



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

Serban Belinschi, Charles Bordenave, Mireille Capitaine, Guillaume Cébron
Cébron

► To cite this version:

Serban Belinschi, Charles Bordenave, Mireille Capitaine, Guillaume Cébron Cébron. Outlier eigenvalues for non-Hermitian polynomials in independent i.i.d. matrices and deterministic matrices. 2020. hal-03023644

HAL Id: hal-03023644

<https://hal.science/hal-03023644v1>

Preprint submitted on 25 Nov 2020

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

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

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

Serban Belinschi*, Charles Bordenave[†],
Mireille Capitaine[‡], Guillaume Cébron[§]

June 25, 2020

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 $\mathcal{C}^*$ -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 $P(0, A)$.

*Institut de Mathématiques de Toulouse; UMR5219; Université de Toulouse; CNRS; UPS, F-31062 Toulouse, FRANCE. E-mail: serban.belinschi@math.univ-toulouse.fr

[†]Institut de Mathématiques de Marseille; CNRS; Aix-Marseille Université, Marseille, France. E-mail: charles.bordenave@univ-amu.fr

[‡]Institut de Mathématiques de Toulouse; UMR5219; Université de Toulouse; CNRS; UPS, F-31062 Toulouse, FRANCE. E-mail: mireille.capitaine@math.univ-toulouse.fr

[§]Institut de Mathématiques de Toulouse; UMR5219; Université de Toulouse; CNRS; UPS, F-31062 Toulouse, FRANCE. E-mail: guillaume.cebron@math.univ-toulouse.fr

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 entries 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 μ_B the empirical spectral measure of B :

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

The following theorem is the culmination of the work of many authors [2, 3, 18, 19, 24, 27, 34, 36].

Theorem 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 μ_c where $d\mu_c = \frac{1}{\pi} \mathbf{1}_{|z| \leq 1} dz$.*

One can prove that when the fourth moment is finite, there are no significant outliers to the circular law.

Theorem 2. *(see Theorem 1.4 in [37]) Let X_N be an i.i.d. random matrix whose entries have finite fourth moment: $\mathbb{E}(|X_{11}|^4) < +\infty$. 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 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 M_N and A_N coincide asymptotically.

Theorem 3. *([37]) 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, [9] investigated the same problem but dealing with full rank additive perturbations. Main terminology related to free probability theory which is used in the following is defined in Section 3 below. Consider the deformed model:

$$S_N = A_N + \frac{X_N}{\sqrt{N}}, \quad (1)$$

where A_N is a $N \times N$ deterministic matrix with operator norm $O(1)$ and such that $A_N \in (\mathcal{M}_N(\mathbb{C}), \text{tr}_N)$ converges in $*$ -moments to some noncommutative random variable a in some $\mathcal{C}^*$ -probability space $(\mathcal{A}, \varphi)$. According to Dozier and Silverstein [16], for any $z \in \mathbb{C}$, almost surely the empirical spectral measure of $(S_N - zI_N)(S_N - zI_N)^*$ converges weakly towards a nonrandom distribution μ_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 4. Note that for any operator K in some $\mathcal{C}^*$ -probability space $(\mathcal{B}, \tau)$, K is invertible if and only if KK^* and K^*K are invertible. If τ is tracial, the distribution μ_{KK^*} of KK^* coincides with the distribution of K^*K . Therefore, if τ 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 φ is faithful and tracial, $\text{spect}(c + a) = \{z \in \mathbb{C} : 0 \in \text{supp}(\mu_z)\}$, where spect denotes the spectrum. Actually, we will present some results of [9] only in terms of the spectrum of $c + a$ so that we do not need the assumption (A3) in [9] on the limiting empirical spectral measure of S_N . The authors in [9] gave a sufficient condition to guarantee that outliers of the deformed model (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 $\{z \in \mathbb{C} : d(z, K) \leq \epsilon\}$.

Definition 5. Let $\Gamma \subset \mathbb{C} \setminus \text{spect}(c + a)$ be a compact set. A_N is well-conditioned in Γ 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 6. ([9]) Assume that A_N is well-conditioned in Γ , Then, a.s. for all N large enough, S_N has no eigenvalue in Γ .

Corollary 7. ([9]) 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 $\varepsilon > 0$, a.s. for all N large enough, all eigenvalues of S_N are in $B(\text{spect}(c + a), \varepsilon)$.

Let us introduce now the notion of *well-conditioned decomposition* of A_N which allowed [9] to exhibit a sufficient condition for stability of outliers.

Definition 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 , $\|A'_N\| + \|A''_N\| \leq M$.
- 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 Γ) and A''_N has a fixed rank r .

Theorem 9. ([9]) 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 $\varepsilon > 0$ and all N large enough,

$$\min_{z \in \partial\Gamma} \left| \frac{\det(A_N - z)}{\det(A'_N - z)} \right| \geq \varepsilon, \quad (2)$$

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

On the other hand, in [15], 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 three following 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 [15], 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

In this paper we generalize the previous results from [9] 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 [15]. Here are the matricial models we deal with. Let t and u be fixed nonzero integer 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 $i = 1, \dots, t$,
$$\sup_N \|A_N^{(i)}\| < \infty, \tag{3}$$

where $\|\cdot\|$ denotes the spectral norm,

2. $(A_N^{(1)}, \dots, A_N^{(t)})$ converges in $*$ -distribution towards a t -tuple $a = (a^{(1)}, \dots, a^{(t)})$ in some C^* -probability space $(\mathcal{A}, \varphi)$ where φ 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 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 23 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 μ_z where μ_z is the distribution of $[P(c, a) - z1][P(c, a) - z1]^*$. Since we can assume that φ is faithful and tracial, we have by Remark 4 that

$$\text{spect}(P(c, a)) = \{z \in \mathbb{C} : 0 \in \text{supp}(\mu_z)\}. \quad (4)$$

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 by describing the individual eigenvalues of M_N outside $B(\text{spect}(P(c, a)), \epsilon)$ for some $\epsilon > 0$. To this end, we shall fix a set $\Gamma \subset \mathbb{C}$. In the lineage of [9], 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.

