \eta_z$ , then, for any  $\epsilon > 0$ , a.s. for all  $N$  large enough, all eigenvalues of  $S_N$  are in  $B(\text{spect}(c + a), \epsilon)$ .

Let us introduce now the notion of well-conditioned decomposition of  $A_N$  which allowed [10] to exhibit a sufficient condition for stability of outliers.

**Definition 1.8.** Let  $\Gamma \subset \mathbb{C} \setminus \text{spect}(c + a)$  be a compact set.  $A_N$  admits a well-conditioned decomposition if:  $A_N = A'_N + A''_N$  where

- There exists  $M > 0$  such that for all  $N$ ,  $\|A'_N\| + \|A''_N\| \leq M$ , where  $\|\cdot\|$  denotes the operator norm.
- For any  $z \in \Gamma$ , there exists  $\eta_z > 0$  such that for all  $N$  large enough,  $s_N(A'_N - zI_N) > \eta_z$  (i.e  $A'_N$  is well-conditioned in  $\Gamma$ ) and  $A''_N$  has a fixed rank  $r$ .

**Theorem 1.9.** ([10]) Let  $\Gamma \subset \mathbb{C} \setminus \text{spect}(c + a)$  be a compact set with continuous boundary. Assume that  $A_N$  admits a well-conditioned decomposition:  $A_N = A'_N + A''_N$ . If for some  $\epsilon > 0$  and all  $N$  large enough,

$$\min_{z \in \partial\Gamma} \left| \frac{\det(A_N - z)}{\det(A'_N - z)} \right| \geq \epsilon, \tag{1.2}$$

then a.s. for all  $N$  large enough, the numbers of eigenvalues of  $A_N$  and  $S_N$  in  $\Gamma$  are equal.

On the other hand, in [16], the authors investigate the outliers of several types of bounded rank perturbations of the product of  $m$  independent random matrices  $X_{N,i}$ ,  $i = 1, \dots, m$  with i.i.d. entries. More precisely, they study the eigenvalues outside the unit disk, of the following three deformed models where  $A_N$  and the  $A_{N,j}$ 's denote  $N \times N$  deterministic matrices with rank  $O(1)$  and norm  $O(1)$ :

1.  $\prod_{k=1}^m \left( \frac{X_{N,k}}{\sqrt{N}} + A_{N,k} \right)$ ;
2. the product, in some fixed order, of the  $m + s$  terms  $\frac{X_{N,k}}{\sqrt{N}}, k = 1, \dots, m, (I_N + A_{N,j}), j = 1, \dots, s$ ;
3.  $\prod_{k=1}^m \frac{X_{N,k}}{\sqrt{N}} + A_N$ .

Set  $\mathbf{Y}_N = \left( \frac{X_{N,k}}{\sqrt{N}}, k = 1, \dots, m \right)$  and denote by  $\mathbf{A}_N$  the tuple of perturbations, that is  $\mathbf{A}_N = (A_{N,k}, k = 1, \dots, m)$  in case 1.,  $\mathbf{A}_N = (A_{N,j}, j = 1, \dots, s)$  in case 2. and  $\mathbf{A}_N = A_N$  in case 3.. In all cases 1.,2.,3., the model is some particular polynomial in  $\mathbf{Y}_N$  and  $\mathbf{A}_N$ , let us say  $P_i(\mathbf{Y}_N, \mathbf{A}_N)$ ,  $i = 1, 2, 3$ . It turns out that, according to [16], for each  $i = 1, 2, 3$ , the

eigenvalues of  $P_i(\mathbf{Y}_N, \mathbf{A}_N)$  and  $P_i(0, \mathbf{A}_N)$  outside the unit disk coincide asymptotically. Note that the unit disk is equal to the spectrum of each  $P_i(c, 0)$ ,  $i = 1, 2, 3$ , where  $c$  is a free  $m$ -circular system.

**1.2 Assumptions and results**

To begin with, we introduce some notations.

- $\mathcal{M}_p(\mathbb{C})$  is the set of  $p \times p$  matrices with complex entries,  $\mathcal{M}_p^{sa}(\mathbb{C})$  the subset of self-adjoint elements of  $\mathcal{M}_p(\mathbb{C})$  and  $I_p$  the identity matrix.
- $\text{Tr}_p$  denotes the trace and  $\text{tr}_p = \frac{1}{p} \text{Tr}_p$  the normalized trace on  $\mathcal{M}_p(\mathbb{C})$ .
- $\|\cdot\|$  denotes the operator norm on  $\mathcal{M}_p(\mathbb{C})$ .
- $\text{id}_p$  denotes the identity operator from  $\mathcal{M}_p(\mathbb{C})$  to  $\mathcal{M}_p(\mathbb{C})$ .

In this paper we generalize the previous results from [10] to non-Hermitian polynomials in several independent i.i.d. matrices and deterministic matrices. Note that our results include in particular the previous results from [16]. Here are the matricial models we deal with. Let  $t$  and  $u$  be fixed nonzero natural numbers independent from  $N$ .

**(A1)**  $(A_N^{(1)}, \dots, A_N^{(t)})$  is a  $t$ -tuple of  $N \times N$  deterministic matrices such that

1. for any  $i = 1, \dots, t$ ,

$$\sup_N \|A_N^{(i)}\| < \infty, \tag{1.3}$$

2. the  $t$ -tuple  $(A_N^{(1)}, \dots, A_N^{(t)})$  converges in  $*$ -distribution to a  $t$ -tuple  $a = (a^{(1)}, \dots, a^{(t)})$  in some  $C^*$ -probability space  $(\mathcal{A}, \varphi)$  where  $\varphi$  is faithful and tracial.

**(X1)** We consider  $u$  independent  $N \times N$  random matrices  $X_N^{(v)} = [X_{ij}^{(v)}]_{i,j=1}^N$ ,  $v = 1, \dots, u$ , where, for each  $v$ ,  $[X_{ij}^{(v)}]_{i \geq 1, j \geq 1}$  is an infinite array of random variables such that  $\{\sqrt{2}\Re(X_{ij}^{(v)}), \sqrt{2}\Im(X_{ij}^{(v)}), i \geq 1, j \geq 1\}$  are independent identically distributed centred random variables with variance 1 and finite fourth moment.

Let  $P$  be a polynomial in  $t + u$  noncommutative indeterminates where  $P$ ,  $t$  and  $u$  are fixed, independent of  $N$ , and define

$$M_N = P \left( \frac{X_N^{(1)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)} \right).$$

Note that we do not need any assumption on the convergence of the empirical spectral measure of  $M_N$ . Let  $c = (c^{(1)}, \dots, c^{(u)})$  be a free noncommutative circular system in  $(\mathcal{A}, \varphi)$  which is free from  $a = (a^{(1)}, \dots, a^{(t)})$ . According to the second assertion of Proposition 5.2 below, for any  $z \in \mathbb{C}$ , almost surely, the empirical spectral measure of  $(M_N - zI_N)(M_N - zI_N)^*$  converges weakly to  $\mu_z$  where  $\mu_z$  is the distribution of  $[P(c, a) - z1][P(c, a) - z1]^*$ . Since we can assume that  $\varphi$  is faithful and tracial, we have by Remark 1.4 that

$$\text{spect}(P(c, a)) = \{z \in \mathbb{C} : 0 \in \text{supp}(\mu_z)\}. \tag{1.4}$$

Define

$$M_N^{(0)} = P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}),$$

where  $0_N$  denotes the  $N \times N$  null matrix. Throughout the whole paper, we will call *outlier* any eigenvalue of  $M_N$  or  $M_N^{(0)}$  outside  $\mathbb{C} \setminus \text{spect}(P(c, a))$ . We are now interested in describing the individual eigenvalues of  $M_N$  outside  $B(\text{spect}(P(c, a)), \epsilon)$  for some  $\epsilon > 0$ . In the lineage of [10], our main result gives a sufficient condition to guarantee that

outliers are stable in the sense that outliers of  $M_N$  and  $M_N^{(0)}$  coincide asymptotically. To state this formally: consider an arbitrary open set  $G$  which is relatively compact in  $\mathbb{C} \setminus \text{spect}(P(c, a))$  (that is, the closure  $\overline{G} \subset \mathbb{C} \setminus \text{spect}(P(c, a))$  is compact). Then one expects that an outlier of  $M_N$  occurs in  $G$  if and only if an outlier of  $M_N^{(0)}$  occurs in  $G$ . For technical reasons, namely the need to be able to apply Rouché’s Theorem, we assume that  $G$  is sufficiently “nice”: specifically, we require that  $\Gamma := \overline{G}$  is a compact set in  $\mathbb{C} \setminus \text{spect}(P(c, a))$ , that the closure of any connected component of  $G$  is a connected component of  $\Gamma$ , that the fundamental group of each component of  $G$  equals the fundamental group of its closure, and that the topological boundary of  $\Gamma$  is a finite union of rectifiable curves (in particular,  $G$  has finitely many connected components). Let us see what kinds of situations these conditions are intended to exclude:

- Say  $\text{spect}(P(c, a)) = \{z \in \mathbb{C} : |z| \leq 1, |z - n^{-1}| \geq 4^{-n}, n \geq 3\}$ , a unit disk with countably many small disks removed from it. Then the set  $G = \bigcup_{n=3}^{\infty} \{z \in \mathbb{C} : |z - n^{-1}| < 8^{-n}\}$  is included in  $\mathbb{C} \setminus \text{spect}(P(c, a))$ , but its closure is not.  $\Gamma = \overline{G}$  is not compact in  $\mathbb{C} \setminus \text{spect}(P(c, a))$ , and  $\partial\Gamma$  is a countable, but not finite, union of rectifiable curves;
- Say  $\text{spect}(P(c, a)) = \{z \in \mathbb{C} : |z - 100| \leq 1\}$ . The set  $G = \{z \in \mathbb{C} : -1 < \Re z < 1, -1 < \Im z < 1, |z \pm 1| > 1\}$  has two connected components (above and below zero), while  $\Gamma = \overline{G}$  is connected;
- For the same  $\text{spect}(P(c, a))$ , the set  $G = \{z \in \mathbb{C} : |z| < 1, |z - 1/2| > 1/2\}$  is simply connected, but  $\Gamma$  isn’t.

While this is a long list of restrictions, they hardly constitute a reduction in generality. To begin with, a statement regarding outliers of  $M_N$  and  $M_N^{(0)}$  coinciding asymptotically means that the outlier of one eventually ends up in a neighborhood of the outlier of the other, so it is natural to choose an open  $G$ . Second, if an open set  $G$  fails to satisfy any of the conditions we require, then there is an arbitrarily small perturbation of  $G$  (in the Hausdorff topology, say) which satisfies all of them, with the possible exception of the requirement of rectifiability for  $\partial\Gamma$ , which is the only somewhat restrictive requirement. And finally, from the perspective of an agent who knows the positions of the outliers of one of  $M_N, M_N^{(0)}$  and wishes to identify the outliers of the other, it is effectively enough to consider the case when  $G$  is a finite union of open disks with mutually disjoint closures.

**Theorem 1.10.** *Assume that hypotheses (A1), (X1) hold. Let  $G$  be an open relatively compact subset of  $\mathbb{C} \setminus \text{spect}(P(c, a))$ , and let  $\Gamma = \overline{G}$ . We assume that (i)  $\partial\Gamma$  is a finite union of rectifiable curves; (ii) the closure of any connected component of  $G$  is a connected component of  $\Gamma$ ; and (iii) the fundamental group of each connected component of  $G$  coincides with the fundamental group of its closure. Assume moreover that*

(A2) for  $k = 1, \dots, t$ ,  $A_N^{(k)} = (A_N^{(k)})' + (A_N^{(k)})''$ ,

where  $(A_N^{(k)})''$  has a bounded rank  $r_k(N) = O(1)$  and  $\left( (A_N^{(1)})', \dots, (A_N^{(t)})' \right)$  satisfies

- (A2') for any  $z$  in  $\Gamma$ , there exists  $\eta_z > 0$  such that for all  $N$  large enough, there is no singular value of

$$P(0_N, \dots, 0_N, (A_N^{(1)})', \dots, (A_N^{(t)})') - zI_N$$

in  $[0, \eta_z]$ .

- for any  $k = 1, \dots, t$ ,

$$\sup_N \|(A_N^{(k)})'\| < +\infty. \tag{1.5}$$

If for some  $\epsilon > 0$ , for all large  $N$ ,

$$\min_{z \in \partial \Gamma} \left| \frac{\det(zI_N - P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}))}{\det(zI_N - P(0_N, \dots, 0_N, (A_N^{(1)})', \dots, (A_N^{(t)})'))} \right| \geq \epsilon \tag{1.6}$$

then almost surely for all large  $N$ , the numbers of eigenvalues of  $M_N^{(0)} = P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)})$  and  $M_N = P\left(\frac{X_N^{(1)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)}\right)$  in  $\Gamma$  are equal.

The next statement is an easy consequence of Theorem 1.10.

**Corollary 1.11.** Assume that (X1) holds and that, for  $k = 1, \dots, t$ ,  $A_N^{(k)}$  are deterministic matrices with rank  $O(1)$  and operator norm  $O(1)$ . Let  $\epsilon > 0$ , and suppose that for all sufficiently large  $N$ , there are no eigenvalues of  $M_N^{(0)} = P(0, \dots, 0, A_N^{(1)}, \dots, A_N^{(t)})$  in  $\{z \in \mathbb{C}, \epsilon < d(z, \text{spect}(P(c, 0))) < 4\epsilon\}$ , and there are  $j$  eigenvalues  $\lambda_1(M_N^{(0)}), \dots, \lambda_j(M_N^{(0)})$  for some  $j = O(1)$  in the region  $\{z \in \mathbb{C}, d(z, \text{spect}(P(c, 0))) \geq 3\epsilon\}$ . Then, almost surely, for all large  $N$ , there are precisely  $j$  eigenvalues of  $M_N = P\left(\frac{X_N^{(1)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)}\right)$  in  $\{z \in \mathbb{C}, d(z, \text{spect}(P(c, 0))) \geq 2\epsilon\}$  and after labeling these eigenvalues properly, for each  $1 \leq i \leq j$ ,

$$\lambda_i(M_N) = \lambda_i(M_N^{(0)}) + o(1).$$

**Remark 1.12.** It is sufficient to prove Theorem 1.10 and Corollary 1.11 for a noncommutative polynomial  $P$  with no constant term, that is, such that  $P(0, \dots, 0) = 0$ ; the general result follows easily by translation.

We will first prove Theorem 1.10 in the case when all ranks  $r_k(N), 1 \leq k \leq t$ , are equal to zero.

**Theorem 1.13.** Suppose that assumptions of Theorem 1.10 hold with, for any  $k = 1, \dots, t$ ,  $(A_N^{(k)})'' = 0, A_N^{(k)} = (A_N^{(k)})'$  and  $\Gamma \subset \mathbb{C} \setminus \text{spect}(P(c, a))$  a compact set. Then, a.s. for all  $N$  large enough,  $M_N$  has no eigenvalue in  $\Gamma$ .

In particular, if assumptions of Theorem 1.10 hold with, for any  $k = 1, \dots, t$ ,  $(A_N^{(k)})'' = 0, A_N^{(k)} = (A_N^{(k)})'$  and  $\Gamma = \mathbb{C} \setminus \text{spect}(P(c, a))$  then for any  $\epsilon > 0$ , a.s. for all  $N$  large enough, all eigenvalues of  $M_N$  are in  $B(\text{spect}(P(c, a)), \epsilon)$ .

While Theorem 1.10 requires supplementary hypotheses on the compact  $\Gamma$ , those hypotheses are clearly not necessary in Theorem 1.13.

To prove Theorems 1.13 and 1.10, we make use of a linearization procedure which brings the study of the polynomial back to that of the sum of matrices in a higher dimensional space. Then, this allows us to follow the approach of [10]. But for this purpose, we need to establish substantial operator-valued free probability results.

In Section 2, we present our theoretical results and corresponding simulations for four examples of random polynomial matrix models. Section 3.2 provides required definitions and preliminary results on operator-valued free probability theory. Section 4 describes the fundamental linearization trick as introduced in [1, Proposition 3]. In Sections 5 and 6, we establish Theorems 1.13 and 1.10 respectively.

## 2 Related results and examples

Recall that we do not need any assumption on the convergence of the empirical spectral measure of  $M_N$ . However, the convergence in  $*$ -distribution of  $\left(\frac{X_N^{(1)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)}\right)$  to  $(c, a) = (c^{(1)}, \dots, c^{(u)}, a^{(1)}, \dots, a^{(t)})$  (see Proposition 5.2) implies the

convergence in \*-distribution of

$$M_N = P \left( \frac{X_N^{(1)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)} \right)$$

to  $P(c, a)$ . In this situation, a good candidate to be the limit of the empirical spectral distribution of  $M_N$  is the *Brown measure*  $\mu_{P(c,a)}$  of  $P(c, a)$  (see [14]). Unfortunately, the convergence of the empirical spectral distribution of  $M_N$  to  $\mu_{P(c,a)}$  is still an open problem for an arbitrary polynomial.

In the following three examples, we will consider the particular situation where we can decompose

$$M_N = \alpha \frac{X_N^{(1)}}{\sqrt{N}} + Q \left( \frac{X_N^{(2)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)} \right),$$

with  $\alpha > 0$ ,  $X_N^{(1)}$  a Ginibre matrix and  $Q$  an arbitrary polynomial. Indeed, in this case, a beautiful result of Śniady [33] ensures that the empirical spectral distribution of  $M_N$  converges to  $\mu_{P(c,a)}$ . Thus, the description of the limiting spectrum of  $M_N$  inside  $\text{supp}(\mu_{P(c,a)})$  is a question of computing explicitly  $\mu_{P(c,a)}$  (a quite hard problem, which can be handled numerically by [7]). On the other hand, Theorem 1.10 explains the behaviour of the spectrum of  $M_N$  outside  $\text{spect}(P(c, a))$ . Thus, we have a complete description of the limiting spectrum of  $M_N$ , except potentially in the set  $\text{spect}(P(c, a)) \setminus \text{supp}(\mu_{P(c,a)})$  which is not necessarily empty (even if it is empty in the majority of the examples known, see [13]).

For an arbitrary polynomial, we only know that any limit point of the empirical spectral distribution of  $M_N$  is a balayée of the measure  $\mu_{P(c,a)}$  (see [13, Corollary 2.2]), which implies that the support of any such limit point is contained in  $\text{supp}(\mu_{P(c,a)})$ , and in particular is contained in  $\text{spect}(P(c, a))$ .

### 2.1 Example 1

We consider the matrix

$$\begin{aligned} M_N &= P_1 \left( \frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N \right) \\ &= \frac{3}{2} \frac{X_N^{(1)}}{\sqrt{N}} + \frac{1}{6} \left( \frac{X_N^{(2)}}{\sqrt{N}} \right)^2 A_N + \frac{1}{6} \frac{X_N^{(2)}}{\sqrt{N}} \frac{X_N^{(3)}}{\sqrt{N}} A_N \frac{X_N^{(3)}}{\sqrt{N}} + A_N^2 \frac{X_N^{(3)}}{\sqrt{N}} + A_N + \frac{1}{8} A_N^2, \end{aligned}$$

where  $X_N^{(1)}, X_N^{(2)}, X_N^{(3)}$  are i.i.d. Gaussian matrices and

$$A_N = \begin{pmatrix} 2 & & & & \\ & 2i & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}.$$

The matrix  $M_N$  converges in \*-distribution to  $\frac{3}{2}c$ , where  $c$  is a circular variable, and the empirical spectral measure of  $M_N$  converges to the Brown measure of  $c$ , which is the uniform law on the centered disk of radius  $3/2$  by [13]. This disk is also the spectrum of  $\frac{3}{2}c$ . Our theorem says that, outside this disk, the outliers of  $M_N$  are close to the eigenvalues  $2.5$  and  $2i - 0.5$  of  $P_1(0_N, 0_N, 0_N, A_N) = A_N + \frac{1}{8}A_N^2$  (see Figure 1).