Theorem 10. *Assume that hypotheses (A1), (X1) hold. Let Γ be a compact subset of $\mathbb{C} \setminus \text{spect}(P(c, a))$. Assume moreover that*

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

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

- $(\mathbf{A}'_2)$ for any z in Γ , 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 $k = 1, \dots, t$,

$$\sup_N \|(A_N^{(k)})'\| < +\infty. \quad (5)$$

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

$$\min_{z \in \partial \Gamma} \left| \frac{\det(zI_N - P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}))}{\det(zI_N - P(0_N, \dots, 0_N, (A_N^{(1)})', \dots, (A_N^{(t)})'))} \right| \geq \epsilon \quad (6)$$

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 Γ are equal.

The next statement is an easy consequence of Theorem 10.

Corollary 11. Assume that $(\mathbf{X1})$ 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 $\{z \in \mathbb{C}, \epsilon < d(z, \text{spect}(P(c, 0))) < 3\epsilon\}$, 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, 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 $\{z \in \mathbb{C}, d(z, \text{spect}(P(c, 0))) \geq 2\epsilon\}$ and after labeling these eigenvalues properly,

$$\max_{j \in J} |\lambda_j(M_N) - \lambda_j(M_N^{(0)})| \rightarrow_{N \rightarrow +\infty} 0.$$

We will first prove Theorem 10 in the case $r = 0$.

Theorem 12. Suppose that assumptions of Theorem 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 Γ .

In particular, if assumptions of Theorem 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 $\varepsilon > 0$, a.s. for all N large enough, all eigenvalues of M_N are in $B(\text{spect}(P(c, a)), \varepsilon)$.

To prove Theorems 12 and 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 [9]. 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 12 and 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 23) 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 [13]). 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 three following 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 [32] 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 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 [12]).

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 [12, 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 [12]. This disk is also the spectrum of $\frac{3}{2}c$. Our theorem says that, outside this disk, the outliers of M_N are closed 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).

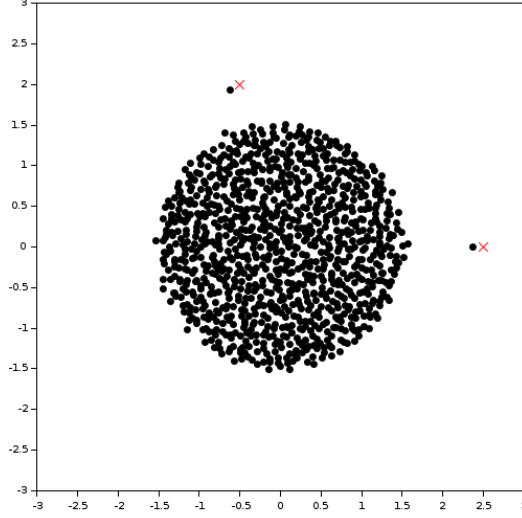


Figure 1: In black, the eigenvalues of $P_1 \left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N \right)$ for $N = 1000$, and in red, the outliers 2.5 and $2i - 0.5$ of $P_1(0_N, 0_N, 0_N, A_N)$.

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,

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

10

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

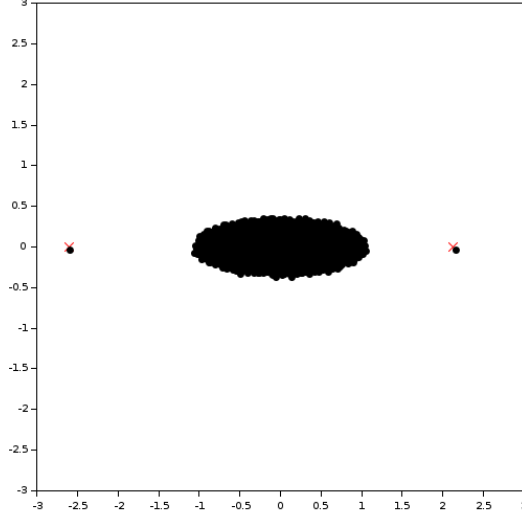


Figure 2: In black, the eigenvalues of $P_2 \left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N^{(1)}, A_N^{(2)} \right)$ for $N = 1000$, and in red, the limiting outliers 2.125 and -2.6 of $P_1(0, 0, 0, A_N^{(1)}, A_N^{(2)})$.

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 [12]. 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 [28].

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 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 [12]. The interior of this ellipse is also the spectrum of $c + a$. Our theorem says that, outside this ellipse, 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).

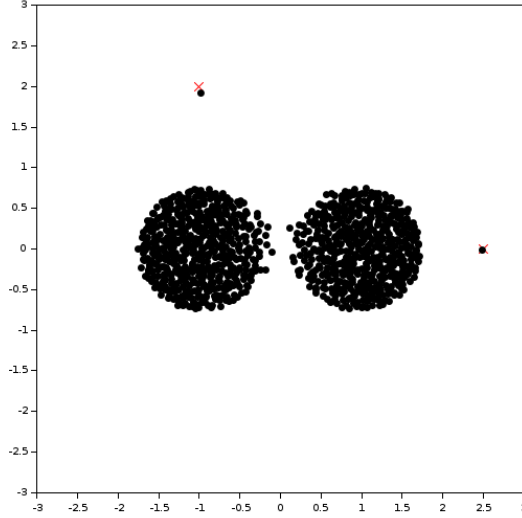


Figure 3: In black, the eigenvalues of $P_3 \left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N^{(1)}, A_N^{(2)} \right)$ for $N = 1000$, and in red, the outliers 2.5 and $-1 + 2i$ of $P_3(0_N, 0_N, 0_N, A_N^{(1)}, A_N^{(2)})$.

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 ex-

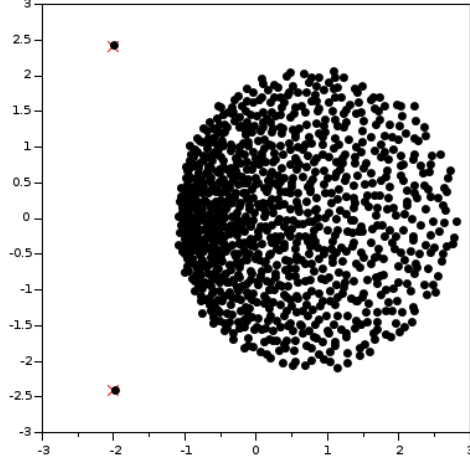


Figure 4: In black, the eigenvalues of $P_4 \left(\frac{X_N^{(1)}}{\sqrt{N}}, \frac{X_N^{(2)}}{\sqrt{N}}, \frac{X_N^{(3)}}{\sqrt{N}}, A_N \right)$ for $N = 1000$, and in red, the outliers $-2 + 2.4i$ and $-2 - 2.4i$ of $P_4(0_N, 0_N, 0_N, A_N)$.

pected (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).

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 [42].

- A $\mathcal{C}^*$ -probability space is a pair $(\mathcal{A}, \varphi)$ consisting of a unital $\mathcal{C}^*$ -algebra $\mathcal{A}$ and a state φ 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 $a \in \mathcal{A}$) φ 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 $\mathcal{C}^*$ -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 ν_{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 μ_{a_1} and ν_{a_1} . If φ is faithful then the support of ν_{a_1} is the spectrum of a_1 and thus $\|a_1\| = \sup\{|z|, z \in \text{support}(\nu_{a_1})\}$.

- A family of elements $(a_i)_{i \in I}$ in a $\mathcal{C}^*$ -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

$$\varphi(p_1(a_{i_1}, a_{i_1}^*) \cdots p_k(a_{i_k}, a_{i_k}^*)) = 0 \quad (7)$$

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 $\mathcal{C}^*$ -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, respectively 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 [38] 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 $E[1] = 1$. 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 μ_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 μ_X - see [41]; since we do not need this extension, we shall not discuss it any further, but refer the reader to [41, 39, 40, 29] for details.) A natural domain for G_X is the upper half-plane of B , $H^+(B) = \{b \in B : \Im b > 0\}$. It follows quite easily that $G_X(H^+(B)) \subseteq H^+(B)$ - see [40].

We warn here the reader that we have changed conventions in our paper compared to [39, 40, 41], 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 [38], 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 [38] 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 [33, 41]). For our purposes, we use the equivalent description of an operator-valued semicircular distribution via its noncommutative Cauchy transform, as in [22]: 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 $\eta(b) = E[SbS] - E[S]bE[S]$. 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 σ^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 13. 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

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

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 [30], and its proof can be found in great detail in [25].

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 [38]). 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

$\begin{bmatrix} s_{1,1}^{(2)} & is_{1,2}^{(2)} \\ -is_{1,2}^{(2)} & s_{2,2}^{(2)} \end{bmatrix}$ are $\mathcal{M}_2(\mathbb{C})$ -valued semicirculars which are free over $\mathcal{M}_2(\mathbb{C})$, so their sum $\begin{bmatrix} s_{1,1}^{(1)} + s_{1,1}^{(2)} & s_{1,2}^{(1)} + is_{1,2}^{(2)} \\ s_{1,2}^{(1)} - is_{1,2}^{(2)} & s_{2,2}^{(1)} + s_{2,2}^{(2)} \end{bmatrix}$ 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 $s_{1,2}^{(1)} \pm is_{1,2}^{(2)}$, 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 14. 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 15. 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 $i = 1, \dots, u$,

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

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

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

Then, in the scalar-valued probability space $(\mathcal{M}_2(\mathcal{A}), \text{tr}_2 \otimes \varphi)$, $\mathbf{s}_1, \dots, \mathbf{s}_u, \{\mathbf{m}_j, j = 1, \dots, t\}$ are free and for $i = 1, \dots, u$, each $\mathbf{s}_i$ is a semicircular variable.