## Outlier eigenvalues for non-Hermitian polynomials

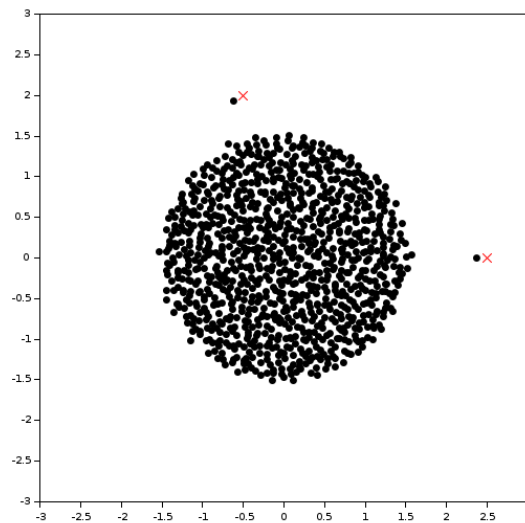


Figure 1: In black, the eigenvalues of  $P_1\left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N\right)$  for  $N = 1000$ , and in red, the outliers  $2.5$  and  $2i - 0.5$  of  $P_1(0_N, 0_N, 0_N, A_N)$ .

### 2.2 Example 2

We consider the matrix

$$\begin{aligned} M_N &= P_2\left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N^{(1)}, A_N^{(2)}\right) \\ &= \frac{1}{2} \frac{X_N^{(1)}}{\sqrt{N}} + \frac{1}{6} A_N^{(1)} \frac{X_N^{(2)}}{\sqrt{N}} \left( A_N^{(2)} + A_N^{(1)} + \frac{X_N^{(3)}}{\sqrt{N}} \right) \frac{X_N^{(2)}}{\sqrt{N}} + A_N^{(2)} \frac{X_N^{(3)}}{\sqrt{N}} A_N^{(1)} \\ &\quad + A_N^{(1)} + \frac{1}{2} A_N^{(2)}, \end{aligned}$$

where  $X_N^{(1)}, X_N^{(2)}, X_N^{(3)}$  are i.i.d. Gaussian matrices,

$$A_N^{(1)} = \begin{pmatrix} 2 & & & & \\ & -2.5 & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}$$

and  $A_N^{(2)}$  is a realization of a G.U.E. matrix.

The matrix  $M_N$  converges in  $*$ -distribution to the elliptic variable  $\frac{1}{2}(c + s)$ , where  $c$  is a circular variable and  $s$  a semicircular variable free from  $c$ . The empirical spectral measure of  $M_N$  converges to the Brown measure of  $\frac{1}{2}(c + s)$ , which is the uniform law on the interior of the ellipse  $\{\frac{3}{2\sqrt{2}} \cos(\theta) + i\frac{1}{2\sqrt{2}} \sin(\theta) : 0 \leq \theta < 2\pi\}$  by [13]. The interior of this ellipse is also the spectrum of  $\frac{1}{2}(c + s)$ . Our theorem says that, outside this ellipse, the outliers of  $M_N$  are closed to the outliers of  $P_2(0_N, 0_N, 0_N, A_N^{(1)}, A_N^{(2)}) = A_N^{(1)} + \frac{1}{2} A_N^{(2)}$  (see Figure 2). Moreover, the outliers of  $A_N^{(1)} + \frac{1}{2} A_N^{(2)}$  are those of an additive perturbation of a G.U.E. matrix, and converges to  $2.125$  and  $-2.6$  by [29].



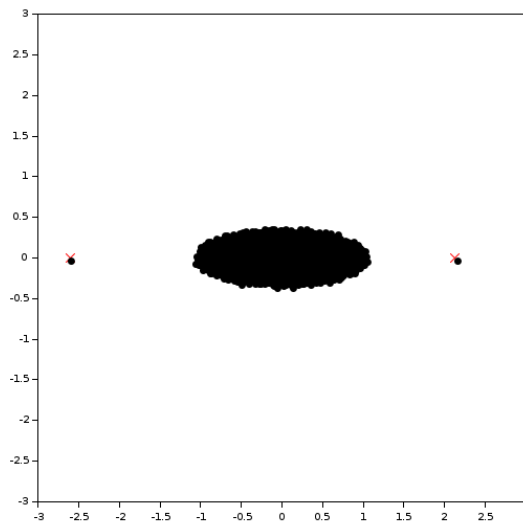


Figure 2: In black, the eigenvalues of  $P_2\left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N^{(1)}, A_N^{(2)}\right)$  for  $N = 1000$ , and in red, the limiting outliers  $2.125$  and  $-2.6$  of  $P_1(0, 0, 0, A_N^{(1)}, A_N^{(2)})$ .

### 2.3 Example 3

We consider the matrix

$$\begin{aligned}
 M_N &= P_3\left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N^{(1)}, A_N^{(2)}\right) \\
 &= \frac{X_N^{(1)}}{\sqrt{N}} + A_N^{(1)} + A_N^{(2)} + A_N^{(1)} \frac{X_N^{(2)}}{\sqrt{N}} A_N^{(2)} \frac{X_N^{(2)}}{\sqrt{N}} + \frac{X_N^{(3)}}{\sqrt{N}} A_N^{(2)} \frac{X_N^{(2)}}{\sqrt{N}},
 \end{aligned}$$

where  $X_N^{(1)}, X_N^{(2)}, X_N^{(3)}$  are i.i.d. Gaussian matrices,

$$A_N^{(1)} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix}$$

is a matrix whose empirical spectral distribution converges to  $\frac{1}{2}(\delta_1 + \delta_{-1})$  and

$$A_N^{(2)} = \begin{pmatrix} 1.5 & & & & \\ & -2 + 2i & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}$$

The matrix  $M_N$  converges in  $*$ -distribution to the random variable  $c + a$ , where  $c$  is a circular variable and  $a$  is a self-adjoint random variable, free from  $c$ , and whose distribution is  $\frac{1}{2}(\delta_1 + \delta_{-1})$ . The empirical spectral measure of  $M_N$  converges to the Brown

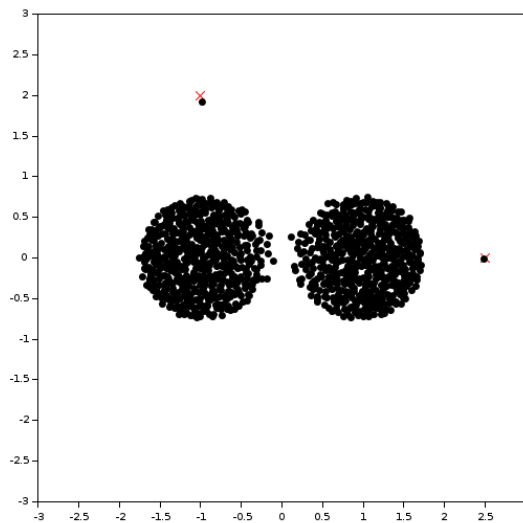


Figure 3: In black, the eigenvalues of  $P_3\left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N^{(1)}, A_N^{(2)}\right)$  for  $N = 1000$ , and in red, the outliers 2.5 and  $-1 + 2i$  of  $P_3(0_N, 0_N, 0_N, A_N^{(1)}, A_N^{(2)})$ .

measure of  $c + a$ , which is absolutely continuous and whose support is the region inside the lemniscate-like curve in the complex plane with the equation  $\{z \in \mathbb{C} : |z^2 + 1|^2 = |z|^2 + 1\}$  by [13]. This region is also the spectrum of  $c + a$ . Our theorem says that, outside this region, the outliers of  $M_N$  are closed to the outliers 2.5 and  $-1 + 2i$  of  $P_3(0_N, 0_N, 0_N, A_N^{(1)}, A_N^{(2)}) = A_N^{(1)} + A_N^{(2)}$  (see Figure 3).

### 2.4 Example 4

We consider the matrix

$$\begin{aligned} M_N &= P_4\left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N\right) \\ &= \frac{1}{5}\left(\frac{X_N^{(1)}}{\sqrt{N}} + 3I_N\right)\left(\frac{X_N^{(2)}}{\sqrt{N}} + A_N + 2I_N\right)\left(\frac{X_N^{(3)}}{\sqrt{N}} + 2I_N\right) - 2I_N, \end{aligned}$$

where  $X_N^{(1)}, X_N^{(2)}, X_N^{(3)}$  are i.i.d. Gaussian matrices and

$$A_N = \begin{pmatrix} 2i & & & & \\ & -2i & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}.$$

The matrix  $M_N$  converges in  $*$ -distribution to the random variable  $(c_1 + 3)(c_2 + 2)(c_3 + 2)/5 - 2$ , where  $c_1, c_2, c_3$  are free circular variables. It is expected (but not proved) that the empirical spectral measure of  $M_N$  converges to the Brown measure of  $(c_1 + 3)(c_2 + 2)(c_3 + 2)/5 - 2$ . The spectrum of  $(c_1 + 3)(c_2 + 2)(c_3 + 2)/5 - 2$  is included in the set  $(B(0, 1) + 3)(B(0, 1) + 2)(B(0, 1) + 2)/5 - 2$ . Our theorem says that, outside this set, the outliers of  $M_N$  are closed to the outliers  $-2 + 2.4i$  and  $-2 - 2.4i$  of  $P_4(0_N, 0_N, 0_N, A_N) = \frac{6}{5}A_N - 2I_N$  (see Figure 4).

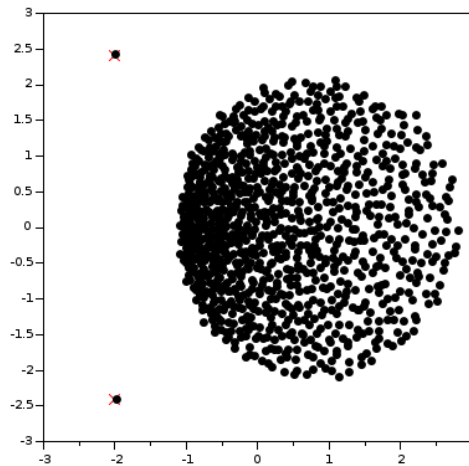


Figure 4: In black, the eigenvalues of  $P_4\left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N\right)$  for  $N = 1000$ , and in red, the outliers  $-2 + 2.4i$  and  $-2 - 2.4i$  of  $P_4(0_N, 0_N, 0_N, A_N)$ .

### 3 Free probability theory

#### 3.1 Scalar-valued free probability theory

For the reader’s convenience, we recall the following basic definitions from free probability theory. For a thorough introduction to free probability theory, we refer to [43].

- A  $C^*$ -probability space is a pair  $(\mathcal{A}, \varphi)$  consisting of a unital  $C^*$ -algebra  $\mathcal{A}$  and a state  $\varphi$  on  $\mathcal{A}$  (i.e. a linear map  $\varphi : \mathcal{A} \rightarrow \mathbb{C}$  such that  $\varphi(1_{\mathcal{A}}) = 1$  and  $\varphi(aa^*) \geq 0$  for all  $a \in \mathcal{A}$ ).  $\varphi$  is a trace if it satisfies  $\varphi(ab) = \varphi(ba)$  for every  $(a, b) \in \mathcal{A}^2$ . A trace is said to be faithful if  $\varphi(aa^*) > 0$  whenever  $a \neq 0$ . An element of  $\mathcal{A}$  is called a noncommutative random variable.
- The  $*$ -noncommutative distribution of a family  $a = (a_1, \dots, a_k)$  of noncommutative random variables in a  $C^*$ -probability space  $(\mathcal{A}, \varphi)$  is defined as the linear functional  $\mu_a : P \mapsto \varphi(P(a, a^*))$  defined on the set of polynomials in  $2k$  noncommutative indeterminates, where  $(a, a^*)$  denotes the  $2k$ -tuple  $(a_1, \dots, a_k, a_1^*, \dots, a_k^*)$ . For any self-adjoint element  $a_1$  in  $\mathcal{A}$ , there exists a probability measure  $\nu_{a_1}$  on  $\mathbb{R}$  such that, for every polynomial  $P$ , we have

$$\mu_{a_1}(P) = \int P(t) d\nu_{a_1}(t).$$

Then, we identify  $\mu_{a_1}$  and  $\nu_{a_1}$ . If  $\varphi$  is faithful then the support of  $\nu_{a_1}$  is the spectrum of  $a_1$  and thus  $\|a_1\| = \sup\{|z|, z \in \text{supp}(\nu_{a_1})\}$ .

- A family of elements  $(a_i)_{i \in I}$  in a  $C^*$ -probability space  $(\mathcal{A}, \varphi)$  is free if for all  $k \in \mathbb{N}$  and all polynomials  $p_1, \dots, p_k$  in two noncommutative indeterminates, one has

$$\varphi(p_1(a_{i_1}, a_{i_1}^*) \cdots p_k(a_{i_k}, a_{i_k}^*)) = 0 \tag{3.1}$$

whenever  $i_1 \neq i_2, i_2 \neq i_3, \dots, i_{n-1} \neq i_k$  and  $\varphi(p_l(a_{i_l}, a_{i_l}^*)) = 0$  for  $l = 1, \dots, k$ .

- A noncommutative random variable  $x$  in a  $C^*$ -probability space  $(\mathcal{A}, \varphi)$  is a standard semicircular variable if  $x = x^*$  and for any  $k \in \mathbb{N}$ ,

$$\varphi(x^k) = \int t^k d\mu_{sc}(t)$$

where  $d\mu_{sc}(t) = \frac{1}{2\pi}\sqrt{4-t^2}\mathbf{1}_{[-2;2]}(t)dt$  is the semicircular standard distribution.

- Let  $k$  be a nonnull integer number. Denote by  $\mathcal{P}$  the set of polynomials in  $2k$  noncommutative indeterminates. A sequence of families of variables  $(a_n)_{n \geq 1} = (a_1(n), \dots, a_k(n))_{n \geq 1}$  in  $\mathcal{C}^*$ -probability spaces  $(\mathcal{A}_n, \varphi_n)$  converges, when  $n$  goes to infinity, in distribution if the map  $P \in \mathcal{P} \mapsto \varphi_n(P(a_n, a_n^*))$  converges pointwise.

### 3.2 Operator-valued free probability theory

#### 3.2.1 Basic definitions

Operator-valued distributions and the operator-valued version of free probability were introduced by Voiculescu in [39] with the main purpose of studying freeness with amalgamation. Thus, an operator-valued noncommutative probability space is a triple  $(M, E, B)$ , where  $M$  is a unital algebra over  $\mathbb{C}$ ,  $B \subseteq M$  is a unital subalgebra containing the unit of  $M$ , and  $E: M \rightarrow B$  is a unit-preserving conditional expectation, that is, a linear  $B$ -bimodule map such that  $E[1] = 1$ .

The subalgebras  $(M_i)_{i \in I}$  of  $M$  containing  $B$  are free with amalgamation over  $B$  (or free with amalgamation with respect to the conditional expectation  $E$ ), if whenever we have  $m \geq 2$  and  $x_1, \dots, x_m \in M$  such that  $E[x_j] = 0$  for  $j = 1, \dots, m$ ,  $x_j \in M_{i(j)}$  with  $i(j) \in I$  and  $i(1) \neq i(2), i(2) \neq i(3), \dots, i(m-1) \neq i(m)$ , then

$$E[x_1 \cdots x_m] = 0.$$

Elements in  $M$  are called free with amalgamation over  $B$  if the algebras generated by  $B$  and the elements are also so.

We will only need the more restrictive context in which  $M$  is a finite von Neumann algebra which is a factor,  $B$  is a finite-dimensional von Neumann subalgebra of  $M$  (and hence isomorphic to an algebra of matrices), and  $E$  is the unique trace-preserving conditional expectation from  $M$  to  $B$ . The  $B$ -valued distribution of an element  $X \in M$  w.r.t.  $E$  is defined to be the family of multilinear maps called the *moments* of  $\mu_X$ :

$$\mu_X = \{B^{n-1} \ni (b_1, b_2, \dots, b_{n-1}) \mapsto E[Xb_1Xb_2 \cdots Xb_{n-1}X] \in B : n \geq 0\},$$

with the convention that the first moment (corresponding to  $n = 1$ ) is the element  $E[X] \in B$ , and the zeroth moment (corresponding to  $n = 0$ ) is the unit 1 of  $B$  (or  $M$ ). The distribution of  $X$  is encoded conveniently by a noncommutative analytic transform defined for certain elements  $b \in B$ , which we agree to call the noncommutative Cauchy transform:

$$G_X(b) = E[(X - b)^{-1}].$$

(To be more precise, it is the noncommutative extension  $G_{X \otimes 1_n}(b) = (E \otimes \text{Id}_{\mathcal{M}_n(B)})[(X \otimes 1_n - b)^{-1}]$ , for elements  $b \in \mathcal{M}_n(B)$ , which completely encodes  $\mu_X$  - see [42]; since we do not need this extension, we shall not discuss it any further, but refer the reader to [42, 40, 41, 30] for details.) A natural domain for  $G_X$  is the upper half-plane of  $B$ ,  $H^+(B) = \{b \in B : \Im b > 0\}$ , where  $\Im b = \frac{b-b^*}{2i}$  is the imaginary part of  $b$ . It follows quite easily that  $G_X(H^+(B)) \subseteq H^+(B)$  - see [41].

We warn here the reader that we have changed conventions in our paper compared to [40, 41, 42], namely we have chosen  $G_X(b) = E[(X - b)^{-1}]$  instead of  $E[(b - X)^{-1}]$ , so that  $G_X$  preserves  $H^+(B)$ .

Among many other results proved in [39], one can find a central limit theorem for random variables which are free with amalgamation. The central limit distribution is called an *operator-valued semicircular*, by analogy with the free central limit for the usual, scalar-valued random variables, which is Wigner's semicircular distribution. It has been shown in [39] that an operator-valued semicircular distribution is entirely described

by its operator-valued free cumulants: only the first and second cumulants of an operator-valued semicircular distribution may be nonzero (see also [34, 42]). For our purposes, we use the equivalent description of an operator-valued semicircular distribution via its noncommutative Cauchy transform, as in [23]:  $S$  is a  $B$ -valued semicircular if and only if

$$G_S(b) = (m_1 - b - \eta(G_S(b)))^{-1}, \quad b \in H^+(B),$$

for some  $m_1 = m_1^* \in B$  and completely positive map  $\eta: B \rightarrow B$ . In that case,  $m_1 = E[S]$  and  $\eta(b) = E[SbS] - E[S]bE[S]$  is the operator-valued variance. The above equation is obviously a generalization of the quadratic equation determining Wigner’s semicircular distribution:  $\sigma^2 G_S(z)^2 + (z - m_1)G_S(z) + 1 = 0$ . Here  $m_1$  is the – classical – first moment of  $S$ , and  $\sigma^2$  its classical variance, which, as a linear completely positive map, is the multiplication with a positive constant. Unless otherwise specified, we shall from now on assume our semicirculars to be centered, i.e.  $m_1 = 0$ .