Proof. Let us prove that $\mathbf{s}_1, \dots, \mathbf{s}_u$ 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 [25, Chapter 9]) that $\mathbf{s}_1, \dots, \mathbf{s}_u$ 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 $\mathbf{s}_1, \dots, \mathbf{s}_u$ 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

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

Using [26, Theorem 3.5], the freeness of $\mathbf{s}_1, \dots, \mathbf{s}_u$ from $\mathcal{M}_2(\mathcal{B})$ over $\mathcal{M}_2(\mathbb{C})$ gives us the free cumulants of $\mathbf{s}_1, \dots, \mathbf{s}_u$ over $\mathcal{M}_2(\mathcal{B})$. More concretely, we get that $\mathbf{s}_1, \dots, \mathbf{s}_u$ 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 [26, Theorem 3.5], $\mathbf{s}_1, \dots, \mathbf{s}_u$ 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 $\mathbf{s}_1, \dots, \mathbf{s}_u$ are free standard semicircular variables. $\square$

Lemma 16. *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 operator-valued probability space $(\mathcal{M}_m(\mathcal{A}), id_m \otimes \varphi)$, $|\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y|^2$ has the same distribution as $|\sum_{j=1}^u \zeta_j \otimes s_j + (I_m \otimes \epsilon) \cdot y|^2$ where ϵ 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 ϵ and the entries of y .*

In the lemma above, we consider the symmetric version ϵy of y , thanks to a noncommutative random variable ϵ which is *tensor-independent* from the entries of y , in the sense that ϵ 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 [25, Chapter 9]). Let us fix $S = \{j : i_j \neq 0\} \subset$

$\{1, \dots, 2n\}$ and use the moment cumulant formula (see [33, 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 π^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 [33, 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 π like $\hat{c}$ and on the blocks of π^c like $\hat{\varphi}$.

Recall that the cumulants of $\zeta_j \otimes c^{(j)}$ are vanishing if π is not a pairing and if π is not alternating (which means that π links two indices with the same parity). Now, let us remark that if π is a pairing which is alternating, then π^c is even (each blocs of π^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 π is not a pairing and that the moment of b_0 is vanishing if π^c is odd. Moreover, if π is a pairing and π^c is even, then π 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 ϵy and y has the same even $\mathcal{M}_m(\mathbb{C})$ -valued moments and $\zeta_j \otimes c^{(j)}$ and $\zeta_j \otimes s^{(j)}$ has the same alternating $\mathcal{M}_m(\mathbb{C})$ -valued cumulants. $\square$

It follows from [4] that the support in $\mathcal{M}_m(\mathbb{C})^{\text{sa}}$ of the addition of a semi-circular s of variance η 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

$$H(w) = w - \eta(G_y(w)) \text{ and } \omega(b) = b + \eta(G_y(\omega(b))), \quad (8)$$

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

Proposition 17. *If $w \in \mathcal{M}_m(\mathbb{C})^{\text{sa}}$ 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 $b \in \mathcal{M}_m(\mathbb{C})^{\text{sa}}$ 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 ω 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 [17, Theorem 2.5] that the spectral radius r of $\eta \circ G'_y(w)$ is reached at a positive element $\xi \in M_m(\mathbb{C})$, 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 $H(w)$, $H'(w) = Id_{M_m(\mathbb{C})} - \eta \circ G'_y(w)$, 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 ω coincides with the local inverse of H on the upper half-plane, so the local inverse of H is the unique

analytic continuation of ω to a neighborhood of $H(w)$. This proves that ω extends analytically to a neighborhood of $H(w)$ and the extension maps selfadjoints from this neighborhood to $\mathcal{M}_m(\mathbb{C})^{\text{sa}}$. 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 $\mathcal{M}_m(\mathbb{C})^{\text{sa}}$. Since $\omega(b) = b + \eta(G_{s+y}(b))$, the same holds for ω . 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 ω around $b = b^*$ implies $\text{spect}(\omega'(0)) \subset \{\Re z \geq 1/2\}$. Thus, ω is invertible wrt composition around zero 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 18. *The proof of the previous proposition, based on [17, 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 19. *Consider the operator-valued $\mathcal{C}^*$ -algebraic noncommutative probability space $(\mathcal{M}_m(\mathcal{A}), \text{id}_m \otimes \varphi, \mathcal{M}_m(\mathbb{C}))$ 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 $\text{id}_m \otimes \varphi$. 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 $G_x(b) = \text{id}_m \otimes \varphi[(x - b)^{-1}]$. 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 $H(i\mathcal{M}_m(\mathbb{C})^+) \subseteq i\mathcal{M}_m(\mathbb{C})^{\text{sa}}$ and $\omega(i\mathcal{M}_m(\mathbb{C})^+) \cup G_y(i\mathcal{M}_m(\mathbb{C})^+) \cup G_{y+s}(i\mathcal{M}_m(\mathbb{C})^+) \subseteq i\mathcal{M}_m(\mathbb{C})^+$.