**Example 3.1.** A rich source of examples of operator-valued semicirculars comes in the case of finite dimensional  $B$  from scalar-valued semicirculars: assume that  $s_{i,j}, 1 \leq i \leq j \leq n$  are scalar-valued centered semicircular random variables of variance one. We do not assume them to be free. Then the matrix

$$\begin{pmatrix} \alpha_1 s_{1,1} & \gamma_{1,2} s_{1,2} & \gamma_{1,3} s_{1,3} & \cdots & \gamma_{1,n-1} s_{1,n-1} & \gamma_{1,n} s_{1,n} \\ \overline{\gamma_{1,2}} s_{1,2} & \alpha_2 s_{2,2} & \gamma_{2,3} s_{2,3} & \cdots & \gamma_{2,n-1} s_{2,n-1} & \gamma_{2,n} s_{2,n} \\ \overline{\gamma_{1,3}} s_{1,3} & \overline{\gamma_{2,3}} s_{2,3} & \alpha_3 s_{3,3} & \cdots & \gamma_{3,n-1} s_{3,n-1} & \gamma_{3,n} s_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \overline{\gamma_{1,n-1}} s_{1,n-1} & \overline{\gamma_{2,n-1}} s_{2,n-1} & \overline{\gamma_{3,n-1}} s_{3,n-1} & \cdots & \alpha_{n-1} s_{n-1,n-1} & \gamma_{n-1,n} s_{n-1,n} \\ \overline{\gamma_{1,n}} s_{1,n} & \overline{\gamma_{2,n}} s_{2,n} & \overline{\gamma_{3,n}} s_{3,n} & \cdots & \overline{\gamma_{n-1,n}} s_{n-1,n} & \alpha_n s_{n,n} \end{pmatrix},$$

where  $\alpha_1, \dots, \alpha_n \in [0, +\infty)$  and  $\gamma_{i,j} \in \mathbb{C}, 1 \leq i < j \leq n$ , is an  $\mathcal{M}_n(\mathbb{C})$ -valued semicircular. Note that we do allow our scalars to be zero. This is a particular case of a result from [31], and its proof can be found in great detail in [26].

An important fact about semicircular elements, both scalar- and operator-valued, is that the sum of two free semicircular elements is again a semicircular element (this follows from the fact that a semicircular is defined by having all its cumulants beyond the first two equal to zero – see [39]). In particular, if  $\{s_{1,1}^{(1)}, s_{1,2}^{(1)}, s_{2,2}^{(1)}, s_{1,1}^{(2)}, s_{1,2}^{(2)}, s_{2,2}^{(2)}\}$  are centered all semicirculars of variance one, and in addition we assume them to be free from each other, then  $\begin{bmatrix} s_{1,1}^{(1)} & s_{1,2}^{(1)} \\ s_{1,2}^{(1)} & s_{2,2}^{(1)} \end{bmatrix}$  and  $\begin{bmatrix} s_{1,1}^{(2)} & i s_{1,2}^{(2)} \\ -i s_{1,2}^{(2)} & s_{2,2}^{(2)} \end{bmatrix}$  are  $\mathcal{M}_2(\mathbb{C})$ -valued semicirculars

which are free over  $\mathcal{M}_2(\mathbb{C})$ , so their sum  $\begin{bmatrix} s_{1,1}^{(1)} + s_{1,1}^{(2)} & s_{1,2}^{(1)} + i s_{1,2}^{(2)} \\ s_{1,2}^{(1)} - i s_{1,2}^{(2)} & s_{2,2}^{(1)} + s_{2,2}^{(2)} \end{bmatrix}$  is also an  $\mathcal{M}_2(\mathbb{C})$ -valued semicircular, despite its off-diagonal elements not being anymore distributed according to the Wigner semicircular law. This is hardly surprising: the two matrices we have added are the limits of the real and imaginary parts of a G.U.E. (Gaussian Unitary Ensemble). The upper right corner of a G.U.E. is known to be a C.U.E. (Circular Unitary Ensemble), and its eigenvalues converge to the uniform law on a disk. On the other hand, direct analytic computations show that the sum  $s_{1,2}^{(1)} \pm i s_{1,2}^{(2)}$ , with  $s_{1,2}^{(1)}$  and  $s_{1,2}^{(2)}$  free from each other, has precisely the same law. Thus, the following definition, due to Voiculescu, is natural.

**Definition 3.2.** An element  $c$  in a \*-noncommutative probability space  $(\mathcal{A}, \varphi)$  is called a circular random variable if  $(c + c^*)/\sqrt{2}$  and  $(c - c^*)/\sqrt{2}i$ , respectively, are free from each other and identically distributed according to standard Wigner’s semicircular law.

**3.2.2 Preliminary results**

We first establish preliminary results in free probability theory that we will need in the following sections.

**Lemma 3.3.** *Let  $\{m_p^{(j)}, p = 1, \dots, 4, j = 1, \dots, t\}$  be noncommutative random variables in some noncommutative probability space  $(\mathcal{A}, \varphi)$ . Let  $s_i^{(1)}, s_i^{(2)}, i = 1, \dots, u$  be semicircular variables and  $c_i, i = 1, \dots, u$  be circular variables such that  $s_1^{(1)}, \dots, s_u^{(1)}, s_1^{(2)}, \dots, s_u^{(2)}, c_1, \dots, c_u, \{m_p^{(j)}, p = 1, \dots, 4, j = 1, \dots, t\}$  are free in  $(\mathcal{A}, \varphi)$ . Define for  $i = 1, \dots, u$ ,*

$$s_i = \frac{1}{\sqrt{2}} \begin{pmatrix} s_i^{(1)} & c_i \\ c_i^* & s_i^{(2)} \end{pmatrix},$$

and for  $j = 1, \dots, t$ ,

$$m_j = \begin{pmatrix} m_1^{(j)} & m_2^{(j)} \\ m_3^{(j)} & m_4^{(j)} \end{pmatrix}.$$

Then, in the scalar-valued probability space  $(\mathcal{M}_2(\mathcal{A}), \text{tr}_2 \otimes \varphi)$ , the random variables  $s_1, \dots, s_u, \{m_j, j = 1, \dots, t\}$  are free and for  $i = 1, \dots, u$ , each  $s_i$  is a semicircular variable.

*Proof.* Let us prove that  $s_1, \dots, s_u$  is free from  $\mathcal{M}_2(\mathcal{B})$ , where  $\mathcal{B}$  is the  $*$ -algebra generated by  $\{m_p^{(j)}, p = 1, \dots, 4, j = 1, \dots, t\}$ . We already now (see [26, Chapter 9]) that  $s_1, \dots, s_u$  are semicircular variables over  $\mathcal{M}_2(\mathbb{C})$  which are free from  $\mathcal{M}_2(\mathcal{B})$ , with respect to  $\text{id}_2 \otimes \varphi$ . Moreover, the covariance mapping of  $s_1, \dots, s_u$  is the function  $(\eta_{i,j}^{\mathcal{M}_2(\mathbb{C})} : \mathcal{M}_2(\mathbb{C}) \rightarrow \mathcal{M}_2(\mathbb{C}))_{1 \leq i, j \leq u}$ , which can be computed as follows: for all  $m = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$ , we have

$$\begin{aligned} \eta_{i,j}^{\mathcal{M}_2(\mathbb{C})}(m) &= (\text{id}_2 \otimes \varphi)(s_i m s_j) \\ &= \frac{1}{2} \left( \frac{\varphi(s_i^{(1)} m_1 s_j^{(1)}) + \varphi(s_i^{(1)} m_2 c_j^*) + \varphi(c_i m_3 s_j^{(1)}) + \varphi(c_i m_4 c_j^*)}{\star} \middle| \star \right) \\ &= \frac{\delta_{ij}}{2} \left( \frac{m_1 + m_4}{0} \middle| \frac{0}{m_1 + m_4} \right) \\ &= \delta_{ij} \text{tr}_2(m) I_2. \end{aligned}$$

Using [27, Theorem 3.5], the freeness of  $s_1, \dots, s_u$  from  $\mathcal{M}_2(\mathcal{B})$  over  $\mathcal{M}_2(\mathbb{C})$  gives us the free cumulants of  $s_1, \dots, s_u$  over  $\mathcal{M}_2(\mathcal{B})$ . More concretely, we get that  $s_1, \dots, s_u$  are semicircular variables over  $\mathcal{M}_2(\mathcal{B})$ , with a covariance mapping  $(\eta_{i,j}^{\mathcal{M}_2(\mathcal{B})} : \mathcal{M}_2(\mathcal{B}) \rightarrow \mathcal{M}_2(\mathcal{B}))_{1 \leq i, j \leq u}$  given by  $\eta_{i,j}^{\mathcal{M}_2(\mathcal{B})} = \eta_{i,j}^{\mathcal{M}_2(\mathbb{C})} \circ (\text{id}_2 \otimes \varphi)$ .

Because of the previous computation, we know that  $\eta_{i,j}^{\mathcal{M}_2(\mathbb{C})} = \text{tr}_2 \circ \eta_{i,j}^{\mathcal{M}_2(\mathbb{C})} \circ \text{tr}_2$ , which means that  $\eta_{i,j}^{\mathcal{M}_2(\mathcal{B})} = (\text{tr}_2 \otimes \varphi) \circ \eta_{i,j}^{\mathcal{M}_2(\mathcal{B})} \circ (\text{tr}_2 \otimes \varphi)$ . As a consequence, using again [27, Theorem 3.5],  $s_1, \dots, s_u$  are semicircular variables over  $\mathbb{C}$  free from  $\mathcal{M}_2(\mathcal{B})$  with respect to  $(\text{tr}_2 \otimes \varphi)$ , and the covariance mapping  $\eta_{i,j}^{\mathbb{C}}$  is given by the restriction of the covariance mapping  $\eta^{\mathcal{M}_2(\mathbb{C})}$  to  $\mathbb{C}$ : for all  $m \in \mathbb{C}$

$$\eta_{i,j}^{\mathbb{C}}(m) = \delta_{ij} m,$$

which means that  $s_1, \dots, s_u$  are free standard semicircular variables. □

Our next lemma is an operator-valued extension of [13, Lemma 5.1].

**Lemma 3.4.** *Let  $y$  be a noncommutative random variable in  $\mathcal{M}_m(\mathcal{A})$  and  $c^{(1)}, \dots, c^{(u)}$  be free circular variables in  $\mathcal{A}$ , free from the entries of  $y$ . Then, in the non-commutative probability space  $(\mathcal{M}_m(\mathcal{A}), \text{tr}_m \otimes \varphi)$ ,  $|\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y|^2$  has the same distribution as  $|\sum_{j=1}^u \zeta_j \otimes s_j + (I_m \otimes \epsilon) \cdot y|^2$  for any scalar matrices  $\zeta_1, \dots, \zeta_u \in \mathcal{M}_m(\mathbb{C})$ , where  $\epsilon$  is a selfadjoint  $\{-1, +1\}$ -Bernoulli variable in  $\mathcal{A}$ , independent from the entries of  $y$ , and  $s_1, \dots, s_u$  are free semicircular variables in  $\mathcal{A}$ , free from  $\epsilon$  and the entries of  $y$ .*

In the lemma above, we consider the symmetric version  $\epsilon y$  of  $y$ , thanks to a non-commutative random variable  $\epsilon$  which is *tensor-independent* from the entries of  $y$ , in the sense that  $\epsilon$  commutes with the entries of  $y$  and  $\varphi(p_1(\epsilon)p_2(y_{i,j}, y_{i,j}^* : i, j)) = \varphi(p_1(\epsilon))\varphi(p_2(y_{i,j}, y_{i,j}^* : i, j))$  for all polynomials  $p_1, p_2$ .

*Proof.* Let  $n \geq 0$ . We compute the  $n$ -th moment of  $|\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y|^2$  with respect to  $id_m \otimes \varphi$ , and compare it to the  $n$ -th moment of  $|\sum_{j=1}^u \zeta_j \otimes s_j + \epsilon y|^2$  with respect to  $id_m \otimes \varphi$ .

Let us set  $a_0 = y$  and  $a_j = \zeta_j \otimes c^{(j)}$ . We compute

$$\begin{aligned} & (id_m \otimes \varphi)(|\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y|^{2n}) \\ &= \sum_{0 \leq i_1, \dots, i_{2n} \leq u} (id_m \otimes \varphi)(a_{i_1} a_{i_2}^* a_{i_3} a_{i_4}^* \dots a_{i_{2n-1}} a_{i_{2n}}^*). \end{aligned}$$

Similarly,

$$\begin{aligned} & (id_m \otimes \varphi)(|\sum_{j=1}^u \zeta_j \otimes s_j + (I_m \otimes \epsilon) \cdot y|^{2n}) \\ &= \sum_{0 \leq i_1, \dots, i_{2n} \leq u} (id_m \otimes \varphi)(b_{i_1} b_{i_2}^* b_{i_3} b_{i_4}^* \dots b_{i_{2n-1}} b_{i_{2n}}^*). \end{aligned}$$

where  $b_0 = (I_m \otimes \epsilon) \cdot y$  and  $b_j = \zeta_j \otimes s_j$ . In order to conclude, it suffices to prove that, for all  $0 \leq i_1, \dots, i_{2n} \leq u$ ,

$$(id_m \otimes \varphi)(a_{i_1} a_{i_2}^* a_{i_3} a_{i_4}^* \dots a_{i_{2n-1}} a_{i_{2n}}^*) = (id_m \otimes \varphi)(b_{i_1} b_{i_2}^* b_{i_3} b_{i_4}^* \dots b_{i_{2n-1}} b_{i_{2n}}^*).$$

Let us fix  $0 \leq i_1, \dots, i_{2n} \leq u$ . Note that  $a_0$  is free over  $\mathcal{M}_m(\mathbb{C})$  from  $a_j$  with respect to  $id_m \otimes \varphi$  (see [26, Chapter 9]). Let us fix  $S = \{j : i_j \neq 0\} \subset \{1, \dots, 2n\}$  and use the moment cumulant formula (see [34, page 36]):

$$\begin{aligned} & (id_m \otimes \varphi)(a_{i_1} a_{i_2}^* a_{i_3} a_{i_4}^* \dots a_{i_{2n-1}} a_{i_{2n}}^*) \\ &= \sum_{\pi \in NC(S)} (\hat{c} \cup \hat{\varphi})(\pi \cup \pi^c)(a_{i_1} \otimes a_{i_2}^* \dots a_{i_{2n-1}} \otimes a_{i_{2n}}^*) \end{aligned}$$

where  $\pi^c$  is the largest partition of  $S^c$  such that  $\pi \cup \pi^c$  is noncrossing and  $\hat{c}$  and  $\hat{\varphi}$  are the  $\mathcal{M}_m(\mathbb{C})$ -valued cumulant function and the  $\mathcal{M}_m(\mathbb{C})$ -valued moment function associated to the conditional expectation  $id_m \otimes \varphi$ . We use here the notation of [34, Notation 2.1.4] which defines  $(\hat{c} \cup \hat{\varphi})(\pi \cup \pi^c)$  as some  $\mathcal{M}_m(\mathbb{C})$ -valued multiplicative function that acts on the blocks of  $\pi$  like  $\hat{c}$  and on the blocks of  $\pi^c$  like  $\hat{\varphi}$ .

Recall that the cumulants of  $\zeta_j \otimes c^{(j)}$  are vanishing if  $\pi$  is not a pairing and if  $\pi$  is not alternating (which means that  $\pi$  links two indices with the same parity). Now, let us remark that if  $\pi$  is a pairing which is alternating, then  $\pi^c$  is even (each blocs of  $\pi^c$  is

even). Thus,

$$\begin{aligned} & (id_m \otimes \varphi)(a_{i_1} a_{i_2}^* a_{i_3} a_{i_4}^* \dots a_{i_{2n-1}} a_{i_{2n}}^*) \\ &= \sum_{\substack{\pi \in NC(S) \\ \pi \text{ pairing and alternating}}} (\hat{c} \cup \hat{\varphi})(\pi \cup \pi^c)(a_{i_1} \otimes a_{i_2}^* \dots a_{i_{2n-1}} \otimes a_{i_{2n}}^*) \\ &= \sum_{\substack{\pi \in NC(S) \\ \pi \text{ pairing and alternating} \\ \pi^c \text{ even}}} (\hat{c} \cup \hat{\varphi})(\pi \cup \pi^c)(a_{i_1} \otimes a_{i_2}^* \dots a_{i_{2n-1}} \otimes a_{i_{2n}}^*). \end{aligned}$$

Similarly, the cumulants of  $\zeta_j \otimes s^{(j)}$  are vanishing if  $\pi$  is not a pairing and that the moment of  $b_0$  is vanishing if  $\pi^c$  is odd. Moreover, if  $\pi$  is a pairing and  $\pi^c$  is even, then  $\pi$  is alternating. As a consequence,

$$\begin{aligned} & (id_m \otimes \varphi)(b_{i_1} b_{i_2}^* b_{i_3} b_{i_4}^* \dots b_{i_{2n-1}} b_{i_{2n}}^*) \\ &= \sum_{\pi \in NC(S)} (\hat{c} \cup \hat{\varphi})(\pi \cup \pi^c)(b_{i_1} \otimes b_{i_2}^* \dots b_{i_{2n-1}} \otimes b_{i_{2n}}^*) \\ &= \sum_{\substack{\pi \in NC(S) \\ \pi \text{ pairing} \\ \pi^c \text{ even}}} (\hat{c} \cup \hat{\varphi})(\pi \cup \pi^c)(b_{i_1} \otimes b_{i_2}^* \dots b_{i_{2n-1}} \otimes b_{i_{2n}}^*) \\ &= \sum_{\substack{\pi \in NC(S) \\ \pi \text{ pairing and alternating} \\ \pi^c \text{ even}}} (\hat{c} \cup \hat{\varphi})(\pi \cup \pi^c)(b_{i_1} \otimes b_{i_2}^* \dots b_{i_{2n-1}} \otimes b_{i_{2n}}^*). \end{aligned}$$

In order to conclude, it suffices to remark that  $\epsilon_y$  and  $y$  have the same even  $\mathcal{M}_m(\mathbb{C})$ -valued moments and  $\zeta_j \otimes c^{(j)}$  and  $\zeta_j \otimes s^{(j)}$  have the same alternating  $\mathcal{M}_m(\mathbb{C})$ -valued cumulants.  $\square$

It follows from [4] that the support in  $\mathcal{M}_m^{\text{sa}}(\mathbb{C})$  of the addition of a semicircular  $s$  of variance  $\eta$  and a selfadjoint noncommutative random variable  $y \in (\mathcal{M}_m(\mathcal{A}), id_m \otimes \varphi)$  which is free with amalgamation over  $\mathcal{M}_m(\mathbb{C})$  with  $s$ , is given via its complement in terms of  $y$  and the functions

$$H(w) = w - \eta(G_y(w)) \text{ and } \omega(b) = b + \eta(G_y(\omega(b))), \tag{3.2}$$

where  $G_x(b) = (id_m \otimes \varphi) [(x - b)^{-1}]$ . Specifically,

**Proposition 3.5.** *If  $w \in \mathcal{M}_m^{\text{sa}}(\mathbb{C})$  is such that  $y - w$  is invertible and  $\text{spect}(\eta \circ G'_y(w)) \subset \overline{\mathbb{D}} \setminus \{1\}$ , then  $s + y - H(w)$  is invertible. Conversely, if  $b \in \mathcal{M}_m^{\text{sa}}(\mathbb{C})$  is such that  $s + y - b$  is invertible, then  $y - \omega(b)$  is invertible.*

It follows quite easily that  $\text{spect}(\eta \circ G'_y(\omega(b))) \subset \overline{\mathbb{D}}$ . Generally, all conditions on the derivatives of  $\omega$  and  $H$  follow from the two functional equations above.

*Proof.* Assume that  $y - w$  is invertible and  $\text{spect}(\eta \circ G'_y(w)) \subset \overline{\mathbb{D}} \setminus \{1\}$ . Since  $w = w^*$ , the derivative  $G'_y(w)$  is completely positive, so  $\eta \circ G'_y(w)$  is completely positive. This means according to [18, Theorem 2.5] that the spectral radius  $r$  of  $\eta \circ G'_y(w)$  is reached at a positive element  $\xi \in \mathcal{M}_m(\mathbb{C})$ , so that necessarily  $r \geq 0$ . Since  $1 \notin \sigma(\eta \circ G'_y(w))$  by hypothesis, it follows that  $r < 1$ , and thus

$$\text{spect}(\eta \circ G'_y(w)) \subseteq r\overline{\mathbb{D}} \subsetneq \mathbb{D}.$$

This forces the derivative of  $H(w)$ ,  $H'(w) = id_m - \eta \circ G'_y(w)$ , to be invertible as a linear operator from  $\mathcal{M}_m(\mathbb{C})$  to itself. By the inverse function theorem,  $H$  has an analytic



inverse on a small enough neighborhood of  $H(w)$  onto a neighborhood of  $w$ . Since  $H$  preserves the selfadjoints near  $w$ , so must the inverse. On the other hand, the map  $v \mapsto H(w) + \eta(G_y(v))$  sends the upper half-plane into itself and has  $w$  as a fixed point. Since its derivative has all its eigenvalues included *strictly* in  $\mathbb{D}$  (recall that the spectral radius  $r < 1$ ), it follows that  $w$  is actually an attracting fixed point for this map. Since for any  $b$  in the upper half-plane,  $\omega(b)$  is given as the attracting fixed point of  $v \mapsto b + \eta(G_y(v))$ , it follows that  $\omega$  coincides with the local inverse of  $H$  on the upper half-plane, so the local inverse of  $H$  is the unique analytic continuation of  $\omega$  to a neighborhood of  $H(w)$ . This proves that  $\omega$  extends analytically to a neighborhood of  $H(w)$  and the extension maps selfadjoints from this neighborhood to  $\mathcal{M}_m^{\text{sa}}(\mathbb{C})$ . In particular,  $\omega(H(v)) = v$  and  $G_{s+y}(H(v)) = G_y(\omega(H(v))) = G_y(v)$  are selfadjoint for all  $v = v^*$  in a small enough neighborhood of  $w$ , showing that  $s + y - H(w)$  is invertible.

Conversely, say  $b = b^*$  and  $s + y - b$  is invertible. Then  $G_{s+y}$  is analytic on a neighborhood of  $b$  and maps selfadjoints from this neighborhood into  $\mathcal{M}_m^{\text{sa}}(\mathbb{C})$ . Since  $\omega(b) = b + \eta(G_{s+y}(b))$ , the same holds for  $\omega$ . Since, by [4, Proposition 4.1],  $\text{spect}(\omega'(v)) \subset \{\Re z > 1/2\}$  for any  $v$  in the upper half-plane, the analyticity of  $\omega$  around  $b = b^*$  implies  $\text{spect}(\omega'(b)) \subset \{\Re z \geq 1/2\}$ . Thus,  $\omega$  is invertible wrt composition around  $b$  by the inverse function theorem. As argued above,  $H$  is its inverse, and extends analytically to a small enough neighborhood of  $\omega(b)$ , with selfadjoint values on the selfadjoints. Composing with  $H$  to the left in Voiculescu's subordination relation  $G_{s+y}(v) = G_y(\omega(v))$  yields  $G_{y+s}(H(w)) = G_y(w)$ , guaranteeing that  $G_y$  is analytic on a neighborhood of  $\omega(b)$ , with selfadjoint values on the selfadjoints, and so  $y - \omega(b)$  must be invertible.  $\square$

**Remark 3.6.** The proof of the previous proposition, based on [18, Theorem 2.5], makes the condition  $\text{spect}(\eta \circ G'_y(0)) \subseteq \overline{\mathbb{D}} \setminus \{1\}$  equivalent to the existence of an  $r \in [0, 1)$  such that  $\text{spect}(\eta \circ G'_y(0)) \subseteq r\overline{\mathbb{D}}$ .

The following lemma is a particular case of the above proposition.

**Lemma 3.7.** Consider the operator-valued  $C^*$ -algebraic noncommutative probability space  $(\mathcal{M}_m(\mathcal{A}), id_m \otimes \varphi, \mathcal{M}_m(\mathbb{C}))$  and a pair of selfadjoint random variables  $s, y \in \mathcal{M}_m(\mathcal{A})$  which are free over  $\mathcal{M}_m(\mathbb{C})$  with respect to  $id_m \otimes \varphi$ . Assume that  $s$  is a centered semicircular of variance  $\eta: \mathcal{M}_m(\mathbb{C}) \rightarrow \mathcal{M}_m(\mathbb{C})$  and that each entry of  $y \in \mathcal{M}_m(\mathcal{A})$  is a noncommutative symmetric random variable in  $(\mathcal{A}, \varphi)$ . We define  $G_x(b) = (id_m \otimes \varphi)[(x - b)^{-1}]$ . Then  $s + y$  is invertible if and only if  $0 \notin \text{spect}(y)$  and  $\text{spect}(\eta \circ G'_y(0))$  is included in  $\overline{\mathbb{D}} \setminus \{1\}$ .

*Proof.* Note that our hypotheses that all entries of the selfadjoint  $y$  are symmetric and that  $s$  is centered imply automatically that  $H(i\mathcal{M}_m(\mathbb{C})^+) \subseteq i\mathcal{M}_m^{\text{sa}}(\mathbb{C})$  and  $\omega(i\mathcal{M}_m(\mathbb{C})^+) \cup G_y(i\mathcal{M}_m(\mathbb{C})^+) \cup G_{y+s}(i\mathcal{M}_m(\mathbb{C})^+) \subseteq i\mathcal{M}_m(\mathbb{C})^+$ . (We have denoted by  $\mathcal{M}_m(\mathbb{C})^+$  the set of positive definite matrices in  $\mathcal{M}_m(\mathbb{C})$ .)

Assume that  $y$  is invertible and  $\text{spect}(\eta \circ G'_y(0)) \subseteq \overline{\mathbb{D}} \setminus \{1\}$ . In particular,  $G_y$  is analytic on a neighborhood of zero in  $\mathcal{M}_m(\mathbb{C})$ . Proposition 3.5 implies that  $s + y - H(0)$  is invertible. Since  $H(i\mathcal{M}_m(\mathbb{C})^+) \subseteq i\mathcal{M}_m(\mathbb{C})^+$ , it follows from the formula of  $H$  that  $H(0) = 0$ . Thus,  $s + y$  is invertible.

Conversely, assume that  $s + y$  is invertible, so that  $G_{s+y}$  extends analytically to a small neighborhood of zero in such a way that it maps selfadjoints to selfadjoints. Since  $\omega(b) = b + \eta(G_{s+y}(b))$ , it follows that  $\omega$  does the same. According to Proposition 3.5,  $y - \omega(0)$  is invertible. Since  $\omega(i\mathcal{M}_m(\mathbb{C})^+) \subseteq i\mathcal{M}_m(\mathbb{C})^+$ , we again have that  $\omega(0) = 0$ , so that  $y$  is invertible.  $\square$

### 4 Linearization trick

A powerful tool to deal with noncommutative polynomials in random matrices or in operators is the so-called “linearization trick.” Its origins can be found in the theory of automata and formal languages (see, for instance, [32]), where it was used to conveniently represent certain categories of formal power series. In the context of operator algebras and random matrices, this procedure goes back to Haagerup and Thorbjørnsen [21, 22] (see [26]). We use the version from [1, Proposition 3], which has several advantages for our purposes, to be described below.

We denote by  $\mathbb{C}\langle X_1, \dots, X_k \rangle$  the complex  $*$ -algebra of polynomials in  $k$  noncommuting indeterminates  $X_1, \dots, X_k$ . The adjoint operation is given by the anti-linear extension of  $(X_{i_1} X_{i_2} \cdots X_{i_l})^* = X_{i_l}^* \cdots X_{i_2}^* X_{i_1}^*$ ,  $(i_1, \dots, i_l) \in \{1, \dots, k\}^l, l \in \mathbb{N} \setminus \{0\}$ . We will sometimes assume that some, or all, of the indeterminates are selfadjoint, i.e.  $X_j^* = X_j$ . Unless we make this assumption explicitly, the adjoints  $X_1^*, \dots, X_k^*$  are assumed to be algebraically free from each other and from  $X_1, \dots, X_k$ .

Given a polynomial  $P \in \mathbb{C}\langle X_1, \dots, X_k \rangle$ , we call *linearization* of  $P$  any  $L_P \in \mathcal{M}_m(\mathbb{C}) \otimes \mathbb{C}\langle X_1, \dots, X_k \rangle$  such that

$$L_P := \begin{pmatrix} 0 & u^* \\ v & Q \end{pmatrix} \in \mathcal{M}_m(\mathbb{C}) \otimes \mathbb{C}\langle X_1, \dots, X_k \rangle$$

where

1.  $m \in \mathbb{N}$ ,
2.  $Q \in \mathcal{M}_{m-1}(\mathbb{C}) \otimes \mathbb{C}\langle X_1, \dots, X_k \rangle$  is invertible in the complex algebra  $\mathcal{M}_{m-1}(\mathbb{C}) \otimes \mathbb{C}\langle X_1, \dots, X_k \rangle$ ,
3.  $u^*$  is a row vector and  $v$  is a column vector, both of length  $m - 1$ , with entries in  $\mathbb{C}\langle X_1, \dots, X_k \rangle$ ,
4. the polynomial entries in  $Q, u$  and  $v$  all have degree  $\leq 1$ , and
5. 
$$P = -u^* Q^{-1} v.$$

We refer to Anderson’s paper [1] for the – constructive – proof of the existence of a linearization  $L_P$  as described above for any given polynomial  $P \in \mathbb{C}\langle X_1, \dots, X_k \rangle$ . It turns out that if  $P$  is selfadjoint, then  $L_P$  can be chosen to be self-adjoint. The well-known result about Schur complements yields then the following invertibility equivalence.

**Lemma 4.1.** [26, Chapter 10, Corollary 3] *Let  $P \in \mathbb{C}\langle X_1, \dots, X_k \rangle$  and let  $L_P \in \mathcal{M}_m(\mathbb{C}\langle X_1, \dots, X_k \rangle)$  be a linearization of  $P$  with the properties outlined above. Let  $e_{11}$  be the  $m \times m$  matrix whose single nonzero entry equals one and occurs in the row 1 and column 1. Let  $y = (y_1, \dots, y_k)$  be a  $k$ -tuple of operators in a unital  $C^*$ -algebra  $\mathcal{A}$ . Then, for any  $z \in \mathbb{C}$ ,  $ze_{11} \otimes 1_{\mathcal{A}} - L_P(y)$  is invertible if and only if  $z1_{\mathcal{A}} - P(y)$  is invertible and we have*

$$(ze_{11} \otimes 1_{\mathcal{A}} - L_P(y))^{-1} = \begin{pmatrix} (z1_{\mathcal{A}} - P(y))^{-1} & \star \\ \star & \star \end{pmatrix}. \tag{4.1}$$

**Lemma 4.2.** *Let  $P \in \mathbb{C}\langle X_1, \dots, X_k \rangle$  and let  $L_P \in \mathcal{M}_m(\mathbb{C}\langle X_1, \dots, X_k \rangle)$  be a linearization of  $P$  with the properties outlined above. There exist two polynomials  $T_1$  and  $T_2$  in  $k$  commutative indeterminates, with nonnegative coefficients, depending only on  $L_P$ , such that, for any  $k$ -tuple  $y = (y_1, \dots, y_k)$  of operators in a unital  $C^*$ -algebra  $\mathcal{A}$ , for any  $z \in \mathbb{C}$  such that  $z1_{\mathcal{A}} - P(y)$  is invertible,*

$$\begin{aligned} \|(ze_{11} \otimes 1_{\mathcal{A}} - L_P(y))^{-1}\| &\leq T_1(\|y_1\|, \dots, \|y_k\|) \times \|(z1_{\mathcal{A}} - P(y))^{-1}\| \\ &\quad + T_2(\|y_1\|, \dots, \|y_k\|). \end{aligned} \tag{4.2}$$

*Proof.* The proof is similar to the proof of [5, Lemma 4.4]. The linearization of  $P$  can be written as

$$L_P = \begin{bmatrix} 0 & u^* \\ v & Q \end{bmatrix} \in \mathcal{M}_m(\mathbb{C}\langle X_1, \dots, X_k \rangle)$$

Now, a matrix calculation in which we suppress the variable  $y$  shows that

$$(ze_{11} \otimes 1_{\mathcal{A}} - L_P)^{-1} = \begin{bmatrix} 1_{\mathcal{A}} & 0 \\ -Q^{-1}v & I_{(m-1)} \otimes 1_{\mathcal{A}} \end{bmatrix} \begin{bmatrix} (z - P)^{-1} & 0 \\ 0 & -Q^{-1} \end{bmatrix} \begin{bmatrix} 1_{\mathcal{A}} & -u^*Q^{-1} \\ 0 & I_{(m-1)} \otimes 1_{\mathcal{A}} \end{bmatrix}.$$

Since  $v$ ,  $u^*$ , and  $Q^{-1}$  are polynomials in  $y_1, \dots, y_k$ , the result readily follows. □

In Section 5.3, we will provide an explicit construction of a linearization that is best adapted to our purposes. In this construction, it is clear that we can always find a linearization such that, for any  $k$ -tuple  $y$  of matrices,

$$\det Q(y) = \pm 1. \tag{4.3}$$

### 5 No outlier; proof of Theorem 1.13

By Bai-Yin’s theorem (see [3, Theorem 5.8]), there exists  $C > 0$  such that, almost surely for all large  $N$ ,  $\|M_N\| \leq C$ , so that the first assertion of Theorem 1.13 readily yields the second one, by choosing

$$\Gamma = \{z \in \mathbb{C}, d(z, \text{spect}(P(c, a))) \geq \epsilon, |z| \leq C\}.$$

Remember that, by (1.4),  $\text{spect}(P(c, a)) = \{z \in \mathbb{C} : 0 \in \text{supp}(\mu_z)\}$ , where  $\mu_z$  is the distribution of  $(P(c, a) - z)(P(c, a) - z)^*$ . The first assertion of Theorem 1.13 is equivalent to the following.

**Proposition 5.1.** *Let  $\Gamma$  be a compact set of  $\{z \in \mathbb{C} : 0 \notin \text{supp}(\mu_z)\}$ ; assume that for any  $z$  in  $\Gamma$ , there exists  $\eta_z > 0$  such that for all  $N$  large enough,*

$$s_N \left( P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}) - zI_N \right) > \eta_z.$$

*Then, for any  $z$  in  $\Gamma$ , there exists  $\gamma_z > 0$ , such that almost surely, for all large  $N$ ,  $s_N(M_N - zI_N) \geq \gamma_z$ . Consequently, there exists  $\gamma_\Gamma > 0$  such that almost surely, for all large  $N$ ,  $\inf_{z \in \Gamma} s_N(M_N - zI_N) \geq \gamma_\Gamma$ .*

#### 5.1 Ideas of the proof

The proof of Proposition 5.1 is based on the two following key results.

**Proposition 5.2.** *Assume that (X1) holds. Let  $K$  be a polynomial in  $u+t$  noncommutative variables. Define*

$$K_N = K \left( \frac{X_N^{(1)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)} \right).$$

- Assume that (1.3) holds. Let  $\{a_N^{(j)}, j = 1, \dots, t\}$  be a set of noncommutative random variables in  $(\mathcal{A}, \varphi)$  which is free from a free circular system  $c = (c^{(1)}, \dots, c^{(u)})$  in  $(\mathcal{A}, \varphi)$  and such that the  $*$ -distribution of  $(A_N^{(j)}, j = 1, \dots, t)$  in the noncommutative probability space  $(\mathcal{M}_N(\mathbb{C}), \frac{1}{N} \text{Tr}_N)$  coincides with the  $*$ -distribution of  $a_N = (a_N^{(j)}, j = 1, \dots, t)$  in  $(\mathcal{A}, \varphi)$ . Let  $\tau_N$  be the the distribution of

$$K(c, a_N) [K(c, a_N)]^*$$

with respect to  $\varphi$ . If the interval  $[x, y]$ ,  $x < y$ , is such that there exists  $\delta > 0$  such that for all large  $N$ ,  $(x - \delta, y + \delta) \subset \mathbb{R} \setminus \text{supp}(\tau_N)$ , then, we have

$$\mathbb{P}[\text{for all large } N, \text{spect}(K_N K_N^*) \subset \mathbb{R} \setminus [x, y]] = 1.$$

- Assume that **(A1)** holds. Then, almost surely, the sequence of  $u + t$ -tuples  $\left(\frac{X_N^{(1)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)}\right)_{N \geq 1}$  converges in  $*$ -distribution towards  $(c, a)$  where  $c = (c_1, \dots, c_u)$  is a free circular system which is free with  $a = (a^{(1)}, \dots, a^{(t)})$  in  $(\mathcal{A}, \varphi)$ .

**Proposition 5.3.** Consider a polynomial  $P(Y_1, Y_2)$ , where  $Y_1$  is a tuple of noncommuting nonselfadjoint indeterminates,  $Y_2$  is a tuple of selfadjoint indeterminates, and no selfadjointness is assumed for  $P$ . We evaluate  $P$  in  $(c, a)$  and  $(c, a_N)$ , where  $c$  is a tuple of free circulars, which is  $*$ -free from the tuples  $a$  and  $a_N$ . We assume that  $a_N \rightarrow a$  in moments and that there exists a number  $\tau > 0$  such that  $\sup_N \|a_N\| \leq \tau$ .

1. We fix  $z_0 \in \mathbb{C}$  such that  $|P(c, a) - z_0|^2 \geq \delta_{z_0} > 0$  for a fixed  $\delta_{z_0}$ .
2. We assume that there exists  $N_{\delta_{z_0}} \in \mathbb{N}$  such that if  $N \geq N_{\delta_{z_0}}$ , then  $|P(c, a_N) - z_0|^2 \geq \delta_{z_0}$ .

Then, there exists  $\epsilon_{z_0} > 0$  for which there exists an  $N_{\epsilon_{z_0}} \in \mathbb{N}$  such that if  $N \geq N_{\epsilon_{z_0}}$ , then  $|P(c, a_N) - z_0|^2 \geq \epsilon_{z_0}$ .

**Remark 5.4.** Of course Proposition 5.3 still holds if the tuples  $a_N$  are nonselfadjoint, by considering instead the selfadjoint tuples  $(\Im(a_N), \Re(a_N))$ .