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 17 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 self-adjoints to selfadjoints. Since $\omega(b) = b + \eta(G_{s+y}(b))$, it follows that ω does the same. According to Proposition 17, $y - \omega(0)$ 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, [31]), 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 [20, 21] (see [25]). 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 $(X_{i_1} X_{i_2} \dots X_{i_l})^* = X_{i_l}^* \dots X_{i_2}^* X_{i_1}^*$, $(i_1, \dots, i_l) \in \{1, \dots, k\}^l$, $l \in \mathbb{N} \setminus \{0\}$. 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 ≤ 1 ,

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 20. [25, 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*

$$(ze_{11} \otimes 1_{\mathcal{A}} - L_P(y))^{-1} = \begin{pmatrix} (z1_{\mathcal{A}} - P(y))^{-1} & \star \\ \star & \star \end{pmatrix}. \quad (9)$$

Lemma 21. *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,*

$$\begin{aligned} \|(ze_{11} \otimes 1_{\mathcal{A}} - L_P(y))^{-1}\| &\leq T_1(\|y_1\|, \dots, \|y_k\|) \times \|(z1_{\mathcal{A}} - P(y))^{-1}\| \\ &\quad + T_2(\|y_1\|, \dots, \|y_k\|). \end{aligned} \quad (10)$$

Proof. 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

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

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

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,

$$\det Q(y) = \pm 1. \quad (11)$$

5 No outlier; proof of Theorem 12

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 for the first assertion of Theorem 12 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 (4), $\text{spect}(P(c, a)) = \{z \in \mathbb{C} : 0 \in \text{supp}(\mu_z)\}$, where μ_z is the distribution of $(P(c, a) - z)(P(c, a) - z)^*$. The first assertion of Theorem 12 is equivalent to the following.

Proposition 22. *Let Γ be a compact set of $\{z, 0 \notin \text{supp}(\mu_z)\}$; assume that for any z in Γ , 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 Γ , 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 22 is based on the two following key results.

Proposition 23. 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 (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 $(\mathcal{M}_N(\mathbb{C}), \frac{1}{N} \text{Tr})$ coincides with the $*$ -distribution of $a_N = (a_N^{(j)}, j = 1, \dots, t)$ in $(\mathcal{A}, \varphi)$. Let τ_N be the the distribution of

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

with respect to φ . If $[x, y]$, $x < y$, is such that there exists a $\delta > 0$ 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 24. 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 $\tau > 0$ 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 δ_{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 $|P(0, a_N) - z_0|^2 \geq \delta_{z_0}$.

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 25. *Of course Proposition 24 still holds dealing with nonselfadjoint tuples a_N by considering the selfadjoint tuples $(\Im(a_N), \Re(a_N))$.*

Let us explain how to deduce Theorem 10 from Proposition 23 and Proposition 24.

Define $\mu_{N,z}$ as the distribution of

$$\left[P(c^{(1)}, \dots, c^{(u)}, a_N^{(1)}, \dots, a_N^{(t)}) - z1 \right] \times \left[P(c^{(1)}, \dots, c^{(u)}, a_N^{(1)}, \dots, a_N^{(t)}) - z1 \right]^*$$

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 24.

Corollary 26. *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 26 and Proposition 23 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 $z \mapsto s_N(M_N - zI)$ 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 ,

$$\inf_{z \in \Gamma} s_N(M_N - zI_N) \geq \gamma_\Gamma, \tag{12}$$

leading to Proposition 22.

5.2 Proof of Proposition 23

Note that

$$\begin{pmatrix} K_N K_N^* & 0 \\ 0 & K_N^* K_N \end{pmatrix} = \begin{pmatrix} 0 & K_N \\ K_N^* & 0 \end{pmatrix}^2, \tag{28}$$

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

$$\begin{aligned} \begin{pmatrix} 0 & K \\ K^* & 0 \end{pmatrix} &= \sum_{i=1}^p \begin{pmatrix} 0 & b_i m_i \\ \bar{b}_i m_i^* & 0 \end{pmatrix} \\ &= \sum_{i=1}^p b_i \begin{pmatrix} 0 & m_i \\ m_i^* & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \bar{b}_i \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & m_i \\ m_i^* & 0 \end{pmatrix} \end{aligned} \quad (13)$$

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 standard Wigner matrices. Now, note that as noticed by [7] for any monomial $x_1 \cdots x_k$,

$$\begin{pmatrix} 0 & x_1 \cdots x_k \\ (x_1 \cdots x_k)^* & 0 \end{pmatrix} = \Pi_{k-1} \begin{pmatrix} 0 & x_k \\ x_k^* & 0 \end{pmatrix} \Pi_{k-1}^* \quad (14)$$

where

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

Indeed, this can be proved by induction noting that

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

Note also that

$$S \begin{pmatrix} 0 & I \\ x_1^* & 0 \end{pmatrix} S = \begin{pmatrix} 0 & x_1^* \\ I & 0 \end{pmatrix}. \quad (15)$$

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

From (13), (14), (15), 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 $i = 1, \dots, u$,

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

Similarly,

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

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

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 15, it turns out that the $\mathbf{s}_i$'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 23 follows by applying [6, Theorem 1.1. and Remark 4]. The second assertion of Proposition 23 can be proven by the same previous arguments. Indeed, there exists a polynomial $\tilde{K}$ such that

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

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

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

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

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

The second assertion of Proposition 23 follows.

5.3 Proof of Proposition 24

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 $P(Y_1, Y_2) - z_0$.

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

$$\left| \begin{bmatrix} 0 & P(c, a_N) - z_0 \\ (P(c, a_N) - z_0)^* & 0 \end{bmatrix} \right|_{31} \geq \epsilon_{z_0} \iff \left| \begin{bmatrix} 0 & L(c, a_N) \\ L(c, a_N)^* & 0 \end{bmatrix} \right| \geq \iota.$$

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 16 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 ϵ is a $\{-1, 1\}$ -Bernoulli distributed random variable which is independent from a_N and free from $\mathcal{S}$. As noted in Example 13, since $C = C^*$, $\mathcal{S}$ is indeed a matrix-valued semicircular random variable.