Let us explain how to deduce Theorem 1.10 from Proposition 5.2 and Proposition 5.3. Define  $\mu_{N,z}$  as the distribution of

$$\left[ P(c^{(1)}, \dots, c^{(u)}, a_N^{(1)}, \dots, a_N^{(t)}) - z \mathbf{1} \right] \times \left[ P(c^{(1)}, \dots, c^{(u)}, a_N^{(1)}, \dots, a_N^{(t)}) - z \mathbf{1} \right]^*$$

where  $\{c^{(1)}, (c^{(1)})^*\}, \dots, \{c^{(u)}, (c^{(u)})^*\}, \{a_N^{(1)}, \dots, a_N^{(t)}\}$  are free sets of noncommutative random variables and the  $*$ -distribution of  $(a_N^{(1)}, \dots, a_N^{(t)})$  in  $(\mathcal{A}, \varphi)$  coincide with the  $*$ -distribution of  $(A_N^{(1)}, \dots, A_N^{(t)})$  in  $(\mathcal{M}_N(\mathbb{C}), \text{tr}_N)$ .  $\mu_{N,z}$  is the so-called deterministic equivalent measure of the empirical spectral measure of  $(M_N - zI_N)(M_N - zI_N)^*$ . The following is a straightforward consequence of Proposition 5.3.

**Corollary 5.5.** Let  $z \in \mathbb{C}$  be such that  $0 \notin \text{supp}(\mu_z)$ . Assume that there exists  $\eta_z > 0$  such that for all  $N$  large enough, there is no singular value of

$$P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}) - zI_N$$

in  $[0, \eta_z]$ . Then, there exists  $\epsilon_z > 0$ , such that, for all large  $N$ ,

$$[0, \epsilon_z] \subset \mathbb{R} \setminus \text{supp}(\mu_{N,z}).$$

Then, we can deduce from Corollary 5.5 and Proposition 5.2 that there exists some  $\gamma_z > 0$  such that almost surely for all large  $N$ , there is no singular value of  $M_N - zI_N$  in  $[0, \gamma_z]$ . By a compactness argument and the fact that  $z \mapsto s_N(M_N - zI_N)$  is 1-Lipschitz, it readily follows that for any compact  $\Gamma \subset \{z : 0 \notin \text{supp}(\mu_z)\}$ , there exists some  $\gamma_\Gamma > 0$  such that almost surely for all large  $N$ ,

$$\inf_{z \in \Gamma} s_N(M_N - zI_N) \geq \gamma_\Gamma, \tag{5.1}$$

leading to Proposition 5.1.

**5.2 Proof of Proposition 5.2**

Note that

$$\begin{pmatrix} K_N K_N^* & 0 \\ 0 & K_N^* K_N \end{pmatrix} = \begin{pmatrix} 0 & K_N \\ K_N^* & 0 \end{pmatrix}^2,$$

so that the spectrum of  $K_N K_N^*$  coincides with the spectrum of  $\begin{pmatrix} 0 & K_N \\ K_N^* & 0 \end{pmatrix}^2$ . Now

$$\begin{aligned} \begin{pmatrix} 0 & K \\ K^* & 0 \end{pmatrix} &= \sum_{i=1}^p \begin{pmatrix} 0 & b_i m_i \\ \bar{b}_i m_i^* & 0 \end{pmatrix} \\ &= \sum_{i=1}^p b_i \begin{pmatrix} 0 & m_i \\ m_i^* & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \bar{b}_i \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & m_i \\ m_i^* & 0 \end{pmatrix} \end{aligned} \tag{5.2}$$

where the  $m_i$ 's are monomials and the  $b_i$ 's are complex numbers. Define  $Q_1 = \begin{pmatrix} I_N & 0 \\ 0 & 0 \end{pmatrix}$ ,

$Q_2 = \begin{pmatrix} 0 & 0 \\ 0 & I_N \end{pmatrix}$  and  $R = \begin{pmatrix} 0 & I_N \\ 0 & 0 \end{pmatrix}$ ,  $S = \begin{pmatrix} 0 & I_N \\ I_N & 0 \end{pmatrix}$ . Note that

$$\begin{pmatrix} 0 & \frac{X_N^{(i)}}{\sqrt{2N}} \\ 0 & 0 \end{pmatrix} = \sqrt{2} Q_1 \frac{\mathcal{W}^{(i)}}{\sqrt{2N}} Q_2$$

where the  $\mathcal{W}^{(i)}$ 's,  $i = 1, \dots, u$ , are  $2N \times 2N$  independent so called Wigner matrices satisfying assumptions of [6]. Now, note that as noticed by [7] for any monomial  $x_1 \cdots x_k$ ,

$$\begin{pmatrix} 0 & x_1 \cdots x_k \\ (x_1 \cdots x_k)^* & 0 \end{pmatrix} = \Pi_{k-1} \begin{pmatrix} 0 & x_k \\ x_k^* & 0 \end{pmatrix} \Pi_{k-1}^* \tag{5.3}$$

where

$$\Pi_{k-1} = \begin{pmatrix} 0 & x_1 \\ I_N & 0 \end{pmatrix} S \begin{pmatrix} 0 & x_2 \\ I_N & 0 \end{pmatrix} S \cdots S \begin{pmatrix} 0 & x_{k-1} \\ I_N & 0 \end{pmatrix} S.$$

Indeed, this can be proved by induction noting that

$$\begin{pmatrix} 0 & x_1 \\ I_N & 0 \end{pmatrix} S \begin{pmatrix} 0 & x_2 \\ x_2^* & 0 \end{pmatrix} S \begin{pmatrix} 0 & I_N \\ x_1^* & 0 \end{pmatrix} = \begin{pmatrix} 0 & x_1 x_2 \\ x_2^* x_1^* & 0 \end{pmatrix}.$$

Note also that

$$S \begin{pmatrix} 0 & I_N \\ x_1^* & 0 \end{pmatrix} S = \begin{pmatrix} 0 & x_1^* \\ I_N & 0 \end{pmatrix}. \tag{5.4}$$

Set for  $j=1, \dots, t$ ,  $\mathbf{A}_N^{(j)} = \begin{pmatrix} 0 & A_N^{(j)} \\ 0 & 0 \end{pmatrix}$ .

From (5.2), (5.3), (5.4), it readily follows that there exists a polynomial  $\hat{K}$  such that  $\begin{pmatrix} 0 & K_N \\ K_N^* & 0 \end{pmatrix}$  is equal to

$$\hat{K} \left( Q_1, Q_2, R, R^*, \mathbf{A}_N^{(j)}, (\mathbf{A}_N^{(j)})^*, j = 1, \dots, t, \frac{\mathcal{W}^{(i)}}{\sqrt{2N}}, i = 1, \dots, u \right).$$

Now, define for  $j = 1, \dots, t$ ,  $\mathbf{a}_N^{(j)} = \begin{pmatrix} 0 & a_N^{(j)} \\ 0 & 0 \end{pmatrix}$ ,  $q_1 = \begin{pmatrix} 1_{\mathcal{A}} & 0 \\ 0 & 0 \end{pmatrix}$ ,  $q_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1_{\mathcal{A}} \end{pmatrix}$  and  $r =$

$\begin{pmatrix} 0 & 1_{\mathcal{A}} \\ 0 & 0 \end{pmatrix}$ . Let  $s_i^{(1)}, s_i^{(2)}, i = 1, \dots, u$  be semicircular variables such that  $\{s_1^{(1)}\}, \dots, \{s_u^{(1)}\}$ ,

$\{s_1^{(2)}, \dots, s_u^{(2)}\}, \{c_1, c_1^*\}, \dots, \{c_u, c_u^*\}, \{a_N^{(j)}, j = 1, \dots, t\}$  are free. Define for  $i = 1, \dots, u$ ,

$$s_i = \frac{1}{\sqrt{2}} \begin{pmatrix} s_i^{(1)} & c^{(i)} \\ (c^{(i)})^* & s_i^{(1)} \end{pmatrix}.$$

Similarly,

$$\begin{aligned} & \begin{pmatrix} 0 & K(c_1, \dots, c_u, a_N^{(1)}, \dots, a_N^{(t)}) \\ [K(c_1, \dots, c_u, a_N^{(1)}, \dots, a_N^{(t)})]^* & 0 \end{pmatrix} \\ &= \hat{K}(q_1, q_2, r, r^*, \mathbf{a}_N^{(j)}, (\mathbf{a}_N^{(j)})^*, j = 1, \dots, t, \mathbf{s}_i, i = 1, \dots, u). \end{aligned}$$

It readily follows that, the spectrum of  $K_N K_N^*$  coincides with the spectrum of  $\hat{K}(Q_1, Q_2, R, R^*, \mathbf{A}_N^{(j)}, (\mathbf{A}_N^{(j)})^*, j = 1, \dots, t, \frac{W^{(i)}}{\sqrt{2N}}, i = 1, \dots, u)^2$  and the spectrum of  $K(c_1, \dots, c_u, a_N^{(1)}, \dots, a_N^{(t)}) [K(c_1, \dots, c_u, a_N^{(1)}, \dots, a_N^{(t)})]^*$  coincides with the spectrum of  $\hat{K}(q_1, q_2, r, r^*, \mathbf{a}_N^{(j)}, (\mathbf{a}_N^{(j)})^*, j = 1, \dots, t, \mathbf{s}_i, i = 1, \dots, u)^2$ .

Now, it is straightforward to see that the  $*$ -distribution of  $(q_1, q_2, r, \mathbf{a}_N^{(j)}, j = 1, \dots, t)$  in  $(\mathcal{M}_2(\mathcal{A}), \text{tr}_2 \otimes \varphi)$  coincides with the  $*$ -distribution of  $(Q_1, Q_2, R, \mathbf{A}_N^{(j)}, j = 1, \dots, t)$  in  $(\mathcal{M}_{2N}(\mathbb{C}), \text{tr}_{2N})$ . Moreover, by Lemma 3.3, it turns out that the  $s_i$ 's are free semicircular variables which are free with  $(q_1, q_2, r, \mathbf{a}_N^{(j)}, j = 1, \dots, t)$  in  $(\mathcal{M}_2(\mathcal{A}), \text{tr}_2 \otimes \varphi)$ . Therefore, the first assertion of Proposition 5.2 follows by applying [6, Theorem 1.1.]. The second assertion of Proposition 5.2 can be proven by the same previous arguments. Indeed, there exists a polynomial  $\tilde{K}$  such that

$$\begin{aligned} & \frac{1}{N} \text{Tr} K \left( \frac{X_N^{(1)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)} \right) \\ &= \frac{1}{N} \text{Tr} \left\{ \begin{pmatrix} 0 & K_N \\ K_N^* & 0 \end{pmatrix} R^* \right\} \\ &= 2 \frac{1}{2N} \text{Tr} \tilde{K} \left( Q_1, Q_2, R, R^*, \mathbf{A}_N^{(j)}, (\mathbf{A}_N^{(j)})^*, j = 1, \dots, t, \frac{W^{(i)}}{\sqrt{2N}}, i = 1, \dots, u \right) \end{aligned}$$

Thus, using [6, Proposition 2.2. and Remark 4], we obtain that

$$\begin{aligned} & \frac{1}{N} \text{Tr} K \left( \frac{X_N^{(1)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)} \right) \\ & \xrightarrow{N \rightarrow +\infty} 2 \text{tr}_2 \otimes \varphi \left[ \tilde{K} \left( q_1, q_2, r, r^*, \mathbf{a}^{(j)}, (\mathbf{a}^{(j)})^*, j = 1, \dots, t, \mathbf{s}_i, i = 1, \dots, u \right) \right] \end{aligned}$$

where, for  $j = 1, \dots, t$ ,  $\mathbf{a}^{(j)} = \begin{pmatrix} 0 & a^{(j)} \\ 0 & 0 \end{pmatrix}$ . Now,

$$\begin{aligned} & 2 \text{tr}_2 \otimes \varphi \left[ \tilde{K} \left( q_1, q_2, r, r^*, \mathbf{a}^{(j)}, (\mathbf{a}^{(j)})^*, j = 1, \dots, t, \mathbf{s}_i, i = 1, \dots, u \right) \right] \\ &= 2 \text{tr}_2 \otimes \varphi \left\{ \begin{pmatrix} 0 & K(c, a) \\ K(c, a)^* & 0 \end{pmatrix} r^* \right\} \\ &= \varphi(K(c, a)). \end{aligned}$$

The second assertion of Proposition 5.2 follows.

**5.3 Proof of Proposition 5.3**

We prove this using linearization and hermitization. Our linearization of a nonselfadjoint polynomial will naturally not be selfadjoint, so the results from [5] do not apply directly to it, but some of the methods will. Before we analyze this linearization, let us lay down the steps that we shall take in order to prove the above result. Let  $L$  be our linearization of  $P(Y_1, Y_2) - z_0$ ; in this section we use a slight modification of the linearization introduced in [1] — see below.

1. We have  $|P(c, a_N) - z_0|^2 \geq \epsilon_{z_0} \iff \begin{bmatrix} 0 & P(c, a_N) - z_0 \\ (P(c, a_N) - z_0)^* & 0 \end{bmatrix}^2 \geq \epsilon_{z_0}$ .

2. There exists  $\iota = \iota(\epsilon_{z_0}, P, \tau) > 0$  such that

$$\left| \begin{bmatrix} 0 & P(c, a_N) - z_0 \\ (P(c, a_N) - z_0)^* & 0 \end{bmatrix} \right| \geq \epsilon_{z_0} \iff \left| \begin{bmatrix} 0 & L(c, a_N) \\ L(c, a_N)^* & 0 \end{bmatrix} \right| \geq \iota.$$

3. We write

$$\begin{bmatrix} 0 & L(c, a_N) \\ L(c, a_N)^* & 0 \end{bmatrix} = \begin{bmatrix} 0 & L(0, a_N) \\ L(0, a_N)^* & 0 \end{bmatrix} + C,$$

where  $C$  is a selfadjoint matrix containing only circular variables and their adjoints.

It will be clear that  $\begin{bmatrix} 0 & L(c, a_N) \\ L(c, a_N)^* & 0 \end{bmatrix}$  contains at most one nonzero element per row/column, except possibly for the first row/column.

4. We use Lemma 3.4 to conclude that the lhs of the previous item is invertible if and only if

$$\begin{bmatrix} 0 & (I_m \otimes \epsilon)L(0, a_N) \\ (I_m \otimes \epsilon)L(0, a_N)^* & 0 \end{bmatrix} + \mathcal{S}$$

is, where  $\mathcal{S}$  is obtained from  $C$  by replacing each circular entry with a semicircular from the same algebra (and hence free from  $a_N$ ), and  $\epsilon$  is a  $\{-1, 1\}$ -Bernoulli distributed random variable which is independent from  $a_N$  and free from  $\mathcal{S}$ . As noted in Example 3.1, since  $C = C^*$ ,  $\mathcal{S}$  is indeed a matrix-valued semicircular random variable.

5. We apply Lemma 3.7 to the above item in order to determine under what conditions the sum in question has a spectrum uniformly bounded away from zero.

6. Finally, we use the convergence in moments of  $a_N$  to  $a$  in order to conclude that the conditions obtained in the previous item are satisfied by

$$\begin{bmatrix} 0 & (I_m \otimes \epsilon)L(0, a_N) \\ (I_m \otimes \epsilon)L(0, a_N)^* & 0 \end{bmatrix} + \mathcal{S}.$$

Part (1) is trivial:

$$\begin{bmatrix} 0 & P(c, a_N) - z_0 \\ (P(c, a_N) - z_0)^* & 0 \end{bmatrix}^2 = \begin{bmatrix} |P(c, a_N) - z_0|^2 & 0 \\ 0 & |(P(c, a_N) - z_0)^*|^2 \end{bmatrix}.$$

Since our variables live in a  $\text{II}_1$ -factor, the two nonzero entries of the right hand side have the same spectrum.

Part (2) requires a careful analysis of the linearization we use. The construction from [1] proceeds by induction on the number of monomials in the given polynomial. If

$P = X_{i_1}X_{i_2}X_{i_3} \cdots X_{i_{\ell-1}}X_{i_\ell}$ , where  $\ell \geq 2$  and  $i_1, \dots, i_\ell \in \{1, \dots, k\}$ , we set  $n = \ell$  and

$$L = - \begin{bmatrix} 0 & 0 & \cdots & 0 & X_{i_1} \\ 0 & 0 & \cdots & X_{i_2} & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & X_{i_{\ell-1}} & \cdots & 0 & 0 \\ X_{i_\ell} & -1 & \cdots & 0 & 0 \end{bmatrix}.$$

However, unlike in [1, 5], we choose here  $L$  to be

$$L = - \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & \cdots & 0 & X_{i_1} & -1 \\ 0 & 0 & \cdots & X_{i_2} & -1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & X_{i_\ell} & \cdots & 0 & 0 & 0 \\ 1 & -1 & \cdots & 0 & 0 & 0 \end{bmatrix}.$$

That is, we apply the procedure from [1], but to  $P = 1X_{i_1}X_{i_2}X_{i_3} \cdots X_{i_{\ell-1}}X_{i_\ell}1$ . If  $\ell = 1$ , we simply complete  $X$  to  $1X1$ . Even if we have a multiple of 1, we choose here to proceed the same way. The lower right  $(\ell + 1) \times (\ell + 1)$  corner of this matrix has an inverse of degree  $\ell$  in the algebra  $\mathcal{M}_{\ell+1}(\mathbb{C}\langle X_1, \dots, X_k \rangle)$ . (The constant term in this inverse is a selfadjoint matrix and its spectrum is contained in  $\{-1, 1\}$ .) The first row contains only zeros and ones, and the first column is the transpose of the first row. Suppose now that  $p = P_1 + P_2$ , where  $P_1, P_2 \in \mathbb{C}\langle X_1, \dots, X_k \rangle$ , and that linear polynomials

$$L_j = \begin{bmatrix} 0 & u_j^* \\ u_j & Q_j \end{bmatrix} \in \mathcal{M}_{n_j}(\mathbb{C}\langle X_1, \dots, X_k \rangle), \quad j = 1, 2,$$

linearize  $P_1$  and  $P_2$ . Then we set  $n = n_1 + n_2 - 1$  and observe that the matrix

$$L = \begin{bmatrix} 0 & u_1^* & u_2^* \\ u_1 & Q_1 & 0 \\ u_2 & 0 & Q_2 \end{bmatrix} = \begin{bmatrix} 0 & u^* \\ u & Q \end{bmatrix} \in \mathcal{M}_{n_1+n_2-1}(\mathbb{C}\langle X_1, \dots, X_k \rangle).$$

is a linearization of  $P_1 + P_2$ .  $L$  is built so that  $[(ze_{1,1} - L)^{-1}]_{1,1} = (z - P)^{-1}$ ,  $z - P$  is invertible if and only if  $(ze_{1,1} - L)$  is invertible, and each row/column of the matrix  $L$ , except possibly for the first, contains at most one nonzero indeterminate (i.e. non-scalar). By applying the linearization process to  $1X_{i_1}X_{i_2}X_{i_3} \cdots X_{i_{\ell-1}}X_{i_\ell}1$  instead of  $X_{i_1}X_{i_2}X_{i_3} \cdots X_{i_{\ell-1}}X_{i_\ell}$ , we have insured that there is at most one nonzero indeterminate in each row/column. An important side benefit is that with this modification, we may assume that, with the notations from item 5 of Section 4,

$$v = u, \text{ and all entries of this vector are either } 0 \text{ or } 1.$$

While this linearization is far from being minimal, and should not be used for practical computations, it turns out to simplify to some extent the notations and arguments of our proofs.

In our arguments below we use several times the following equivalences regarding inequalities involving operators and their norms: let  $A$  be a bounded linear operator on a Hilbert space and let  $\mathbf{1}$  denote the identity operator on the same Hilbert space. Then  $\|A\|^2 = \|A^*\|^2 = \|A^*A\| = \|AA^*\| = \| \|A\|^2$  and

$$\|A\|^2 \leq M \iff A^*A \leq M \cdot \mathbf{1} \iff AA^* \leq M \cdot \mathbf{1}.$$



In particular, if  $A = A^*$ , then  $-\|A\| \cdot \mathbf{1} \leq A \leq \|A\| \cdot \mathbf{1}$ . If  $A > 0$  (that is,  $A = A^*$  and the spectrum of  $A$  is included in  $(0, +\infty)$ ), then  $A \geq \frac{1}{\|A^{-1}\|} \cdot \mathbf{1}$ . All these relations follow from functional calculus in  $C^*$ -algebras and the definition of positivity for operators (as the reader has by now noticed, we use the same symbol to denote inequalities between real numbers and inequalities between operators). In the future, we will sometimes suppress the identity in our notations and write, for instance,  $A \leq \|A\|$  instead of  $A \leq \|A\| \cdot \mathbf{1}$ .

The concrete expression of the inverse of  $ze_{1,1} - L$  in terms of  $L = \begin{bmatrix} 0 & u^* \\ u & Q \end{bmatrix}$  is provided by the Schur complement formula as

$$(ze_{1,1} - L)^{-1} = \begin{bmatrix} (z - u^*Q^{-1}u)^{-1} & -(z - u^*Q^{-1}u)^{-1}u^*Q^{-1} \\ -Q^{-1}u(z - u^*Q^{-1}u)^{-1} & Q^{-1} + Q^{-1}u(z - u^*Q^{-1}u)^{-1}u^*Q^{-1} \end{bmatrix}.$$

It follows easily from this formula that  $z - P$  is invertible if and only if  $ze_{1,1} - L$  is invertible. It was established in [5, Lemma 4.1] that  $Q$ , and hence  $Q^{-1}$ , is of the form  $T(1 + N)$  for some permutation scalar matrix  $T$  and nilpotent matrix  $N$ , which may contain non-scalar entries. Let us establish a version of [5, Lemma 4.3] suitable for our purposes.

**Lemma 5.6.** *Assume that  $P \in \mathbb{C}\langle Y_1, Y_2 \rangle$  is an arbitrary polynomial in the non-selfadjoint indeterminates  $Y_1$  and selfadjoint indeterminates  $Y_2$ . Let  $L$  be a linearization of  $P$  constructed as above. Given tuples of noncommutative random variables  $c$  and  $a$ , for all  $\delta > 0$  such that  $|P(c, a)|^2 > \delta$ , there exists  $\epsilon > 0$  such that  $|L(c, a)|^2 > \epsilon$ , and the number  $\epsilon$  only depends on  $\delta > 0$ ,  $P$ , and the supremum of the norms of  $c, a$ . Conversely, for all  $\epsilon > 0$  such that  $|L(c, a)|^2 > \epsilon$ , there exists  $q > 0$  such that  $|P(c, a)|^2 > q > 0$  and  $q$  depends only on  $\epsilon, P$ , and the supremum of the norms of  $c, a$ .*

*Proof.* With the decomposition  $L = \begin{bmatrix} 0 & u^* \\ u & Q \end{bmatrix}$ , we have  $|L|^2 = \begin{bmatrix} u^*u & u^*Q^* \\ Qu & uu^* + QQ^* \end{bmatrix}$ . Recall that  $|P|^2 = u^*Q^{-1}uu^*(Q^{-1})^*u$ . Now consider these expressions evaluated in the tuples of operators mentioned in the statement of the lemma. In order to save space, we will nevertheless suppress them from the notation. We assume that  $|P|^2 > \delta$ . Strangely enough, it will be more convenient to estimate an upper bound for  $|L|^{-2}$  rather than a lower bound for  $|L|^2$ . The entries of  $|L|^{-2}$  expressed in terms of the above decomposition are

$$\begin{aligned} (|L|^{-2})_{1,1} &= (u^*u - u^*Q^*(uu^* + QQ^*)^{-1}Qu)^{-1}, \\ (|L|^{-2})_{1,2} &= -(u^*u - u^*Q^*(uu^* + QQ^*)^{-1}Qu)^{-1}u^*Q^*(uu^* + QQ^*)^{-1}, \\ (|L|^{-2})_{2,1} &= -(uu^* + QQ^*)^{-1}Qu(u^*u - u^*Q^*(uu^* + QQ^*)^{-1}Qu)^{-1}, \\ (|L|^{-2})_{2,2} &= (uu^* + QQ^*)^{-1}Qu(u^*u - u^*Q^*(uu^* + QQ^*)^{-1}Qu)^{-1}u^*Q^*(uu^* + QQ^*)^{-1} \\ &\quad + (uu^* + QQ^*)^{-1}. \end{aligned}$$

We only need to estimate the norms of the above elements in terms of  $\delta, P$ , and the norms of the variables in which we have evaluated the above. It is clear that

$$\begin{aligned} (|L|^{-2})_{1,1} &= (u^*Q^{-1}u(1 + u^*(Q^*)^{-1}Q^{-1}u)^{-1}u^*(Q^*)^{-1}u)^{-1} \\ &= (P(1 + u^*(Q^*)^{-1}Q^{-1}u)^{-1}P^*)^{-1} \\ &\leq (P(\|1 + u^*(Q^*)^{-1}Q^{-1}u\|)^{-1}P^*)^{-1} \\ &= (1 + \|u^*(Q^*)^{-1}Q^{-1}u\|)|P|^{-2}. \end{aligned}$$

Similarly,  $(uu^* + QQ^*)^{-1} \leq (QQ^*)^{-1} \leq \|Q^{-1}\|^2$ . We obtain this way the following majorizations for each of the entries, which will allow us to estimate  $\epsilon$  (these majorizations

are not optimal, but close to):

$$\begin{aligned} \|(|L|^{-2})_{1,1}\| &\leq (1 + \|u^*(Q^*)^{-1}Q^{-1}u\|)\|P\|^{-2}, \\ \|(|L|^{-2})_{1,2}\| &\leq (1 + \|u^*(Q^*)^{-1}Q^{-1}u\|)\|P\|^{-2}\|u^*\| \|Q^*\| \|Q^{-1}\|^2, \\ \|(|L|^{-2})_{2,1}\| &\leq \|Q^{-1}\|^2 \|Q\| \|u\| (1 + \|u^*(Q^*)^{-1}Q^{-1}u\|)\|P\|^{-2}, \\ \|(|L|^{-2})_{2,2}\| &\leq \|Q^{-1}\|^4 \|Q\|^2 \|u\|^2 (1 + \|u^*(Q^*)^{-1}Q^{-1}u\|)\|P\|^{-2} + \|Q^{-1}\|^2. \end{aligned}$$

We shall not be much more explicit than this, but let us nevertheless comment on why the above satisfies the corresponding conclusion of our lemma. As noted before,  $u$  is a vector of zeros and ones. It follows immediately from the construction of  $L$  that the number of ones is dominated by the number of monomials of  $P$ , quantity clearly depending only on  $P$ . Recall that  $Q$  is of the form  $T(1 + N)$ , with  $T$  a permutation matrix, and  $N$  a nilpotent matrix. The norm of  $T$  is necessarily one. The nilpotent matrix corresponding to  $Q$  is simply a block upper diagonal matrix (i.e. a matrix which has on its diagonal a succession of blocks, each block being itself an upper diagonal matrix) with entries which are operators from the tuples  $a$  and  $c$  in which we evaluate  $P$  (and  $L$ ). Its norm is trivially bounded by the supremum of all the norms of the operators involved times the supremum of all the scalar coefficients. Since  $\|Q^{-1}\| = \|T^{-1}(1 + N)^{-1}\| \leq 1 + \sum_{j=1}^m \|N\|^j$ , where  $m$  is the size of the linearization, we obtain an estimate for  $\|Q^{-1}\|$  from above by  $(m + 1)(1 + \|Q\|)^m$ . Finally,  $\|P\|^{-2} \leq \delta^{-1}$ . This guarantees that  $\|L\|^{-2}$  is bounded from above, so that  $|L|^2$  is bounded from below, by a number  $\epsilon$  depending on  $\delta$ ,  $P$ , and the norms of the entries of  $P$ .

Conversely, assume that  $|L|^2 > \epsilon$  for a given strictly positive constant  $\epsilon$ . As before, this is equivalent to  $\|L\|^{-2} < \frac{1}{\epsilon}$ , which allows for the estimate of the  $(1, 1)$  entry of  $|L|^{-2}$  by  $\left\| (P(1 + u^*(Q^*)^{-1}Q^{-1}u)^{-1}P^*)^{-1} \right\| < \frac{1}{\epsilon}$ , so that

$$(P(1 + u^*(Q^*)^{-1}Q^{-1}u)^{-1}P^*)^{-1} < \frac{1}{\epsilon},$$

as inequality of operators. This tells us that  $P(1 + u^*(Q^*)^{-1}Q^{-1}u)^{-1}P^* > \epsilon$ , so that

$$PP^* > \frac{\epsilon}{\|(1 + u^*(Q^*)^{-1}Q^{-1}u)^{-1}\|} \geq \epsilon.$$

This concludes the proof. □

Part (3) is a simple formal step.

Step (4) becomes a direct consequence of Lemma 3.4.

Now, in step (5), we finally involve our variables  $c, a, a_N$  directly. We have assumed that  $|P(c, a) - z_0|^2 > \delta_{z_0} > 0$ , so that, according to steps (1) and (2), we have  $\left\| \begin{bmatrix} 0 & L(c, a) \\ L(c, a)^* & 0 \end{bmatrix} \right\| > \iota$  for a  $\iota > 0$  depending, according to step (2), only on  $P$ ,  $\delta_{z_0}$ , and the norms of  $c, a$ . According to step (4), it follows that

$$\begin{bmatrix} 0 & (I_m \otimes \epsilon)L(0, a) \\ (I_m \otimes \epsilon)L(0, a)^* & 0 \end{bmatrix} + \mathcal{S}$$

is invertible; moreover, the norm of the inverse is bounded in terms of  $P$ ,  $\delta_{z_0}$ , and the norms of  $c, a$ . According to Lemma 3.7 and Remark 3.6, denoting

$$\mathcal{Y} = \begin{bmatrix} 0 & (I_m \otimes \epsilon)L(0, a) \\ (I_m \otimes \epsilon)L(0, a)^* & 0 \end{bmatrix},$$

the condition of invertibility of  $\mathcal{S} + \mathcal{Y}$  is equivalent to the invertibility of  $\mathcal{Y}$  together with the existence of an  $r \in (0, 1)$  such that  $\text{spect}(\eta \circ G'_y(0)) \subset (1 - r)\mathbb{D}$ . We naturally denote

$\mathcal{Y}_N = \begin{bmatrix} 0 & (I_m \otimes \epsilon)L(0, a_N) \\ (I_m \otimes \epsilon)L(0, a_N)^* & 0 \end{bmatrix}$ . We have assumed that  $|P(0, a_N) - z_0|^2 > \delta_{z_0}$  for all (sufficiently large)  $N \in \mathbb{N}$ , so that  $|\mathcal{Y}_N|^2 > \zeta$  for a  $\zeta$  that only depends on  $P, \delta_{z_0}$ , and the supremum of the norms of  $a_N$ , which is assumed to be bounded. Thus,  $|\mathcal{Y}_N|^2$  is uniformly bounded from below as  $N \rightarrow \infty$ . In order to insure the invertibility of  $\mathcal{S} + \mathcal{Y}_N$ , we also need that  $\text{spect}(\eta \circ G'_{\mathcal{Y}_N}(0)) \subset \overline{\mathbb{D}} \setminus \{1\}$ , for all  $N$  sufficiently large. The existence of  $G'_{\mathcal{Y}_N}(0)$  is guaranteed by the hypothesis of invertibility of  $\mathcal{Y}_N$ . Since

$$G'_{\mathcal{Y}_N}(0)(c) = (id_m \otimes \varphi) [\mathcal{Y}_N^{-1}c\mathcal{Y}_N^{-1}],$$

and

$$\mathcal{Y}_N^{-1} = \begin{bmatrix} 0 & (I_m \otimes \epsilon)(L(0, a_N)^*)^{-1} \\ (I_m \otimes \epsilon)L(0, a_N)^{-1} & 0 \end{bmatrix},$$

we only need to remember that all entries of  $L(0, a_N)^{-1}$  are products of polynomials in  $a_N$  and  $(P(0, a_N) - z_0)^{-1}$  in order to conclude that the convergence in moments of  $a_N$  to  $a$  implies the convergence in norm of  $G'_{\mathcal{Y}_N}(0)$  to  $G'_Y(0)$  (recall that, according to hypothesis 2. in the statement of our proposition,  $|P(0, a_N) - z_0|^2 > \delta_{z_0} > 0$  uniformly). Thus, for  $N$  sufficiently large, all eigenvalues of  $\eta \circ G'_{\mathcal{Y}_N}(0)$  are included in  $(1 - \frac{\epsilon}{2})\mathbb{D}$ . This guarantees the invertibility of all  $\mathcal{S} + \mathcal{Y}_N$  for  $N$  sufficiently large.

To prove item (6) and conclude our proof, we only need to show that for  $N$  sufficiently large,  $|\mathcal{S} + \mathcal{Y}_N|^2 > \frac{\epsilon}{2}$ . There is a simple abstract shortcut for this: as Proposition 3.5 shows, the endpoint of the support of the (scalar) distribution of  $\mathcal{S} + \mathcal{Y}_N$  is given by that smallest  $x_N \in (0, +\infty)$  for which  $1 \in \text{spect}(\eta \circ G'_{\mathcal{Y}_N}(x_N))$  (as usual in this context, by  $G_{\mathcal{Y}_N}(x_N)$  we mean  $G_{\mathcal{Y}_N}(x_N \cdot I_{2m})$ ). On the one hand,  $G_{\mathcal{Y}_N}$  is guaranteed to be analytic on  $[0, \delta_{z_0}]$ . On the other, since  $\mathcal{Y}_N \rightarrow \mathcal{Y}$  in distribution, we have  $G_{\mathcal{Y}_N} \rightarrow G_Y$  uniformly on  $[0, \delta_{z_0} - \epsilon]$  for any fixed  $\epsilon > 0$ . In particular,  $G'_{\mathcal{Y}_N}(x) \rightarrow G'_Y(x)$  for any  $x$  in this interval. Thus,  $x_N$  is bounded away from zero uniformly in  $N$  as  $N \rightarrow \infty$ . A second application of the convergence of  $G_{\mathcal{Y}_N}$  allows us to conclude.

### 6 Stable outliers; proof of Theorem 1.10

Making use of a linearization procedure, the proof closely follows the approach of [10]. The most significant novelty is Proposition 6.1 which substantially generalizes Theorem 1.3. A. in [15] (see also Proposition 2.1 in [10]) and whose proof relies on operator-valued free probability results established in Section 3.2.2. Nevertheless, we provide all arguments for the reader's convenience.

Let

$$L_P = \gamma \otimes 1 + \sum_{j=1}^u \zeta_j \otimes y_j + \sum_{k=1}^t \beta_k \otimes y_{u+k},$$

be a linearization of  $P(y_1, \dots, y_{u+t})$  with coefficients in  $\mathcal{M}_m(\mathbb{C})$  such that, for any  $u + t$ -tuple  $y$  of matrices,  $|\det Q(y)| = 1$  (see (4.3)).

Let  $G$  be a relatively compact set in  $\mathbb{C} \setminus \text{spect}(P(c, a))$  satisfying the hypotheses of Theorem 1.10, and  $\Gamma = \overline{G}$ . Note that

$$\min_{z \in \partial \Gamma} \left| \frac{\det(zI_N - P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}))}{\det(zI_N - P(0_N, \dots, 0_N, (A_N^{(1)})', \dots, (A_N^{(t)})'})} \right| \geq \epsilon$$

is equivalent to

$$\begin{aligned} \min_{z \in \partial \Gamma} \left| \frac{\det(zI_N - P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}))}{\det(zI_N - P(0_N, \dots, 0_N, (A_N^{(1)})', \dots, (A_N^{(t)})'})} \right| & \quad (6.1) \\ \times \frac{\det(Q(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}))}{\det Q(0_N, \dots, 0_N, (A_N^{(1)})', \dots, (A_N^{(t)})')} & \geq \epsilon, \end{aligned}$$

since  $|\det Q(y)|$  is constant. Now, following the proof of Lemma 4.3 in [5], one can see that this is also equivalent to

$$\min_{z \in \partial\Gamma} \left| \frac{\det(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes A_N^{(k)})}{\det(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})')} \right| \geq \epsilon. \tag{6.2}$$

According to Lemma 4.1, the eigenvalues of  $M_N$  are the zeroes of  $z \mapsto \det(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{j=1}^u \zeta_j \otimes \frac{X_N^{(j)}}{\sqrt{N}} - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)}))$ . By Assumption  $(A'_2)$ , Proposition 5.1 and Lemma 4.1, almost surely for all large  $N$ , for any  $z \in \Gamma$ , we can define

$$R_N(z) = (ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{j=1}^u \zeta_j \otimes \frac{X_N^{(j)}}{\sqrt{N}} - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})')^{-1},$$

$$R'_N(z) = (ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})')^{-1}.$$

Since each  $(A_N^{(k)})''$  has a bounded rank  $r_k(N) = O(1)$ , there exist matrices  $P_N \in \mathcal{M}_{mN,p}$ ,  $Q_N \in M_{p,mN}$ , where  $p$  is fixed, such that

$$\sum_{k=1}^t \beta_k \otimes (A_N^{(k)})'' = P_N Q_N. \tag{6.3}$$

Recall the Weinstein–Aronszajn identity: if  $P, Q^\top \in \mathcal{M}_{d_1, d_2}(\mathbb{C})$ ,

$$\det(I_{d_1} + PQ) = \det(I_{d_2} + QP),$$

where  $Q^\top$  denotes the transpose of  $Q$ . Using this identity, it is clear that, almost surely for all large  $N$ , the eigenvalues of  $M_N$  in  $\Gamma$  are precisely the zeros of the random analytic function  $z \mapsto \det(I_p - Q_N R_N(z) P_N)$  in that set.

Similarly, for any  $z$  in  $\Gamma$ ,

$$\det(I_p - Q_N R'_N(z) P_N) = \frac{\det(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes A_N^{(k)})}{\det(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})')}. \tag{6.4}$$

Thus, the zeroes of  $z \mapsto \det(I_p - Q_N R'_N(z) P_N)$  in  $\Gamma$  are the zeroes of  $z \mapsto \det(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_k \beta_k \otimes A_N^{(k)})$  in  $\Gamma$ , that is, the eigenvalues in  $\Gamma$  of

$$M_N^{(0)} = P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}).$$

The rest of the proof is devoted to establish that  $\det(I_p - Q_N R_N(z) P_N) - \det(I_p - Q_N R'_N(z) P_N)$  converges uniformly in  $\Gamma$  to zero.