5. We apply Lemma 19 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

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

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 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 $\ell \times \ell$ corner of this matrix has an inverse of degree ℓ in the algebra $\mathcal{M}_\ell(\mathbb{C}\langle X_1, \dots, X_k \rangle)$. (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,

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

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.

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 non-selfadjoint (and thus necessarily weaker) version of [5, Lemma 4.3].

Lemma 27. *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 ϵ 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 ϵ , 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 δ , 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 $\mathfrak{e}$ (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 $\mathfrak{e}$ depending on δ , P , and the norms of the entries of P .

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

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

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

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

This concludes the proof. $\square$

Part (3) is a simple formal step.

Step (4) becomes a direct consequence of Lemma 16.

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 , δ_{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 , δ_{z_0} , and the norms of c, a . According to Lemma 19 and Remark 18, 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 $\text{spect}(\eta \circ G'_{\mathcal{Y}}(0)) \subset (1 - r)\mathbb{D}$. 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 ζ that only depends on P, δ_{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 $G'_{\mathcal{Y}}(0)$ (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 $(1 - \frac{\epsilon}{2})\overline{\mathbb{D}}$. 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 17 shows, the endpoint of the support of the (scalar) distribution of $\mathcal{S} + \mathcal{Y}_N$ is given by that first $x_N \in (0, +\infty)$ for which $1 \in \text{spect}(\eta \circ G'_{\mathcal{Y}_N}(x_N))$. 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 $G_{\mathcal{Y}_N} \rightarrow G_{\mathcal{Y}}$ uniformly on $[0, \delta_{z_0} - \epsilon]$ for any fixed $\epsilon > 0$. In particular, $G'_{\mathcal{Y}_N}(x) \rightarrow G'_{\mathcal{Y}}(x)$ 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 10

Making use of a linearization procedure, the proof closely follows the approach of [9]. The most significant novelty is Proposition 28 which substantially generalizes Theorem 1.3. A. in [14] (see also Proposition 2.1 in [9]) and whose proof relies on operator-valued free probability results established in Section 3.2.2. Nevertheless, we precise 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 (11)).

Let Γ be a compact set in $\mathbb{C} \setminus \text{spect}(P(c, a))$. 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

$$\begin{aligned} \min_{z \in \partial\Gamma} \left| \frac{\det(zI_N - P(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}))}{\det(zI_N - P(0_N, \dots, 0_N, (A_N^{(1)})', \dots, (A_N^{(t)})'))} \right| & \quad (16) \\ & \times \frac{\det(Q(0_N, \dots, 0_N, A_N^{(1)}, \dots, A_N^{(t)}))}{\det Q(0_N, \dots, 0_N, (A_N^{(1)})', \dots, (A_N^{(t)})')} \Big| \geq \epsilon, \end{aligned}$$

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

$$\min_{z \in \partial\Gamma} \left| \frac{\det(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes A_N^{(k)})}{\det(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})')} \right| \geq \epsilon. \quad (17)$$

According to Lemma 20, 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 $(\mathbf{A}'_2)$, Proposition 22 and Lemma 20, almost surely for all large N , for any $z \in \Gamma$, we can define

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

Note that, 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

$$\sum_{k=1}^t \beta_k \otimes (A_N^{(k)})'' = P_N Q_N. \quad (18)$$

Recall Sylvester's identity: if $P, Q^\top \in \mathcal{M}_{d_1, d_2}(\mathbb{C})$,

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

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

Now, similarly, for any z in Γ ,

$$\det(I_p - Q_N R'_N(z) P_N) = \frac{\det(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes A_N^{(k)})}{\det(ze_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})')}. \quad (19)$$

Thus, the zeroes of $z \mapsto \det(I_p - Q_N R'_N(z) P_N)$ in Γ 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 Γ , that is, the eigenvalues in Γ 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 Γ 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 $K \geq 2$,

$$\begin{aligned} & Q_N R_N(z) P_N - Q_N R'_N(z) P_N \\ &= \sum_{k=1}^{K-1} Q_N (R'_N(z) Y_N)^k R'_N(z) P_N + Q_N (R'_N(z) Y_N)^K R_N(z) P_N. \end{aligned} \quad (20)$$

The following two steps will be of basic use to prove the uniform convergence in Γ of the right hand side of (20) 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 32 which establishes an upper bound strictly smaller than 1 of the spectral norm of $R'_N(z)Y_N$. The proof of Lemma 32 is based on Proposition 22 and the characterization, provided by Lemma 19, 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 μ_z is the distribution of

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

Define ν_z as the distribution of

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

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

Proposition 28. *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 ϵ 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 ϵ and the entries of y_z . Define

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

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

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

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 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 20, 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 $\left[\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y_z \right] \left[\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y_z \right]^*$ and $\left[\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y_z \right]^* \left[\sum_{j=1}^u \zeta_j \otimes c^{(j)} + y_z \right]$ are invertible, and then, by Lemma 16, 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 28 follows from Example 13 and Lemma 19. $\square$

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

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

Lemma 29. $0 \notin \text{supp}(\mu_{w,z})$ 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 28.