**Step 1: Iterated resolvent identities.**

Set

$$Y_N = \sum_{j=1}^u \zeta_j \otimes \frac{X_N^{(j)}}{\sqrt{N}}.$$

Using repeatedly the resolvent identity,

$$R_N(z) = R'_N(z) + R'_N(z) Y_N R_N(z),$$

we find that, for any integer number  $K \geq 2$ ,

$$\begin{aligned}
 & Q_N R_N(z) P_N - Q_N R'_N(z) P_N \\
 &= \sum_{k=1}^{K-1} Q_N (R'_N(z) Y_N)^k R'_N(z) P_N + Q_N (R'_N(z) Y_N)^K R_N(z) P_N. \tag{6.5}
 \end{aligned}$$

The following two steps will be of basic use to prove the uniform convergence in  $\Gamma$  of the right hand side of (6.5) towards zero.

**Step 2: Study of the spectral radius of  $R'_N(z) Y_N$ .**

The aim of this second step is to prove Lemma 6.5 which establishes an upper bound strictly smaller than 1 of the spectral norm of  $R'_N(z) Y_N$ . The proof of Lemma 6.5 is based on Proposition 5.1 and the characterization, provided by Lemma 3.7, of the invertibility of the sum of a centered  $\mathcal{M}_m(\mathbb{C})$ -valued semi-circular  $s$  and some selfadjoint  $y \in \mathcal{M}_m(\mathcal{A})$  with non-commutative symmetric entries such that  $s$  and  $y$  are free over  $\mathcal{M}_m(\mathbb{C})$ . Recall that  $\mu_z$  is the distribution of

$$\left[ P(c^{(1)}, \dots, c^{(u)}, a^{(1)}, \dots, a^{(t)}) - z1_{\mathcal{A}} \right] \left[ P(c^{(1)}, \dots, c^{(u)}, a^{(1)}, \dots, a^{(t)}) - z1_{\mathcal{A}} \right]^*.$$

Define  $\nu_z$  as the distribution of

$$\left[ P(0_{\mathcal{A}}, \dots, 0_{\mathcal{A}}, a^{(1)}, \dots, a^{(t)}) - z1_{\mathcal{A}} \right] \left[ P(0_{\mathcal{A}}, \dots, 0_{\mathcal{A}}, a^{(1)}, \dots, a^{(t)}) - z1_{\mathcal{A}} \right]^*,$$

and

$$S_0 = \{z \in \mathbb{C}, 0 \in \text{supp}(\nu_z)\}.$$

**Proposition 6.1.** *Let*

$$L_P = \gamma \otimes 1 + \sum_{j=1}^u \zeta_j \otimes y_j + \sum_{k=1}^t \beta_k \otimes y_{u+k},$$

be a linearization of  $P(y_1, \dots, y_{u+t})$  with coefficients in  $\mathcal{M}_m(\mathbb{C})$ . Set

$$y_z = \sum_{k=1}^t \beta_k \otimes a^{(k)} + (\gamma - ze_{11}) \otimes 1_{\mathcal{A}}.$$

Let  $\epsilon$  be some selfadjoint  $\{-1, +1\}$ -Bernoulli variable in  $\mathcal{A}$  independent from the entries of  $y_z$ . Let  $s_1, \dots, s_u$  be free semicircular variables in  $\mathcal{A}$  free from  $\epsilon$  and the entries of  $y_z$ . Define

$$\mathcal{Y}_z = \begin{pmatrix} 0 & (I_m \otimes \epsilon)y_z \\ (I_m \otimes \epsilon)y_z^* & 0 \end{pmatrix} \text{ and } S = \begin{pmatrix} 0 & \sum_{j=1}^u \zeta_j \otimes s_j \\ \sum_{j=1}^u \zeta_j^* \otimes s_j & 0 \end{pmatrix}.$$

If  $z \notin S_0$ , let  $\Delta_1(z)$  be the operator

$$\begin{aligned}
 & \mathcal{M}_{2m}(\mathbb{C}) \rightarrow \mathcal{M}_{2m}(\mathbb{C}) \\
 & b \mapsto id_{2m} \otimes \varphi \left( \mathcal{S}([id_{2m} \otimes \varphi((\mathcal{Y}_z)^{-1}(b \otimes 1)(\mathcal{Y}_z)^{-1})] \otimes 1) \mathcal{S} \right).
 \end{aligned}$$

We have  $0 \notin \text{supp}(\mu_z)$  iff  $z \notin S_0$  and  $\text{spect}(\Delta_1(z)) \subseteq \overline{\mathbb{D}} \setminus \{1\}$ .

*Proof.* According to Remark 1.4, we have that  $0 \notin \text{supp}(\mu_z)$  if and only if  $P(c^{(1)}, \dots, c^{(u)}, a^{(1)}, \dots, a^{(t)}) - z1$  is invertible. According to Lemma 4.1, it follows that  $0 \notin \text{supp}(\mu_z)$  if and only if  $\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y_z$  is invertible. Now,  $\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y_z$  is invertible if

and only if both  $\left[\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y_z\right] \left[\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y_z\right]^*$  and  $\left[\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y_z\right]^* \times \left[\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y_z\right]$  are invertible, and then, by Lemma 3.4, since  $\text{tr}_m \otimes \varphi$  is faithful, if and only if  $\left[\sum_{j=1}^u \zeta_j \otimes s_j + (I_m \otimes \epsilon)y_z\right] \left[\sum_{j=1}^u \zeta_j \otimes s_j + (I_m \otimes \epsilon)y_z\right]^*$  and  $\left[\sum_{j=1}^u \zeta_j \otimes s_j + (I_m \otimes \epsilon)y_z\right]^* \left[\sum_{j=1}^u \zeta_j \otimes s_j + (I_m \otimes \epsilon)y_z\right]$  are invertible, that is if and only if  $\mathcal{S} + \mathcal{Y}_z$  is invertible. Thus, Proposition 6.1 follows from Example 3.1 and Lemma 3.7.  $\square$

Define for any  $w, z$  in  $\mathbb{C}$ ,  $\mu_{w,z}$  as the distribution of

$$\left[P(wc^{(1)}, \dots, wc^{(u)}, a^{(1)}, \dots, a^{(t)}) - zI\right] \left[P(wc^{(1)}, \dots, wc^{(u)}, a^{(1)}, \dots, a^{(t)}) - zI\right]^*.$$

**Lemma 6.2.**  $0 \notin \text{supp}(\mu_{w,z})$  if and only if  $z \notin S_0$  and  $\text{spect}(|w|^2 \Delta_1(z)) \subseteq \overline{\mathbb{D}} \setminus \{1\}$ , where  $S_0$  and  $\Delta_1(z)$  are defined in Proposition 6.1.

*Proof.* Note that  $(c^{(1)}, \dots, c^{(u)})$  and  $(\exp(i \arg w)c^{(1)}, \dots, \exp(i \arg w)c^{(u)})$  have the same  $*$ -distribution so that  $\mu_{w,z}$  is the distribution of

$$\begin{aligned} &\left[P(|w|c^{(1)}, \dots, |w|c^{(u)}, a^{(1)}, \dots, a^{(t)}) - zI\right] \\ &\quad \times \left[P(|w|c^{(1)}, \dots, |w|c^{(u)}, a^{(1)}, \dots, a^{(t)}) - zI\right]^*. \end{aligned}$$

Then the result follows from Proposition 6.1.  $\square$

**Lemma 6.3.** Let  $\Gamma$  be a compact subset in  $\{z \in \mathbb{C} : 0 \notin \text{supp}(\mu_z)\}$ . Then there exists  $\rho > 1$  such that for any  $w \in \mathbb{C}$  such that  $|w| \leq \rho$  and any  $z \in \Gamma$ , we have  $0 \notin \text{supp}(\mu_{w,z})$ .

*Proof.* Let  $z$  be in  $\Gamma$ . According to Proposition 6.1,  $z \notin S_0$  and  $\text{spect}(\Delta_1(z)) \subseteq \overline{\mathbb{D}} \setminus \{1\}$ . According to [18, Theorem 2.5], if  $r(z)$  is the spectral radius of the positive linear map  $\Delta_1(z)$ , then there exists a nonzero positive element  $\xi$  in  $\mathcal{M}_{2m}(\mathbb{C})$  such that  $\Delta_1(z)(\xi) = r(z)\xi$ . Thus, we can deduce that  $r(z) < 1$ . Now, since  $\{z \in \mathbb{C} : 0 \notin \text{supp}(\mu_z)\} \subset \mathbb{C} \setminus S_0$ , using Remark 1.4 and Lemma 4.1, it is easy to see that  $(z \mapsto r(z))$  is continuous on  $\{z \in \mathbb{C} : 0 \notin \text{supp}(\mu_z)\}$ . Thus, there exists  $0 < \gamma < 1$  such that for any  $z \in \Gamma$ , we have  $0 \leq r(z) < 1 - \gamma$ . It readily follows that if  $|w| \leq \frac{1}{\sqrt{1-\gamma}}$  then  $|w|^2 r(z) < 1$  and according to Lemma 6.2,  $0 \notin \text{supp}(\mu_{w,z})$ .  $\square$

**Lemma 6.4.** Let  $\Gamma$  be a compact subset in  $\{z \in \mathbb{C}, 0 \notin \text{supp}(\mu_z)\}$ . Assume that  $(\mathbf{A}'_2)$  holds. Then there exists  $\rho > 1$  and  $\eta > 0$  such that a.s. for all large  $N$ , for any  $w \in \mathbb{C}$  such that  $|w| \leq \rho$  and any  $z \in \Gamma$ , there is no singular value of

$$P \left( w \frac{X_N^{(1)}}{\sqrt{N}}, \dots, w \frac{X_N^{(u)}}{\sqrt{N}}, (A_N^{(1)})', \dots, (A_N^{(t)})' \right) - zI_N$$

in  $[0, \eta]$ .

*Proof.* Let  $\tilde{\Gamma} = \{(w, z) \in \mathbb{C}^2, |w| \leq \rho, z \in \Gamma\}$  where  $\rho$  is defined in Lemma 6.3. According to Lemma 6.3,  $\forall (w, z) \in \tilde{\Gamma}$ ,  $0 \notin \text{supp}(\mu_{w,z})$ . Therefore, using  $(\mathbf{A}'_2)$ , according to Proposition 5.1, there exists  $\gamma(w, z)$  such that a.s. for all large  $N$ , there is no singular value of

$$P \left( w \frac{X_N^{(1)}}{\sqrt{N}}, \dots, w \frac{X_N^{(u)}}{\sqrt{N}}, (A_N^{(1)})', \dots, (A_N^{(t)})' \right) - zI_N$$

in  $[0, \gamma(w, z)]$ . The conclusion follows by a compactness argument (using Bai-Yin's theorem and (1.5)).  $\square$

**Lemma 6.5.** *Let  $\Gamma$  be a compact subset in  $\{z \in \mathbb{C} : 0 \notin \text{supp}(\mu_z)\}$ . Assume that  $(\mathbf{A}'_2)$  and (1.5) hold. There exists  $0 < \epsilon_0 < 1$  such that almost surely for all large  $N$ , we have,*

$$\sup_{z \in \Gamma} \rho(R'_N(z)Y_N) \leq 1 - \epsilon_0,$$

where  $\rho(M)$  denotes the spectral radius of a matrix  $M$ .

*Proof.* Now, assume that  $\lambda \neq 0$  is an eigenvalue of  $R'_N(z)Y_N$ . Then there exists  $v \in \mathbb{C}^{Nm}$ ,  $v \neq 0$  such that  $(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})')^{-1} Y_N v = \lambda v$  and thus  $(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})' - \sum_{j=1}^u \zeta_j \otimes \lambda^{-1} \frac{X_N^{(j)}}{\sqrt{N}}) v = 0$ . This means that  $z$  is an eigenvalue of

$$P \left( \lambda^{-1} \frac{X_N^{(1)}}{\sqrt{N}}, \dots, \lambda^{-1} \frac{X_N^{(u)}}{\sqrt{N}}, (A_N^{(1)})', \dots, (A_N^{(t)})' \right),$$

or equivalently that 0 is a singular value of

$$P \left( \lambda^{-1} \frac{X_N^{(1)}}{\sqrt{N}}, \dots, \lambda^{-1} \frac{X_N^{(u)}}{\sqrt{N}}, (A_N^{(1)})', \dots, (A_N^{(t)})' \right) - zI_N.$$

By Lemma 6.4, we can deduce that almost surely for all large  $N$ , the nonnull eigenvalues of  $R'_N(z)Y_N$  must satisfy  $1/|\lambda| > \rho$ . The result follows.  $\square$

**Step 3: Study of the moments of  $R'_N(z)Y_N$ .**

**Proposition 6.6.** *Let  $\Gamma$  be a compact subset in  $\{z \in \mathbb{C}, 0 \notin \text{supp}(\mu_z)\}$ . Assume that  $(\mathbf{A}'_2)$  and (1.5) hold. There exists  $0 < \epsilon_0 < 1$  and  $C > 0$  such that almost surely for all large  $N$ , for any  $k \geq 1$ ,*

$$\sup_{z \in \Gamma} \left\| (R'_N(z)Y_N)^k \right\| \leq C(1 - \epsilon_0)^k.$$

*Proof.* For  $z \in \Gamma$ , we set  $T_N(z) = R'_N(z)Y_N$ . Let  $\epsilon_0$  be as defined by Lemma 6.5 and  $\rho$  be as defined in Lemma 6.4. Choose  $0 < \epsilon < \min(\epsilon_0, 1 - \frac{1}{\rho})$ . Therefore, according to Lemma 6.5 and using Dunford-Riesz calculus, we have almost surely for all large  $N$ , for any  $z$  in  $\Gamma$ ,

$$\forall k \geq 0, (T_N(z))^k = \frac{1}{2i\pi} \int_{|w|=1-\epsilon} w^k (w - T_N(z))^{-1} dw,$$

and therefore

$$\forall k \geq 0, \|(T_N(z))^k\| \leq \sup_{|w|=1-\epsilon} \|(w - T_N(z))^{-1}\| (1 - \epsilon)^{k+1}. \tag{6.6}$$

Now, note that, for any  $w$  such that  $|w| = 1 - \epsilon$ , we have  $\frac{1}{|w|} < \rho$  and  $(w - T_N(z)) =$

$$wR'_N(z) \left( ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{j=1}^u \zeta_j \otimes w^{-1} \frac{X_N^{(j)}}{\sqrt{N}} - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})' \right),$$

so that

$$(w - T_N(z))^{-1} =$$

$$\left( ze_{11} \otimes I_N - L_P(w^{-1} \frac{X_N^{(1)}}{\sqrt{N}}, \dots, w^{-1} \frac{X_N^{(u)}}{\sqrt{N}}, (A_N^{(1)})', \dots, (A_N^{(t)})') \right)^{-1}$$

$$\times \frac{1}{w} \left( ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})' \right). \tag{6.7}$$

Lemma 6.4 readily implies that almost surely for all large  $N$ ,

$$\left\| \left( zI_N - P(w^{-1} \frac{X_N^{(1)}}{\sqrt{N}}, \dots, w^{-1} \frac{X_N^{(t)}}{\sqrt{N}}, (A_N^{(1)})', \dots, (A_N^{(t)})' \right)^{-1} \right\| \leq 1/\eta, \tag{6.8}$$

where  $\eta$  is defined in Lemma 6.4. It readily follows from (6.7), Lemma 4.2, (6.8), (1.5) and Bai-Yin’s theorem that there exists  $C > 0$  such that we have almost surely for all large  $N$ , for any  $z$  in  $\Gamma$ ,

$$\sup_{|w|=1-\epsilon} \|(w - T_N(z))^{-1}\| \leq C. \tag{6.9}$$

Proposition 6.6 follows from (6.6) and (6.9). □

**Step 4: Conclusion.**

We will use the following proposition from [10] to establish Lemma 6.8 below.

**Proposition 6.7** ([10]). *Let  $n \geq 1$  be an integer and  $Q \in \mathbb{C}\langle X_1, \dots, X_n \rangle$  such that the total exponent of  $X_n$  in each monomial of  $Q$  is nonzero. We consider a sequence  $(B_N^{(1)}, \dots, B_N^{(n-1)}) \in \mathcal{M}_N(\mathbb{C})^{n-1}$  of matrices with operator norm uniformly bounded in  $N$  and  $u_N, v_N$  in  $\mathbb{C}^N$  with unit norm. Then if  $X_N$  is a  $N \times N$  matrix with i.i.d. entries centered with variance 1 and finite fourth moment a.s.*

$$u_N^* Q \left( B_N^{(1)}, \dots, B_N^{(n-1)}, \frac{X_N}{\sqrt{N}} \right) v_N \rightarrow 0.$$

**Lemma 6.8.** *Assume  $(\mathbf{X}_1)$ , (1.3) and  $(\mathbf{A}_2)$ . For any  $z$  in  $\Gamma \subset \mathbb{C} \setminus \text{spect}(P(c, a))$ , almost surely, the series  $\sum_{k \geq 1} Q_N (R'_N(z) Y_N)^k R'_N(z) P_N$  converges in norm to zero when  $N$  goes to infinity, where  $P_N$  and  $Q_N$  are defined by (6.3). Here we assume that  $\Gamma$  satisfies the hypotheses of Theorem 1.10.*

*Proof.* The singular value decomposition of  $\sum_k \beta_k \otimes (A_N^{(k)})''$  gives that for any  $i, j \in \{1, \dots, p\}$ ,

$$(Q_N (R'_N(z) Y_N)^k R'_N(z) P_N)_{ij} = s_i v_i^* (R'_N(z) Y_N)^k R'_N(z) u_j,$$

where  $u_j$  and  $v_j$  are unit vectors in  $\mathbb{C}^{Nm}$  and  $s_i$  is a singular value of  $\sum_{k=1}^t \beta_k \otimes (A_N^{(k)})''$ . According to (1.3) and (1.5), the  $s_i$ ’s are uniformly bounded. Using  $(\mathbf{A}_2)$ , (1.5) and (4.2), almost surely for any  $z$  in  $\Gamma$ , there exists  $\tilde{\eta}_z > 0$  such that for all large  $N$ ,

$$\|R'_N(z)\| \leq \frac{1}{\tilde{\eta}_z}. \tag{6.10}$$

Using (6.10) and Bai-Yin’s theorem, we deduce from Proposition 6.7 that  $v_i^* (R'_N(z) Y_N)^k \times R'_N(z) u_j$  converges almost surely to zero. The result follows by applying the dominated convergence theorem thanks to Proposition 6.6. □

We are going to prove that, assuming  $(\mathbf{X}_1)$ , (1.3) and  $(\mathbf{A}_2)$ , we have for any  $z$  in  $\Gamma$ , almost surely, as  $N \rightarrow \infty$ ,

$$\|Q_N R_N(z) P_N - Q_N R'_N(z) P_N\| \rightarrow 0. \tag{6.11}$$

Let  $C' > 0$  such that

$$\|P_N\| \|Q_N\| \leq C'. \tag{6.12}$$



According to Proposition 5.1 and (4.2), for any  $z \in \Gamma$ , there exists  $\tilde{\gamma}_z > 0$  such that almost surely for all large  $N$

$$\|R_N(z)\| \leq \frac{1}{\tilde{\gamma}_z}. \tag{6.13}$$

Then using also Proposition 6.6 and (6.10), for any  $k \geq 1$ , we have

$$\begin{aligned} \|Q_N (R'_N(z)Y_N)^k R'_N(z)P_N\| &\leq \frac{CC'}{\tilde{\eta}_z}(1 - \epsilon_0)^k, \\ \|Q_N (R'_N(z)Y_N)^k R_N(z)P_N\| &\leq \frac{CC'}{\tilde{\gamma}_z}(1 - \epsilon_0)^k. \end{aligned}$$

Let  $\eta > 0$ . Choose  $K \geq 1$  such that  $\frac{CC'}{\tilde{\gamma}_z}(1 - \epsilon_0)^K < \eta/2$  and  $\sum_{k \geq K} \frac{CC'}{\tilde{\eta}_z}(1 - \epsilon_0)^k < \eta/2$ . Thus, using (6.5), we have that, for any  $\eta > 0$ ,

$$\left\| Q_N R_N(z)P_N - Q_N R'_N(z)P_N - \sum_{k \geq 1} Q_N (R'_N(z)Y_N)^k R'_N(z)P_N \right\| < \eta$$

and then, letting  $\eta$  going to zero, that we have

$$Q_N R_N(z)P_N - Q_N R'_N(z)P_N = \sum_{k \geq 1} Q_N (R'_N(z)Y_N)^k R'_N(z)P_N. \tag{6.14}$$

Applying Lemma 6.8, we obtain (6.11).