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

$$[P(|w|c^{(1)}, \dots, |w|c^{(u)}, a^{(1)}, \dots, a^{(t)}) - zI] [P(|w|c^{(1)}, \dots, |w|c^{(u)}, a^{(1)}, \dots, a^{(t)}) - zI]^*.$$

Then the result follows from Proposition 28. $\square$

Lemma 30. Let Γ 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 Γ . According to Proposition 28, $z \notin S_0$ and $\text{spect}(\Delta_1(z)) \subseteq \overline{\mathbb{D}} \setminus \{1\}$. According to [17, 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 ξ in $\mathcal{M}_m(\mathbb{C})$ 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 4 and Lemma 20, 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 29, $0 \notin \text{supp}(\mu_{w,z})$. $\square$

Lemma 31. *Let Γ 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 ρ is defined in Lemma 30. According to Lemma 30, $\forall (w, z) \in \tilde{\Gamma}$, $0 \notin \text{supp}(\mu_{w,z})$. Therefore, using $(\mathbf{A}'_2)$, according to Proposition 22, 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 (5)). $\square$

Lemma 32. *Let Γ be a compact subset in $\{z \in \mathbb{C}: 0 \notin \text{supp}(\mu_z)\}$. Assume that $(\mathbf{A}'_2)$ and (5) hold. There exists $0 < \epsilon_0 < 1$ such that almost surely for all large N , we have for any z in Γ ,*

$$\rho(R'_N(z)Y_N) \leq 1 - \epsilon_0,$$

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 31, we can deduce that almost surely for all large N , the nonnul 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 33. *Let Γ be a compact subset in $\{z \in \mathbb{C}, 0 \notin \text{supp}(\mu_z)\}$. Assume that $(\mathbf{A}'_2)$ and (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 ϵ_0 be as defined by Lemma 32 and ρ be as defined in Lemma 31. Choose $0 < \epsilon < \min(\epsilon_0, 1 - \frac{1}{\rho})$. Therefore, according to Lemma 32 and using Dunford-Riesz calculus, we have almost surely for all large N , for any z in Γ ,

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

and therefore

$$\forall k \geq 0, \|(T_N(z))^k\| \leq \sup_{|w|=1-\epsilon} \|(w - T_N(z))^{-1}\| (1 - \epsilon)^{k+1}. \quad (21)$$

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

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

$$\times \frac{1}{w} \left(z e_{11} \otimes I_N - \gamma \otimes I_N - \sum_{k=1}^t \beta_k \otimes (A_N^{(k)})' \right). \quad (22)$$

Lemma 31 readily implies that almost surely for all large N ,

$$\left\| \left(z I_N - P(w^{-1} \frac{X_N^{(1)}}{\sqrt{N}}, \dots, w^{-1} \frac{X_N^{(u)}}{\sqrt{N}}, (A_N^{(1)})', \dots, (A_N^{(t)})') \right)^{-1} \right\| \leq 1/\eta, \quad (23)$$

where η is defined in Lemma 31.

It readily follows from (22), Lemma 21, (23), (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 Γ ,

$$\sup_{|w|=1-\epsilon} \|(w - T_N(z))^{-1}\| \leq C. \quad (24)$$

Proposition 33 follows from (21) and (24). $\square$

Step 4: Conclusion.

We will use the following proposition from [9] to establish Lemma 35 below.

Proposition 34 ([9]). *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 iid 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 35. *Assume $(\mathbf{X}_1)$, (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 towards zero when N goes to infinity, where P_N and Q_N are defined by (18).*

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 (3) and (5), the s_i 's are uniformly bounded. Using

($\mathbf{A}_2'$), (5) and (10), almost surely for any z in Γ , there exists $\tilde{\eta}_z > 0$ such that for all large N ,

$$\|R'_N(z)\| \leq \frac{1}{\tilde{\eta}_z}. \quad (25)$$

Using (25) and Bai-Yin's theorem, we deduce from Proposition 34 that $v_i^*(R'_N(z)Y_N)^k R'_N(z)u_j$ converges almost surely to zero. The result follows by applying the dominated convergence theorem thanks to Proposition 33. $\square$

We are going to prove that, assuming ($\mathbf{X}_1$), (3) and ($\mathbf{A}_2$), we have for any z in Γ , almost surely, as $N \rightarrow \infty$,

$$\|Q_N R_N(z)P_N - Q_N R'_N(z)P_N\| \rightarrow 0. \quad (26)$$

Let $C' > 0$ such that $\|P_N\| \|Q_N\| \leq C'$. According to Proposition 22 and (10), for any $z \in \Gamma$, there exists $\tilde{\gamma}_z > 0$ such that almost surely for all large N

$$\|R_N(z)\| \leq \frac{1}{\tilde{\gamma}_z}. \quad (27)$$

Then using also Proposition 33 and (25), for any $k \geq 1$, we have

$$\left\| Q_N (R'_N(z)Y_N)^k R'_N(z)P_N \right\| \leq \frac{CC'}{\tilde{\eta}_z} (1 - \epsilon_0)^k,$$

$$\left\| Q_N (R'_N(z)Y_N)^k R_N(z)P_N \right\| \leq \frac{CC'}{\tilde{\gamma}_z} (1 - \epsilon_0)^k.$$

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 (20), 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 η going to zero, that we have

$$Q_N R_N(z)P_N - Q_N R'_N(z)P_N = \sum_{k \geq 1} Q_N (R'_N(z)Y_N)^k R'_N(z)P_N. \quad (28)$$

Applying Lemma 35, we obtain (26).