**Proposition 6.9.** *Let  $\Gamma$  be a compact subset of  $\mathbb{C} \setminus \text{spect}(P(c, a))$  which satisfies the hypotheses of Theorem 1.10. Assume  $(\mathbf{X}_1)$ , (1.3) and  $(\mathbf{A}_2)$ . Then, almost surely,  $\det(I_p - Q_N R_N(z)P_N) - \det(I_p - Q_N R'_N(z)P_N)$  converges to zero uniformly on  $\Gamma$ , when  $N$  goes to infinity.*

*Proof.* It sufficient to check that for any  $\delta > 0$ , a.s., for all large  $N$ ,

$$\sup_{z \in \Gamma} \|Q_N R_N(z)P_N - Q_N R'_N(z)P_N\| \leq 3\delta. \tag{6.15}$$

We set  $\zeta_z = \tilde{\eta}_z \wedge \tilde{\gamma}_z$  and  $r_z = (\zeta_z/2) \wedge (\delta(\zeta_z^2/2C'))$  where  $\tilde{\eta}_z, \tilde{\gamma}_z$  and  $C'$  are defined in (6.10), (6.13) and (6.12). Using the resolvent identity, (6.10) and (6.13), if  $(z, w) \in \Gamma^2$  are such that  $|z - w| \leq r_z$ , then

$$\begin{aligned} \|Q_N R_N(z)P_N - Q_N R_N(w)P_N\| &\leq \frac{2C'}{\zeta_z^2}|z - w| \leq \delta, \\ \|Q_N R'_N(z)P_N - Q_N R'_N(w)P_N\| &\leq \frac{2C'}{\zeta_z^2}|z - w| \leq \delta. \end{aligned}$$

Since  $\Gamma \subset \cup_{z \in \Gamma} B(z, r_z)$  and  $\Gamma$  compact, there is a finite covering and the proposition follows from (6.11).  $\square$

Theorem 1.10 follows from Proposition 6.9 by Rouché’s Theorem, using (6.4) and (6.2).

### 7 Proof of Corollary 1.11

Using Remark 1.12, assume that  $P$  has no constant term. Note that, by Proposition 6.1 and (1.4), this implies that 0 belongs to the spectrum of  $P(c, 0)$ . There exists  $C > 0$  large enough such that  $\text{spect}(P(c, 0)) \subset \{z: |z| \leq C\}$  and, by Bai-Yin’s theorem (see [3, Theorem 5.8]) and the fact that for  $k = 1, \dots, t$ ,  $A_N^{(k)}$  are deterministic matrices

with norm  $O(1)$ , such that, almost surely for all large  $N$ ,  $\|M_N\| \leq C$ ,  $\|M_N^{(0)}\| \leq C$ . Denote  $C_\iota = \{z \in \mathbb{C} : d(z, \text{spect}(P(c, 0))) \leq \iota\}$ . This is a compact set for any  $\iota \geq 0$ . Given  $\delta > 0$ , the set  $\{z \in \mathbb{C} : d(z, C_\iota) < \delta\} = \bigcup_{v \in C_\iota} \{z : |z - v| < \delta\}$  includes the compact  $C_\iota$ , so that  $\{\{z : |z - v| < \delta\}\}_{v \in C_\iota}$  is an open cover. Extract a finite subcover  $\{\{z : |z - v_j| < \delta\}\}_{j=1}^n$ . Denote  $K_\delta^\iota = \bigcup_{j=1}^n \{z : |z - v_j| < \delta\}$ . It is obvious that  $\partial K_\delta^\iota$  is a finite union of rectifiable curves for any  $\delta > 0, \iota \geq 0$ . Assume that  $\delta > 0$  is sufficiently small so that  $C_0 \subset C_{2\epsilon} \subset K_\delta^{2\epsilon} \subset C_{5\epsilon/2}$ . By increasing if necessary the value of  $\delta$  by an arbitrarily small positive amount, we may assume that  $C_0 \subset C_{2\epsilon} \subset K_\delta^{2\epsilon} \subset C_{3\epsilon}$  and in addition  $(\mathbb{C} \setminus K_\delta^\epsilon) \cap \{z \in \mathbb{C} : |z| < R\}$  satisfies the hypotheses imposed in Theorem 1.10 for any  $R > 0$  sufficiently large. Choose  $K > C + 4\epsilon$  and consider the set

$$G = \{z \in \mathbb{C} \setminus K_\delta^{2\epsilon}, |z| < K\}, \quad \Gamma = \overline{G}.$$

This set satisfies the hypotheses of Theorem 1.10. Since 0 belongs to the spectrum of  $P(c, 0)$ , any  $z$  in  $\Gamma$  satisfies  $|z| \geq 2\epsilon$ . Now,  $s_N(P(0, 0) - zI_N) = |z|$ , so that  $(A_2')$  is satisfied with for any  $k = 1, \dots, t$ ,  $(A_N^{(k)})' = 0$ . If  $|z| = K$  then for any  $l = 1, \dots, N$ ,

$$|z - \lambda_l(M_N^{(0)})| \geq \epsilon.$$

Moreover, since for all sufficiently large  $N$ , there are no eigenvalues of  $M_N^{(0)} = P(0, \dots, 0, A_N^{(1)}, \dots, A_N^{(t)})$  in  $\{z \in \mathbb{C}, \epsilon < d(z, \text{spect}(P(c, 0))) < 4\epsilon\}$ , we can deduce that, if  $z \in \partial K_\delta^{2\epsilon}$ , we also have that, for all sufficiently large  $N$ , for any  $l = 1, \dots, N$ ,  $|z - \lambda_l(M_N^{(0)})| \geq \epsilon$ . Since for  $k = 1, \dots, t$ ,  $A_N^{(k)}$  are deterministic matrices with rank  $O(1)$ ,  $M_N^{(0)}$  has rank  $r(N) = O(1)$ . Let  $m_0(N)$  the multiplicity of 0 as a root of the characteristic polynomial of  $M_N^{(0)}$ . Note that  $N - m_0(N) \leq r(N) \leq r$  for some  $r > 0$ . Thus,

$$\begin{aligned} & \min_{z \in \partial \Gamma} \left| \frac{\det(zI_N - P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}))}{\det(zI_N - P(0_N, \dots, 0_N, (A_N^{(1)})', \dots, (A_N^{(t)})'))} \right| \\ &= \min_{z \in \partial \Gamma} \frac{\prod_{l=1}^{N-m_0(N)} |z - \lambda_l(M_N^{(0)})|}{|z|^{N-m_0(N)}} \\ &\geq \left(\frac{\epsilon}{K}\right)^{N-m_0(N)} \\ &\geq \left(\frac{\epsilon}{K}\right)^r. \end{aligned}$$

Thus (1.6) is satisfied and we can deduce from Theorem 1.10 that a.s, for all large  $N$ , there are precisely  $j$  eigenvalues of  $M_N = P\left(\frac{X_N^{(1)}}{\sqrt{N}}, \dots, \frac{X_N^{(u)}}{\sqrt{N}}, A_N^{(1)}, \dots, A_N^{(t)}\right)$  in  $\Gamma \subseteq \{z \in \mathbb{C}, d(z, \text{spect}(P(c, 0))) \geq 2\epsilon\}$ .

Denote by  $\{\lambda_1(M_N^{(0)}), \dots, \lambda_j(M_N^{(0)})\}$  the set of eigenvalues that belong to the set  $\{z \in \mathbb{C}, d(z, \text{spect}(P(c, 0))) \geq 2\epsilon\}$ . By passing if necessary to a subsequence, we may assume that for any  $l = 1, \dots, j$ ,  $\lambda_l(M_N^{(0)})$  converges to  $\lambda_l \in \{z \in \mathbb{C}, d(z, \text{spect}(P(c, 0))) \geq 2\epsilon\}$ . Let  $0 < \delta < \epsilon$  such that  $2\delta < \min_{\lambda_i \neq \lambda_{i'}} |\lambda_i - \lambda_{i'}|$  and for  $l = 1, \dots, j$ , let  $\Gamma_l = B(\lambda_l; \delta) \subset \{z \in \mathbb{C}, d(z, \text{spect}(P(c, 0))) \geq 2\epsilon\}$ .

$\mathbb{C}, |z| \leq K, d(z, \text{spect}(P(c, 0))) \geq 2\epsilon\}$ .

$$\begin{aligned} & \min_{z \in \partial\Gamma_l} \left| \frac{\det(zI_N - P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}))}{\det(zI_N - P(0_N, \dots, 0_N, (A_N^{(1)})', \dots, (A_N^{(t)})'))} \right| \\ &= \min_{z \in \partial\Gamma_l} \frac{\prod_{i=1}^{N-m_0(N)} |z - \lambda_i(M_N^{(0)})|}{|z|^{N-m_0(N)}} \\ &\geq \left(\frac{\delta}{K}\right)^{N-m_0(N)} \\ &\geq \left(\frac{\delta}{K}\right)^r. \end{aligned}$$

Then, we may apply Theorem 1.10 to each of the  $\Gamma_l$ . Since  $\delta$  can be arbitrarily small, the conclusion follows.

## References

- [1] G. W. Anderson. Convergence of the largest singular value of a polynomial in independent Wigner matrices. *Ann. Probab.*, 41(3B), 2103–2181, 2013. MR-3098069
- [2] Z. D. Bai. Circular law. *Ann. Probab.* 25, 494–529, 1997. MR-1428519
- [3] Z. Bai and J. Silverstein. *Spectral analysis of large dimensional random matrices*. Second edition. Springer Series in Statistics. Springer, New York, 2010. MR-2567175
- [4] S. T. Belinschi. Some Geometric Properties of the Subordination Function Associated to an Operator-Valued Free Convolution Semigroup. *Complex Anal. Oper. Theory* DOI 10.1007/s11785-017-0688-y. MR-3905583
- [5] S. T. Belinschi, H. Bercovici, and M. Capitaine. On the outlying eigenvalues of a polynomial in large independent random matrices. *Int. Math. Res. Notices*. <https://doi.org/10.1093/imrn/rnz080>, 2019. MR-4218332
- [6] S. T. Belinschi and M. Capitaine, Spectral properties of polynomials in independent Wigner and deterministic matrices. *Journal of Functional Analysis* 273: 3901–3963, 2017. MR-3711884
- [7] S. T. Belinschi, P. Śniady, and R. Speicher. Eigenvalues of non-Hermitian random matrices and Brown measure of non-normal operators: Hermitian reduction and linearization method. *Linear Algebra and its Applications* 537, 48–83, 2018. MR-3716236
- [8] R. Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics, 169. Springer-Verlag, New York, 1997. MR-1477662
- [9] P. Biane. Processes with free increments. *Math. Z.* 227(1), 143–174, 1998. MR-1605393
- [10] C. Bordenave and M. Capitaine. Outlier eigenvalues for deformed i.i.d. matrices. *Comm. Pure Appl. Math.*, Vol. 69, Issue 11, 2131–2194 (2016). MR-3552011
- [11] C. Bordenave and D. Chafaï. Around the circular law. *Probab. Surv.*, 9:1–89, 2012. MR-2908617
- [12] C. Bordenave, D. Chafaï and D. García-Zelada. Convergence of the spectral radius of a random matrix through its characteristic polynomial. arXiv:2012.05602.
- [13] P. Biane and F. Lehner. Computation of some examples of Brown’s spectral measure in free probability. *Colloq. Math.* 90, 181–211, 2001. MR-1876844
- [14] L. G. Brown. Lidskii’s theorem in the type II case. Geometric methods in operator algebras (Kyoto, 1983), Longman Sci. Tech., Harlow, pp. 1–35, 1986. MR-0866489
- [15] M. Capitaine. Exact separation phenomenon for the eigenvalues of large Information-Plus-Noise type matrices. Application to spiked models. *Indiana Univ. Math. J.* 63 (6), 1875–1910, 2014. MR-3298725
- [16] N. Coston, P. Wood, and S. O’Rourke. Outliers in the spectrum for products of independent random matrices. *Ann. Inst. H. Poincaré Probab. Statist.* Volume 56, Number 2, 1284–1320, 2020. MR-4076784

- [17] R.B. Dozier and J.W. Silverstein. On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices. *J. Multivariate. Anal.*, vol. 98, no. 4, 678–694, 2007. MR-2322123
- [18] D. Evans and R. Høegh-Krohn. Spectral properties of positive maps on  $C^*$ -algebras." *J. London Math. Soc. (2)* 17, 345–355, 1978. MR-0482240
- [19] J. Ginibre. Statistical Ensembles of Complex, Quaternion and Real Matrices. *J. Math. Phys.* 6, 440–449, 1965. MR-0173726
- [20] F. Götze and A.N. Tikhomirov. The Circular Law for Random Matrices. *Ann. Probab.* 38, no. 4, 1444–1491, 2010. MR-2663633
- [21] U. Haagerup and S. Thorbjørnsen. A new application of random matrices:  $Ext(C_{\text{red}}^*(\mathbb{F}_2))$  is not a group. *Ann. of Math. (2)*, 162(2), 711–775, 2005. MR-2183281
- [22] U. Haagerup, H. Schultz, and S. Thorbjørnsen. A random matrix approach to the lack of projections in  $C_{\text{red}}^*(\mathbb{F}_2)$ . *Adv. Math.*, 204(1), 1–83, 2006. MR-2233126
- [23] J. W. Helton, R. Rashidi Far, and R. Speicher. Operator-valued semicircular elements: solving a quadratic matrix equation with positivity constraints. *Int. Math. Res. Not.*, 2007. MR-2376207
- [24] T. Mai. On the Analytic Theory of Non-commutative Distributions in Free Probability, PhD thesis, Universitaät des Saarlandes, 2017, <http://scidok.sulb.uni-saarland.de/volltexte/2017/6809>.
- [25] M. L. Mehta. *Random Matrices and the Statistical Theory of Energy Levels*, Academic Press, New York, NY, 1967. MR-0220494
- [26] J. Mingo and R. Speicher. *Free Probability and Random Matrices*. Fields Institute Monographs, Volume 35, Springer, New York, 2017. MR-3585560
- [27] A. Nica, D. Shlyakhtenko, R. Speicher. Operator-valued distributions. I. Characterizations of freeness. *International Mathematics Research Notices*, (29):1509–1538, 2002. MR-1907203
- [28] G. Pan and W. Zhou. Circular law, Extreme singular values and potential theory. *J. Multivariate Anal.* 101, no. 3, 645–656, 2010. MR-2575411
- [29] S. Péché. The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields* 134, no. 1, 127–173, 2006. MR-2221787
- [30] M. Popa and V. Vinnikov. Non-commutative functions and non-commutative free Lévy-Hinçin formula. *Adv. Math.* 236, 131–157, 2013. MR-3019719
- [31] R. Rashidi Far, T. Oraby, W. Bryc, and R. Speicher. On slow-fading MIMO systems with nonseparable correlation. *IEEE Transactions on Information Theory* 54, no. 2, 544–553, 2008. MR-2444540
- [32] M. P. Schützenberger. On the definition of a family of automata. *Information and Control*, 4, 245–270, 1961. MR-0135680
- [33] P. Sniady. Random regularization of Brown spectral measure. *J. Funct. Anal.*, 193(2):291–313, 2002. MR-1929504
- [34] R. Speicher. *Combinatorial theory of the free product with amalgamation and operator-valued free probability theory*, Mem. AMS 132, no. 627, 1998. MR-1407898
- [35] T. Tao and V. Vu. Random matrices: the circular law. *Commun. Contemp. Math.* 10, 261–307, 2008. MR-2409368
- [36] T. Tao and V. Vu. From the Littlewood-Offord problem to the circular law: universality of the spectral distribution of random matrices. *Bull. Amer. Math. Soc. (N.S.)*, 46(3):377–396, 2009. MR-2507275
- [37] T. Tao and V. Vu. Random matrices: universality of ESDs and the circular law. *Ann. Probab.*, 38(5):2023–2065, 2010. With an appendix by Manjunath Krishnapur. MR-2722794
- [38] T. Tao. Outliers in the spectrum of iid matrices with bounded rank perturbations. *Probab. Theory Related Fields*, 155(1-2):231–263, 2013. MR-3010398
- [39] D. V. Voiculescu. Operations on certain non-commutative operator-valued random variables. *Astérisque* 232, 243–275, 1995. MR-1372537
- [40] D. V. Voiculescu. The coalgebra of the free difference quotient and free probability. *Internat. Math. Res. Not.* 2000, no. 2, 79–106. MR-1744647

- [41] D. V. Voiculescu. Free Analysis Questions I: Duality Transform for the Coalgebra of  $\partial_{X,B}$ . *Internat. Math. Res. Not.* 16, 793–822, 2004. MR-2036956
- [42] D. V. Voiculescu. Free analysis questions II: The Grassmannian completion and the series expansions at the origin. *J. reine angew. Math.* 645, 155–236, 2010. MR-2673426
- [43] D. V. Voiculescu, K. Dykema, and A. Nica. *Free random variables* (CRM Monograph Series, vol. 1, American Mathematical Society, Providence, RI, 1992, ISBN 0-8218-6999-X, A non-commutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups). MR-1217253

**Acknowledgments.** The authors want to thank an anonymous referee for his very careful reading and his pertinent comments.

---

# Electronic Journal of Probability

## Electronic Communications in Probability

---

### Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS<sup>1</sup>)
- Easy interface (EJMS<sup>2</sup>)

### Economical model of EJP-ECP

- Non profit, sponsored by IMS<sup>3</sup>, BS<sup>4</sup>, ProjectEuclid<sup>5</sup>
- Purely electronic

### Help keep the journal free and vigorous

- Donate to the IMS open access fund<sup>6</sup> (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

---

<sup>1</sup>LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

<sup>2</sup>EJMS: Electronic Journal Management System <http://www.vtex.lt/en/ejms.html>

<sup>3</sup>IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

<sup>4</sup>BS: Bernoulli Society <http://www.bernoulli-society.org/>

<sup>5</sup>Project Euclid: <https://projecteuclid.org/>

<sup>6</sup>IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>