Proposition 36. *Let Γ be a compact subset of $\mathbb{C} \setminus \text{spect}(P(c, a))$. Assume $(\mathbf{X}_1)$, (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 Γ , when N goes to infinity.*

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

$$\sup_{z \in \Gamma} \|Q_N R_N(z) P_N - Q_N R'_N(z) P_N\| \leq 3\delta. \quad (29)$$

We set $\zeta_z = \tilde{\eta}_z \wedge \tilde{\gamma}_z$ and $r_z = (\zeta_z/2) \wedge (\delta(\zeta_z^2/2C'))$. Using the resolvent identity, (25) and (27), 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 Γ compact, there is a finite covering and the proposition follows from (26). $\square$

Theorem 10 follows from Proposition 36 by Rouché's Theorem, using (19) and (17).

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.
- [2] Z. D. Bai. Circular law. *Ann. Probab.* 25, 494–529, 1997.
- [3] Z. Bai and J. Silverstein. *Spectral analysis of large dimensional random matrices*. Second edition. Springer Series in Statistics. Springer, New York, 2010.
- [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
- [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).

- [6] S.T. Belinschi and M. Capitaine, “Spectral properties of polynomials in independent Wigner and deterministic matrices.” *Journal of Functional Analysis* 273 (2017): 3901–3963.
- [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 (2018) 48–83.
- [8] R. Bhatia, *Matrix Analysis*. Graduate Texts in Mathematics, 169. Springer-Verlag, New York, 1997.
- [9] C. Bordenave and M. Capitaine, “Outlier eigenvalues for deformed i.i.d. matrices.” *Comm. Pure Appl. Math.* , Vol. 69, Issue 11, 2131-2194 (2016).
- [10] Philippe Biane, “Processes with free increments. *Math. Z.* 227(1), 143–174 (1998).
- [11] C. Bordenave and D. Chafaï. “Around the circular law.” *Probab. Surv.*, 9:1–89, 2012.
- [12] P. Biane, F. Lehner. “Computation of some examples of Brown’s spectral measure in free probability.” *Colloq. Math.* 90, 181–211 (2001).
- [13] 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.
- [14] 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.
- [15] N. Coston, P. Wood, and S. O’Rourke. “Outliers in the spectrum for products of independent random matrices.” arxiv:1711.07420.
- [16] 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.
- [17] D. Evans and R. Høegh-Krohn, “Spectral properties of positive maps on C^* -algebras.” *J. London Math. Soc. (2)* 17, (1978): 345–355.

- [18] J. Ginibre, “Statistical Ensembles of Complex, Quaternion and Real Matrices,” *J.Math. Phys.* 6, 440–449, 1965.
- [19] F. Götze and A.N. Tikhomirov. “The Circular Law for Random Matrices.” *Ann. Probab.* 38, no. 4, 1444–1491, 2010.
- [20] 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.
- [21] 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.
- [22] J.W. Helton, R. Rashidi Far, and R. Speicher, “Operator-valued semi-circular elements: solving a quadratic matrix equation with positivity constraints. *Int. Math. Res. Not.* (2007).
- [23] 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>.
- [24] M.L. Mehta. *Random Matrices and the Statistical Theory of Energy Levels*, Academic Press, New York, NY, 1967.
- [25] J. Mingo and R. Speicher, *Free Probability and Random Matrices*. Fields Institute Monographs, Volume 35, Springer, New York (2017).
- [26] A. Nica, D. Shlyakhtenko, R. Speicher, “Operator-valued distributions. I. Characterizations of freeness.” *International Mathematics Research Notices*, 2002(29):1509–1538.
- [27] G. Pan and W. Zhou, Circular law, Extreme singular values and potential theory. *J. Multivariate Anal.* 101, no. 3, 645–656, 2010.
- [28] S. Péché. “The largest eigenvalue of small rank perturbations of Hermitian random matrices.” *Probab. Theory Related Fields* 134, no. 1, 127–173. (2006)
- [29] M. Popa and V. Vinnikov, “Non-commutative functions and non-commutative free Lévy-Hinčin formula”. *Adv. Math.* 236, (2013) 131–157

- [30] 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 (2008): 544–553.
- [31] M. P. Schützenberger. “On the definition of a family of automata.” *Information and Control*, 4, (1961) 245–270.
- [32] P. Sniady. “Random regularization of Brown spectral measure.” *J. Funct. Anal.*, 193(2):291–313, 2002.
- [33] R. Speicher, *Combinatorial theory of the free product with amalgamation and operator-valued free probability theory*, Mem. AMS 132 (1998), no. 627.
- [34] T. Tao and V. Vu. “Random matrices: the circular law.” *Commun. Contemp. Math.* 10, 261–307, 2008.
- [35] 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.
- [36] 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.
- [37] T. Tao. “Outliers in the spectrum of iid matrices with bounded rank perturbations.” *Probab. Theory Related Fields*, 155(1-2):231–263, 2013.
- [38] D. V. Voiculescu, “Operations on certain non-commutative operator-valued random variables.” *Astérisque* 232 (1995), 243–275.
- [39] D. V. Voiculescu, “The coalgebra of the free difference quotient and free probability”. *Internat. Math. Res. Not.* 2000 (2000), no. 2, 79–106.
- [40] D. V. Voiculescu, “Free Analysis Questions I: Duality Transform for the Coalgebra of $\partial_{X:B}$ ”. *Internat. Math. Res. Not.* 16 (2004), 793–822.
- [41] D. V. Voiculescu, “Free analysis questions II: The Grassmannian completion and the series expansions at the origin”. *J. reine angew. Math.* 645 (2010), 155–236.

- [42] 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 noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups).