Free Probability, Random Matrices, and Non-Commutative Rational Functions

Roland Speicher

Saarland University
Saarbrücken, Germany

supported by ERC Advanced Grant
"Non-Commutative Distributions in Free Probability"

(joint work with S. Belinschi, J. W. Helton, and T. Mai)

European Research Council
Established by the European Commission

Section 1

Introduction

Goal: We want to understand distributions of functions in non-commuting variables

Goal: We want to understand distributions of functions in non-commuting variables

which "non-commuting variables"

- (random) matrices of size $N \times N$
- operators on Hilbert spaces (corresponding to $N \to \infty$)

Goal: We want to understand distributions of functions in non-commuting variables

which "non-commuting variables"

- (random) matrices of size $N \times N$
- operators on Hilbert spaces (corresponding to $N \to \infty$)

which "functions"

- non-commutative polynomials
- non-commutative rational functions
- (maybe even: non-commutative analytic functions)

European Research Council
Established by the European Commission

Goal: We want to understand distributions of functions in non-commuting variables

which "non-commuting variables"

- (random) matrices of size $N \times N$
- operators on Hilbert spaces (corresponding to $N \to \infty$)

which "functions"

- non-commutative polynomials
- non-commutative rational functions
- (maybe even: non-commutative analytic functions)

what means "distribution"

- algebraic/combinatorial distribution: collection of moments
- analytic distribution: probability measure

Non-commutative probability spaces

Definition

A non-commutative probability space $(\mathcal{A}, \varphi)$ consists of

- a complex algebra $\mathcal{A}$ with unit $1_{\mathcal{A}}$ and
- a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ satisfying $\varphi(1_{\mathcal{A}}) = 1$ (expectation).

Elements $x \in \mathcal{A}$ are called non-commutative random variables.

European Research Council
Established by the European Commission

Non-commutative probability spaces

Definition

A non-commutative probability space $(\mathcal{A}, \varphi)$ consists of

- a complex algebra $\mathcal{A}$ with unit $1_\mathcal{A}$ and
- a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ satisfying $\varphi(1_\mathcal{A}) = 1$ (expectation).

Elements $x \in \mathcal{A}$ are called non-commutative random variables.

Example

- $(L^\infty(\Omega, \mathbb{P}), \mathbb{E})$, where $(\Omega, \Sigma, \mathbb{P})$ is a classical probability space and $\mathbb{E}$ the usual expectation that is given by $\mathbb{E}[X] = \int_\Omega X(\omega)\, d\mathbb{P}(\omega)$

Non-commutative probability spaces

Definition

A non-commutative probability space $(\mathcal{A}, \varphi)$ consists of

- a complex algebra $\mathcal{A}$ with unit $1_{\mathcal{A}}$ and
- a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ satisfying $\varphi(1_{\mathcal{A}}) = 1$ (expectation).

Elements $x \in \mathcal{A}$ are called non-commutative random variables.

Example

- $(L^\infty(\Omega, \mathbb{P}), \mathbb{E})$, where $(\Omega, \Sigma, \mathbb{P})$ is a classical probability space and $\mathbb{E}$ the usual expectation that is given by $\mathbb{E}[X] = \int_\Omega X(\omega)\, d\mathbb{P}(\omega)$
- $(M_N(\mathbb{C}), \mathrm{tr}_N)$, where tr_N is the normalized trace on $M_N(\mathbb{C})$

European Research Council
Established by the European Commission

Non-commutative probability spaces

Definition

A non-commutative probability space $(\mathcal{A}, \varphi)$ consists of

- a complex algebra $\mathcal{A}$ with unit $1_{\mathcal{A}}$ and
- a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ satisfying $\varphi(1_{\mathcal{A}}) = 1$ (expectation).

Elements $x \in \mathcal{A}$ are called non-commutative random variables.

Example

- random matrices $((L^\infty(\Omega) \otimes M_n(\mathbb{C}), \mathbb{E} \otimes \mathrm{tr}_n)$

European Research Council
Established by the European Commission

Non-commutative probability spaces

Definition

A non-commutative probability space $(\mathcal{A}, \varphi)$ consists of

- a complex algebra $\mathcal{A}$ with unit $1_{\mathcal{A}}$ and
- a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ satisfying $\varphi(1_{\mathcal{A}}) = 1$ (expectation).

Elements $x \in \mathcal{A}$ are called non-commutative random variables.

Example

- random matrices $((L^{\infty}(\Omega) \otimes M_n(\mathbb{C}), \mathbb{E} \otimes \mathrm{tr}_n)$

We call $(\mathcal{A}, \varphi)$ a C^*-probability space if

- $\mathcal{A}$ is a unital C^*-algebra;
  i.e., $\mathcal{A}$ consists of bounded operators on a Hilbert space $\mathcal{H}$
- φ is positive (i.e. $\varphi(x^*x) \geq 0$ for $x \in \mathcal{A}$) and hence a state on $\mathcal{A}$;
  i.e., φ is of the form $\varphi(x) = \langle \Omega, x\Omega \rangle$ for some unit vector $\Omega \in \mathcal{H}$

Non-commutative distributions

Definition ("combinatorial distribution")

Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space.
Let $(x_i)_{i \in I}$ be a family of non-commutative random variables. We call the collection of all mixed moments

$$\{\varphi(x_{i_1} \cdots x_{i_k}) \mid k \in \mathbb{N}, i_1, \ldots, i_k \in I\}$$

their (joint) distribution.

Non-commutative distributions

Definition ("combinatorial distribution")

Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space.
Let $(x_i)_{i \in I}$ be a family of non-commutative random variables. We call the collection of all mixed moments

$$\{\varphi(x_{i_1} \cdots x_{i_k}) \mid k \in \mathbb{N}, i_1, \dots, i_k \in I\}$$

their (joint) distribution.

Definition ("analytic distribution")

Let $(\mathcal{A}, \varphi)$ be a C^*-probability space.
For any $x = x^* \in \mathcal{A}$, the distribution of x can be identified with the unique Borel probability measure μ_x on the real line $\mathbb{R}$ that satisfies

$$\varphi(x^k) = \int_{\mathbb{R}} t^k \, d\mu_x(t) \qquad \text{for all } k \in \mathbb{N}_0.$$

Section 2

Random Matrices

What are random matrices

- random matrices are 'typical' sequences of $N \times N$ matrices, with growing N

What are random matrices

- random matrices are 'typical' sequences of $N \times N$ matrices, with growing N

History

- 1928: Wishart introduced random matrices in statistics, for finite N

What are random matrices

- random matrices are 'typical' sequences of $N \times N$ matrices, with growing N

History

- 1928: Wishart introduced random matrices in statistics, for finite N
- 1955: Wigner introduced random matrices in physics, for a statistical description of nuclei of heavy atoms, and investigated the $N \to \infty$ asymptotics of these "Wigner matrices"

What are random matrices

- random matrices are 'typical' sequences of $N \times N$ matrices, with growing N

History

- 1928: Wishart introduced random matrices in statistics, for finite N
- 1955: Wigner introduced random matrices in physics, for a statistical description of nuclei of heavy atoms, and investigated the $N \to \infty$ asymptotics of these "Wigner matrices"
- 1967: Marchenko and Pastur described $N \to \infty$ asymptotics of "Wishart matrices"

What are random matrices

- random matrices are 'typical' sequences of $N \times N$ matrices, with growing N

History

- 1928: Wishart introduced random matrices in statistics, for finite N
- 1955: Wigner introduced random matrices in physics, for a statistical description of nuclei of heavy atoms, and investigated the $N \to \infty$ asymptotics of these "Wigner matrices"
- 1967: Marchenko and Pastur described $N \to \infty$ asymptotics of "Wishart matrices"
- 1972: Montgomery and Dyson discovered relation between zeros of the Riemann zeta function and eigenvalues of random matrices

What are random matrices

- random matrices are 'typical' sequences of $N \times N$ matrices, with growing N

History

- 1928: Wishart introduced random matrices in statistics, for finite N
- 1955: Wigner introduced random matrices in physics, for a statistical description of nuclei of heavy atoms, and investigated the $N \to \infty$ asymptotics of these "Wigner matrices"
- 1967: Marchenko and Pastur described $N \to \infty$ asymptotics of "Wishart matrices"
- 1972: Montgomery and Dyson discovered relation between zeros of the Riemann zeta function and eigenvalues of random matrices
- since 2000: random matrix theory developed into a central subject in mathematics, with many different connections

Oberwolfach workshop "Random Matrices" 2000

What are random matrices

- random matrices are 'typical' sequences of $N \times N$ matrices, with growing N
- more precisely: random matrices are sequences of $N \times N$ matrices whose entries are chosen randomly (according to a prescribed distribution)

What are random matrices

- random matrices are 'typical' sequences of $N \times N$ matrices, with growing N
- more precisely: random matrices are sequences of $N \times N$ matrices whose entries are chosen randomly (according to a prescribed distribution)

Fundamental observation

Many random matrices show for $N \to \infty$ almost surely a *deterministic* (and interesting) behavior

European Research Council
Established by the European Commission

Wigner random matrices (Wigner 1955)

Definition

A **Wigner random matrix** $X_N = \frac{1}{\sqrt{N}}(x_{ij})_{i,j=1}^N$

- is symmetric: $X_N^* = X_N$, i.e. $x_{ij} = x_{ji}$ for all i, j
- entries $\{x_{ij} \mid 1 \le i \le j \le N\}$ are chosen according to independent coin tosses head $= +1$, tail $= -1$

Wigner random matrices (Wigner 1955)

Definition

A **Wigner random matrix** $X_N = \frac{1}{\sqrt{N}} (x_{ij})_{i,j=1}^N$

- is symmetric: $X_N^* = X_N$, i.e. $x_{ij} = x_{ji}$ for all i,j
- entries $\{x_{ij} \mid 1 \leq i \leq j \leq N\}$ are chosen according to independent coin tosses head $= +1$, tail $= -1$

$$\begin{pmatrix} 1 & -1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 \\ -1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\ -1 & -1 & 1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 \end{pmatrix}$$

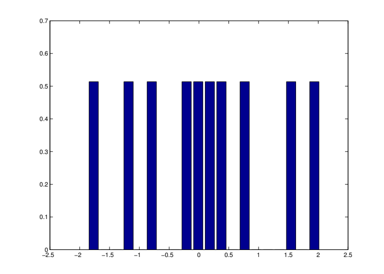
Wigner random matrices (Wigner 1955)

Definition

A **Wigner random matrix** $X_N = \frac{1}{\sqrt{N}}(x_{ij})_{i,j=1}^N$

- is symmetric: $X_N^* = X_N$, i.e. $x_{ij} = x_{ji}$ for all i, j
- entries $\{x_{ij} \mid 1 \le i \le j \le N\}$ are chosen according to independent coin tosses head $= +1$, tail $= -1$

Example: eigenvalue distribution for $N = 10$



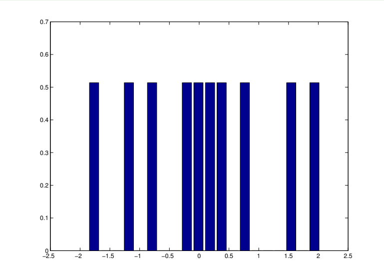
Wigner random matrices (Wigner 1955)

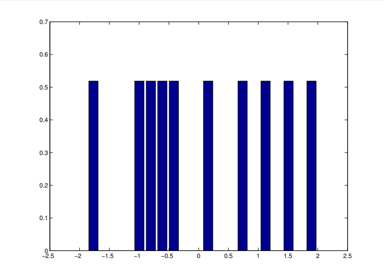
Definition

A **Wigner random matrix** $X_N = \frac{1}{\sqrt{N}} (x_{ij})_{i,j=1}^N$

- is symmetric: $X_N^* = X_N$, i.e. $x_{ij} = x_{ji}$ for all i, j
- entries $\{x_{ij} \mid 1 \leq i \leq j \leq N\}$ are chosen according to independent coin tosses head $= +1$, tail $= -1$

Example: eigenvalue distribution for $N = 10$





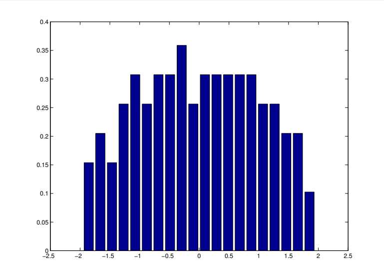
Wigner random matrices (Wigner 1955)

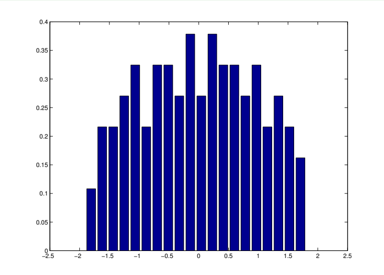
Definition

A **Wigner random matrix** $X_N = \frac{1}{\sqrt{N}}\left(x_{ij}\right)_{i,j=1}^{N}$

- is symmetric: $X_N^* = X_N$, i.e. $x_{ij} = x_{ji}$ for all i, j
- entries $\{x_{ij} \mid 1 \le i \le j \le N\}$ are chosen according to independent coin tosses head $= +1$, tail $= -1$

Example: eigenvalue distribution for $N = 100$





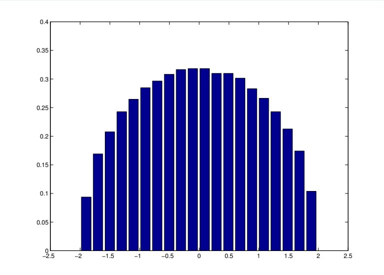
Wigner random matrices (Wigner 1955)

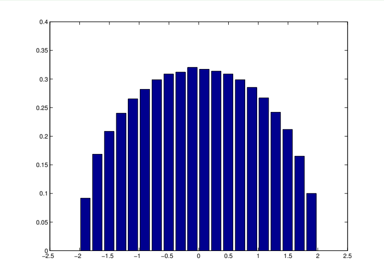
Definition

A **Wigner random matrix** $X_N = \frac{1}{\sqrt{N}} \left(x_{ij} \right)_{i,j=1}^{N}$

- is symmetric: $X_N^* = X_N$, i.e. $x_{ij} = x_{ji}$ for all i, j
- entries $\{x_{ij} \mid 1 \leq i \leq j \leq N\}$ are chosen according to independent coin tosses head $= +1$, tail $= -1$

Example: eigenvalue distribution for $N = 3000$





How to describe the deterministic limit

one-matrix case: limit of X_N

- the almost sure limit of μ_{X_N} is given by a probability measure μ

European Research Council
Established by the European Commission

How to describe the deterministic limit

one-matrix case: limit of X_N

- the almost sure limit of μ_{X_N} is given by a probability measure μ
- alternatively: try to find nice operator x on Hilbert space $\mathcal{H}$ with state φ such that $\mu = \mu_x$; then we can say that $X_N \to x$

How to describe the deterministic limit

one-matrix case: limit of X_N

- the almost sure limit of μ_{X_N} is given by a probability measure μ
- alternatively: try to find nice operator x on Hilbert space $\mathcal{H}$ with state φ such that $\mu = \mu_x$; then we can say that $X_N \to x$

multi-matrix case: limit of X_N, Y_N

- try to find (combinatorial) description of almost sure limit of μ_{X_N,Y_N}

How to describe the deterministic limit

one-matrix case: limit of X_N

- the almost sure limit of μ_{X_N} is given by a probability measure μ
- alternatively: try to find nice operator x on Hilbert space $\mathcal{H}$ with state φ such that $\mu = \mu_x$; then we can say that $X_N \to x$

multi-matrix case: limit of X_N, Y_N

- try to find (combinatorial) description of almost sure limit of μ_{X_N,Y_N}
- alternatively: try to find some nice operators x, y on a Hilbert space such that almost surely

$$\lim_{N\to\infty} \mathrm{tr}_N(q(X_N, Y_N)) = \varphi(q(x,y)) \qquad \text{for all monomials } q$$

How to describe the deterministic limit

one-matrix case: limit of X_N

- the almost sure limit of μ_{X_N} is given by a probability measure μ
- alternatively: try to find nice operator x on Hilbert space $\mathcal{H}$ with state φ such that $\mu = \mu_x$; then we can say that $X_N \to x$

multi-matrix case: limit of X_N, Y_N

- try to find (combinatorial) description of almost sure limit of μ_{X_N,Y_N}
- alternatively: try to find some nice operators x, y on a Hilbert space such that almost surely

$$\lim_{N\to\infty} \mathrm{tr}_N(q(X_N, Y_N)) = \varphi(q(x,y)) \qquad \text{for all monomials } q$$

- this means then in particular, that

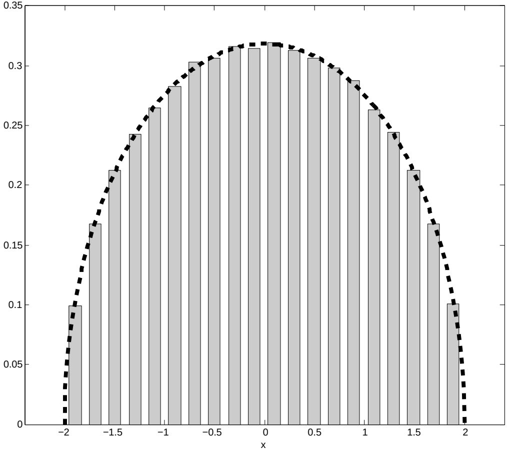
$$p(X_N, Y_N) \to p(x,y) \qquad \text{for **all** polynomials } p$$

Random matrices and operators

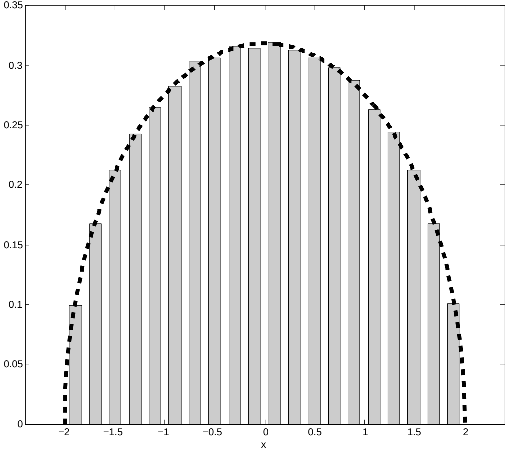
Fundamental observation of Voiculescu (1991)

limit of random matrices can often be described by "nice" and "interesting" operators on Hilbert spaces
(which, in the case of several matrices, describe interesting von Neumann algebras)

One-matrix case: classical random matrix case



One-matrix case: classical random matrix case



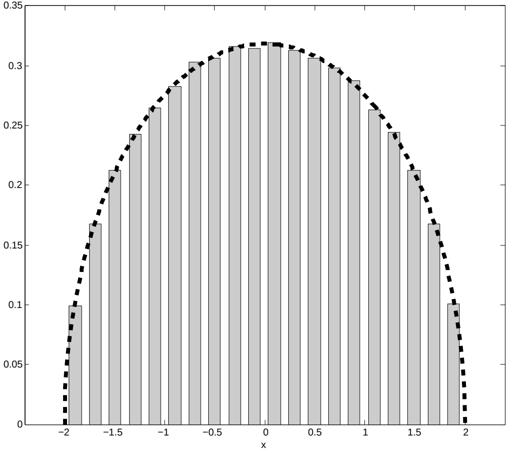
$x = l + l^*$,

l one-sided shift on $\bigoplus_{n\geq 0} \mathbb{C}e_n$

$le_n = e_{n+1}$
$l^* e_{n+1} = e_n, \; l^* e_0 = 0$

$\varphi(a) = \langle e_0, ae_0 \rangle$

One-matrix case: classical random matrix case



$x = l + l^*$,

l one-sided shift on $\bigoplus_{n \geq 0} \mathbb{C} e_n$

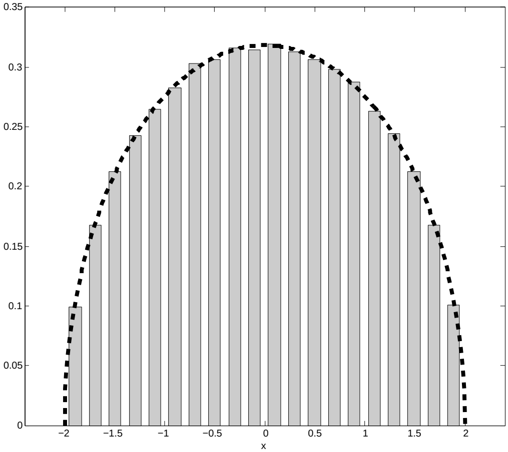
$le_n = e_{n+1}$
$l^* e_{n+1} = e_n, \; l^* e_0 = 0$

$\varphi(a) = \langle e_0, a e_0 \rangle$

- $X_N \to x$ in distribution (Wigner 1955)

European Research Council
Established by the European Commission

One-matrix case: classical random matrix case



$x = l + l^*$,

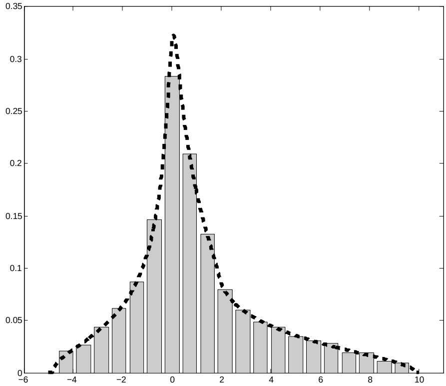
l one-sided shift on $\bigoplus_{n\geq 0} \mathbb{C}e_n$

$le_n = e_{n+1}$
$l^* e_{n+1} = e_n, \; l^* e_0 = 0$

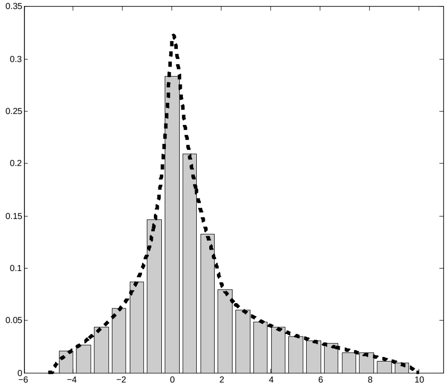
$\varphi(a) = \langle e_0, ae_0 \rangle$

- $X_N \to x$ in distribution (Wigner 1955)
- $\|X_N\| \to \|x\| = 2$ (Füredi, Komlós 1981)

Multi-matrix case: non-commutative case; $p(x, y) = xy + yx + x^2$



Multi-matrix case: non-commutative case; $p(x,y) = xy + yx + x^2$



X_N, Y_N independent Wigner

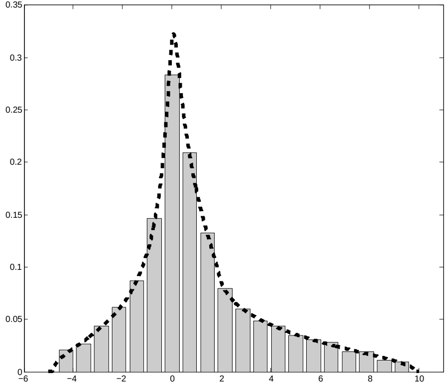
$x = l_1 + l_1^*$, $y = l_2 + l_2^*$

two copies of one-sided shift in different directions (creation and annihilation operators on full Fock space; Cuntz algebra)

$\varphi(a) = \langle \Omega, a\Omega \rangle$

European Research Council
Established by the European Commission

Multi-matrix case: non-commutative case; $p(x, y) = xy + yx + x^2$



X_N, Y_N independent Wigner

$x = l_1 + l_1^*$, $y = l_2 + l_2^*$

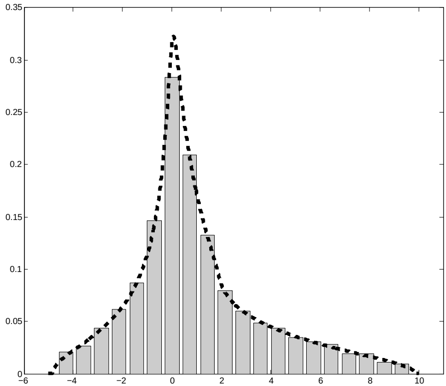
two copies of one-sided shift in different directions (creation and annihilation operators on full Fock space; Cuntz algebra)

$\varphi(a) = \langle \Omega, a\Omega \rangle$

- $p(X_N, Y_N) \to p(x, y)$ in distribution (Voiculescu 1991)

European Research Council
Established by the European Commission

Multi-matrix case: non-commutative case; $p(x,y) = xy + yx + x^2$



X_N, Y_N independent Wigner

$x = l_1 + l_1^*,\ y = l_2 + l_2^*$

two copies of one-sided shift in different directions (creation and annihilation operators on full Fock space; Cuntz algebra)

$\varphi(a) = \langle \Omega, a\Omega \rangle$

- $p(X_N, Y_N) \to p(x,y)$ in distribution (Voiculescu 1991)
- $\|p(X_N, Y_N)\| \to \|p(x,y)\|$ (Haagerup and Thorbjørnsen 2005)

European Research Council
Established by the European Commission

What are those limit operators x, y good for?

Basic Theorem (Voiculescu 1991)

For many random matrix models X_N, Y_N (like for independent Wigner matrices) the limit operators x, y are

- free in the sense of Voiculescu's free probability theory

European Research Council
Established by the European Commission

What are those limit operators x, y good for?

Basic Theorem (Voiculescu 1991)

For many random matrix models X_N, Y_N (like for independent Wigner matrices) the limit operators x, y are

- free in the sense of Voiculescu's free probability theory

Tools from free probability theory: for x and y free we have

- **free convolution**: the distribution of $x + y$ can effectively be calculated in terms of the distribution of x and the distribution of y
- **matrix-valued free convolution**: the matrix-valued distribution of $\alpha_0 \otimes 1 + \alpha_1 \otimes x + \alpha_2 \otimes y$ can be calculated in terms of the distribution of x and the distribution of y

European Research Council
Established by the European Commission

What are those limit operators x, y good for?

Basic Theorem (Voiculescu 1991)

For many random matrix models X_N, Y_N (like for independent Wigner matrices) the limit operators x, y are

- free in the sense of Voiculescu's free probability theory

Tools from free probability theory: for x and y free we have

- **free convolution**: the distribution of $x + y$ can effectively be calculated in terms of the distribution of x and the distribution of y
- **matrix-valued free convolution**: the matrix-valued distribution of $\alpha_0 \otimes 1 + \alpha_1 \otimes x + \alpha_2 \otimes y$ can be calculated in terms of the distribution of x and the distribution of y

Problem: what are such sums good for?

We are actually interested not just in the distribution of $x + y$, but much more general in the distribution of non-linear polynomials $p(x, y)$!

Section 3

Linearization

Idea of linearization

non-linear problem $\rightarrow$ operator-valued linear problem

Idea of linearization

non-linear problem $\rightarrow$ operator-valued linear problem

$$p(x_1, \dots, x_m) \rightarrow \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

European Research Council
Established by the European Commission

Idea of linearization

non-linear problem $\to$ operator-valued linear problem

$$p(x_1, \dots, x_m) \to \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

Example

Is $p(x, y) = xy$ invertible?

Idea of linearization

non-linear problem $\rightarrow$ operator-valued linear problem

$$p(x_1,\dots,x_m) \rightarrow \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

Example

Is $p(x,y) = xy$ invertible? We have

$$\begin{pmatrix} xy & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x \\ y & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix}$$

Idea of linearization

non-linear problem $\rightarrow$ operator-valued linear problem

$$p(x_1, \dots, x_m) \rightarrow \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

Example

Is $p(x, y) = xy$ invertible? We have

$$\begin{pmatrix} xy & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x \\ y & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -x \\ 0 & 1 \end{pmatrix}$$

Idea of linearization

non-linear problem → operator-valued linear problem

$$p(x_1, \dots, x_m) \to \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

Example

Is $p(x, y) = xy$ invertible? We have

$$\begin{pmatrix} xy & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x \\ y & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -x \\ 0 & 1 \end{pmatrix}$$

Hence xy is invertible if and only if

$$\begin{pmatrix} 0 & x \\ y & -1 \end{pmatrix}$$

is invertible

Idea of linearization

non-linear problem $\to$ operator-valued linear problem

$$p(x_1, \dots, x_m) \to \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

Example

Is $p(x,y) = xy$ invertible? We have

$$\begin{pmatrix} xy & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x \\ y & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -x \\ 0 & 1 \end{pmatrix}$$

Hence xy is invertible if and only if

$$\begin{pmatrix} 0 & x \\ y & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \otimes 1 + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes x + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes y = \hat{p}$$

is invertible

Idea of linearization

non-linear problem $\rightarrow$ operator-valued linear problem

$$p(x_1, \dots, x_m) \rightarrow \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

European Research Council
Established by the European Commission

Idea of linearization

non-linear problem $\to$ operator-valued linear problem

$$p(x_1, \dots, x_m) \to \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

Example

Is $p(x, y) = xy + yx + x^2$ invertible?

Idea of linearization

non-linear problem $\to$ operator-valued linear problem

$$p(x_1,\dots,x_m) \to \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

Example

Is $p(x,y) = xy + yx + x^2$ invertible? We have

$$\begin{pmatrix} xy+yx+x^2 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & y+\frac{x}{2} & x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x & y+\frac{x}{2} \\ x & 0 & -1 \\ y+\frac{x}{2} & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ y+\frac{x}{2} & 1 & 0 \\ x & 0 & 1 \end{pmatrix}$$

Idea of linearization

non-linear problem $\to$ operator-valued linear problem

$$p(x_1, \dots, x_m) \to \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

Example

Is $p(x,y) = xy + yx + x^2$ invertible? We have

$$\begin{pmatrix} xy + yx + x^2 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & y + \frac{x}{2} & x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x & y + \frac{x}{2} \\ x & 0 & -1 \\ y + \frac{x}{2} & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ y + \frac{x}{2} & 1 & 0 \\ x & 0 & 1 \end{pmatrix}$$

Hence $p = xy + yx + x^2$ is invertible if and only if

$$\begin{pmatrix} 0 & x & y + \frac{x}{2} \\ x & 0 & -1 \\ y + \frac{x}{2} & -1 & 0 \end{pmatrix}$$

is invertible.

Idea of linearization

non-linear problem $\to$ operator-valued linear problem

$$p(x_1, \dots, x_m) \to \hat{p} := \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots \alpha_m \otimes x_m$$

Example

Is $p(x,y) = xy + yx + x^2$ invertible? We have

$$\begin{pmatrix} xy+yx+x^2 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & y+\frac{x}{2} & x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x & y+\frac{x}{2} \\ x & 0 & -1 \\ y+\frac{x}{2} & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ y+\frac{x}{2} & 1 & 0 \\ x & 0 & 1 \end{pmatrix}$$

Hence $p = xy + yx + x^2$ is invertible if and only if

$$\begin{pmatrix} 0 & x & y+\frac{x}{2} \\ x & 0 & -1 \\ y+\frac{x}{2} & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \otimes 1 + \begin{pmatrix} 0 & 1 & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix} \otimes x + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \otimes y$$

is invertible.

Theorem (Haagerup, Thorbjørnsen 2005 (+Schultz 2006); Anderson 2012)

Every polynomial $p(x_1, \dots, x_m)$ *has a (non-unique) linearization* $\hat{p} = \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \dots + \alpha_m \otimes x_m$ *such that*

$$(z - p)^{-1} = [(\Lambda(z) - \hat{p})^{-1}]_{1,1}, \quad \textit{where} \quad \Lambda(z) = \begin{pmatrix} z & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Theorem (Haagerup, Thorbjørnsen 2005 (+Schultz 2006); Anderson 2012)

Every polynomial $p(x_1, \dots, x_m)$ *has a (non-unique) linearization* $\hat{p} = \alpha_0 \otimes 1 + \alpha_1 \otimes x_1 + \cdots + \alpha_m \otimes x_m$ *such that*

$$(z-p)^{-1} = [(\Lambda(z) - \hat{p})^{-1}]_{1,1}, \quad \textit{where} \quad \Lambda(z) = \begin{pmatrix} z & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

and hence also

$$G_p(z) = \varphi((z-p)^{-1}) = [\varphi \otimes 1(\Lambda(z) - \hat{p})^{-1}]_{1,1} = [G_{\hat{p}}(\Lambda(z))]_{1,1}.$$

- $G_p(z) = \varphi[(z-p)^{-1}]$ is the Cauchy transform of p
- $G_{\hat{p}}(b) = \varphi \otimes 1[(b-\hat{p})^{-1}]$ is the operator-valued Cauchy transform of $\hat{p}$

Distribution of $p(x, y)$ for x and y free

Calculation of distribution of p by linearization and operator-valued convolution (Belinschi, Mai, Speicher 2013)

- we want the distribution of $p(x, y) = xy + yx + x^2$

Distribution of $p(x,y)$ for x and y free

Calculation of distribution of p by linearization and operator-valued convolution (Belinschi, Mai, Speicher 2013)

- we want the distribution of $p(x,y) = xy + yx + x^2$
- this is determined by the operator-valued distribution of its linearization

$$\hat{p}(x,y) = \begin{pmatrix} 0 & x & y + \frac{x}{2} \\ x & 0 & -1 \\ y + \frac{x}{2} & -1 & 0 \end{pmatrix}$$

Distribution of $p(x, y)$ for x and y free

Calculation of distribution of p by linearization and operator-valued convolution (Belinschi, Mai, Speicher 2013)

- we want the distribution of $p(x, y) = xy + yx + x^2$
- this is determined by the operator-valued distribution of its linearization
- but this is now an additive (operator-valued) problem

$$\hat{p}(x, y) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \otimes 1 + \begin{pmatrix} 0 & 1 & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix} \otimes x + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \otimes y$$

Distribution of $p(x,y)$ for x and y free

Calculation of distribution of p by linearization and operator-valued convolution (Belinschi, Mai, Speicher 2013)

- we want the distribution of $p(x,y) = xy + yx + x^2$
- this is determined by the operator-valued distribution of its linearization
- but this is now an additive (operator-valued) problem

$$\hat{p}(x,y) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \otimes 1 + \begin{pmatrix} 0 & 1 & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix} \otimes x + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \otimes y$$

- for this we have analytic theory of operator-valued free convolution
- this allows to calculate the operator-valued Cauchy transform of $\hat{p}$, and thus the Cauchy transform of p

How to calculate operator-valued free convolution

Theorem (Belinschi, Mai, Speicher 2013)

Consider

$$\hat{p}(x,y) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \otimes 1 + \begin{pmatrix} 0 & 1 & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix} \otimes x + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \otimes y =: X + Y$$

How to calculate operator-valued free convolution

Theorem (Belinschi, Mai, Speicher 2013)

Consider $\hat{p} = X + Y$ Then X and Y are free in the operator-valued sense and there exists a unique pair of (Fréchet-)holomorphic maps $\omega_1, \omega_2 : \mathbb{H}^+(\mathcal{B}) \to \mathbb{H}^+(\mathcal{B})$, such that

$$G_X(\omega_1(b)) = G_Y(\omega_2(b)) = G_{X+Y}(b), \quad b \in \mathbb{H}^+(\mathcal{B}).$$

How to calculate operator-valued free convolution

Theorem (Belinschi, Mai, Speicher 2013)

Consider $\hat{p} = X + Y$ Then X and Y are free in the operator-valued sense and there exists a unique pair of (Fréchet-)holomorphic maps $\omega_1, \omega_2 : \mathbb{H}^+(\mathcal{B}) \to \mathbb{H}^+(\mathcal{B})$, such that

$$G_X(\omega_1(b)) = G_Y(\omega_2(b)) = G_{X+Y}(b), \quad b \in \mathbb{H}^+(\mathcal{B}).$$

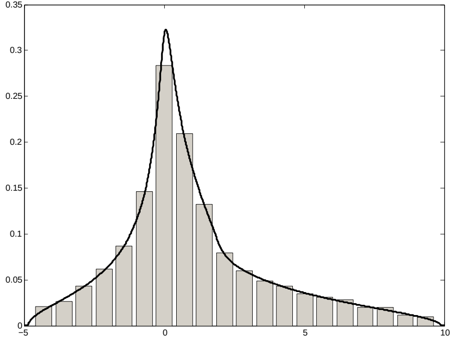
Moreover, ω_1 and ω_2 can easily be calculated via the following fixed point iterations on $\mathbb{H}^+(\mathcal{B})$

$$w \mapsto h_Y(b + h_X(w)) + b \qquad \text{for } \omega_1(b)$$
$$w \mapsto h_X(b + h_Y(w)) + b \qquad \text{for } \omega_2(b)$$

where we put $h_X(b) := G_X(b)^{-1} - b$ and $h_Y(b) := G_Y(b)^{-1} - b$;

$$\mathcal{B} = M_3(\mathbb{C}), \qquad \mathbb{H}(\mathcal{B}) := \{b \in \mathcal{B} \mid \Im b = \frac{b - b^*}{2i} > 0\}$$

Distribution of $p(x,y) = xy + yx + x^2$



$x = l_1 + l_1^*$, $y = l_2 + l_2^*$

two copies of one-sided shift in different directions (creation and annihilation operators on full Fock space; Cuntz algebra relations)

$\varphi(a) = \langle \Omega, a\Omega \rangle$

Historical remark

Note that this linearization trick is also a well-known idea in many other mathematical communities, known under various names like

- Higman's trick (Higman "The units of group rings", 1940)
- recognizable power series (automata theory, Kleene 1956, Schützenberger 1961)
- linearization by enlargement (ring theory, Cohn 1985; Cohn and Reutenauer 1994, Malcolmson 1978)
- descriptor realization (control theory, Kalman 1963; Helton, McCullough, Vinnikov 2006)

European Research Council
Established by the European Commission

Historical remark

Note that this linearization trick is also a well-known idea in many other mathematical communities, known under various names like

- Higman's trick (Higman "The units of group rings", 1940)
- recognizable power series (automata theory, Kleene 1956, Schützenberger 1961)
- linearization by enlargement (ring theory, Cohn 1985; Cohn and Reutenauer 1994, Malcolmson 1978)
- descriptor realization (control theory, Kalman 1963; Helton, McCullough, Vinnikov 2006)

$\curvearrowright$ **Linearization even works for non-commutative rational functions!**

European Research Council
Established by the European Commission

Section 4

Non-Commutative Rational Functions

Non-commutative rational functions

Non-commutative rational functions (Amitsur 1966, Cohn 1971)

- are given by rational expressions in non-commuting variables, like

$$r(x,y) := (4-x)^{-1} + (4-x)^{-1} y \big((4-x) - y(4-x)^{-1} y \big)^{-1} y (4-x)^{-1}$$

European Research Council
Established by the European Commission

Non-commutative rational functions

Non-commutative rational functions (Amitsur 1966, Cohn 1971)

- are given by rational expressions in non-commuting variables, like

$$r(x,y) := (4-x)^{-1} + (4-x)^{-1}y\big((4-x) - y(4-x)^{-1}y\big)^{-1}y(4-x)^{-1}$$

- form a skew field, so each $r \neq 0$ is invertible

European Research Council
Established by the European Commission

Non-commutative rational functions

Non-commutative rational functions (Amitsur 1966, Cohn 1971)

- are given by rational expressions in non-commuting variables, like

$$r(x,y) := (4-x)^{-1} + (4-x)^{-1}y\big((4-x) - y(4-x)^{-1}y\big)^{-1}y(4-x)^{-1}$$

- form a skew field, so each $r \neq 0$ is invertible
- deciding whether $r = 0$ is not easy, for example

$$x_2^{-1} + x_2^{-1}(x_3^{-1}x_1^{-1} - x_2^{-1})^{-1}x_2^{-1} - (x_2 - x_3x_1)^{-1} = 0$$

European Research Council
Established by the European Commission

Non-commutative rational functions

Non-commutative rational functions (Amitsur 1966, Cohn 1971)

- are given by rational expressions in non-commuting variables, like

$$r(x,y) := (4-x)^{-1} + (4-x)^{-1}y\big((4-x) - y(4-x)^{-1}y\big)^{-1}y(4-x)^{-1}$$

- form a skew field, so each $r \neq 0$ is invertible
- deciding whether $r = 0$ is not easy, for example

$$x_2^{-1} + x_2^{-1}(x_3^{-1}x_1^{-1} - x_2^{-1})^{-1}x_2^{-1} - (x_2 - x_3x_1)^{-1} = 0$$

- there is no standard form for a rational function
- in general nested inversions are needed

European Research Council
Established by the European Commission

Non-commutative rational functions

Non-commutative rational functions

- $r(x,y) := (4-x)^{-1} + (4-x)^{-1} y \big((4-x) - y(4-x)^{-1} y\big)^{-1} y (4-x)^{-1}$

Non-commutative rational functions

Non-commutative rational functions

- $r(x,y) := (4-x)^{-1} + (4-x)^{-1} y \big((4-x) - y(4-x)^{-1} y\big)^{-1} y (4-x)^{-1}$
- can always be realized in terms of matrices of (linear) polynomials, like

$$r = \begin{pmatrix} \frac{1}{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 + \frac{1}{4}x & \frac{1}{4}y \\ \frac{1}{4}y & -1 + \frac{1}{4}x \end{pmatrix}^{-1} \cdot \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}$$

Non-commutative rational functions

Non-commutative rational functions

- $r(x,y) := (4-x)^{-1} + (4-x)^{-1}y\big((4-x) - y(4-x)^{-1}y\big)^{-1}y(4-x)^{-1}$
- can always be realized in terms of matrices of (linear) polynomials, like

$$r = \begin{pmatrix} \frac{1}{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} -1+\frac{1}{4}x & \frac{1}{4}y \\ \frac{1}{4}y & -1+\frac{1}{4}x \end{pmatrix}^{-1} \cdot \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}$$

- this gives also linearizations $\hat{r}$ of any non-commutative rational function r, like

$$\hat{r} = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & -1+\frac{1}{4}x & \frac{1}{4}y \\ 0 & \frac{1}{4}y & -1+\frac{1}{4}x \end{pmatrix}$$

Non-commutative rational functions

Non-commutative rational functions

- $r(x,y) := (4-x)^{-1} + (4-x)^{-1}y\big((4-x) - y(4-x)^{-1}y\big)^{-1}y(4-x)^{-1}$
- can always be realized in terms of matrices of (linear) polynomials, like

$$r = \begin{pmatrix} \frac{1}{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 + \frac{1}{4}x & \frac{1}{4}y \\ \frac{1}{4}y & -1 + \frac{1}{4}x \end{pmatrix}^{-1} \cdot \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}$$

- this gives also linearizations $\hat{r}$ of any non-commutative rational function r, like

$$\hat{r} = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \otimes 1 + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix} \otimes x + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{pmatrix} \otimes y$$

Non-commutative rational functions

Non-commutative rational functions

- $r(x,y) := (4-x)^{-1} + (4-x)^{-1}y\big((4-x) - y(4-x)^{-1}y\big)^{-1}y(4-x)^{-1}$
- can always be realized in terms of matrices of (linear) polynomials, like

$$r = \begin{pmatrix} \frac{1}{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} -1+\frac{1}{4}x & \frac{1}{4}y \\ \frac{1}{4}y & -1+\frac{1}{4}x \end{pmatrix}^{-1} \cdot \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}$$

- this gives also linearizations $\hat{r}$ of any non-commutative rational function r, like

$$\hat{r} = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \otimes 1 + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix} \otimes x + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{pmatrix} \otimes y$$

- if x and y are free, this is an operator-valued free convolution problem and can be solved as before (Helton, Mai, Speicher 2016)

Rational functions of random matrices and their limit

Proposition (Sheng Yin 2016)

Consider selfadjoint random matrices X_N, Y_N which converge to selfadjoint operators x, y in the following strong sense: for any selfadjoint polyomial p we have

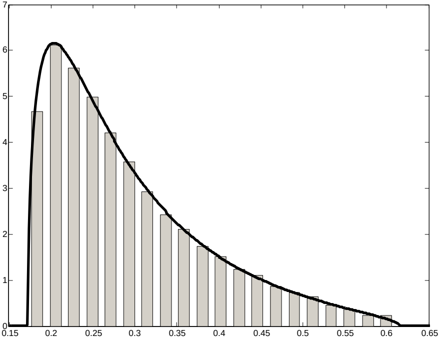
- $p(X_N, Y_N) \to p(x, y)$ in distribution
- $\lim_{N\to\infty} \|p(X_N, Y_N)\| = \|p(x, y)\|$

Then this strong convergence remains also true for rational functions: Let r be a non-commutative rational function, such that $r(x, y)$ is defined. Then we have almost surely that

- $r(X_N, Y_Y)$ is defined eventually for large N
- $r(X_N, Y_N) \to r(x, y)$ in distribution
- $\lim_{N\to\infty} \|r(X_N, Y_N)\| = \|r(x, y)\|$

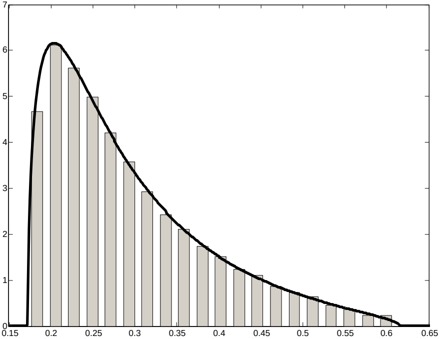
Distribution of random matrices and their limit for

$$r(x,y) = (4-x)^{-1} + (4-x)^{-1}y\big((4-x) - y(4-x)^{-1}y\big)^{-1}y(4-x)^{-1}$$



Distribution of random matrices and their limit for

$$r(x,y) = (4-x)^{-1} + (4-x)^{-1}y\big((4-x) - y(4-x)^{-1}y\big)^{-1}y(4-x)^{-1}$$



X_N, Y_N independent Wigner

$x = l_1 + l_1^*$, $y = l_2 + l_2^*$

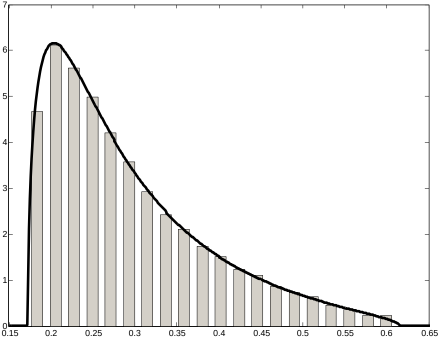
two copies of one-sided shift in different directions (creation and annihilation operators on full Fock space; Cuntz algebra)

$\varphi(a) = \langle \Omega, a\Omega \rangle$

European Research Council
Established by the European Commission

Distribution of random matrices and their limit for

$r(x,y) = (4-x)^{-1} + (4-x)^{-1} y \big((4-x) - y(4-x)^{-1} y\big)^{-1} y (4-x)^{-1}$



X_N, Y_N independent Wigner

$x = l_1 + l_1^*$, $y = l_2 + l_2^*$

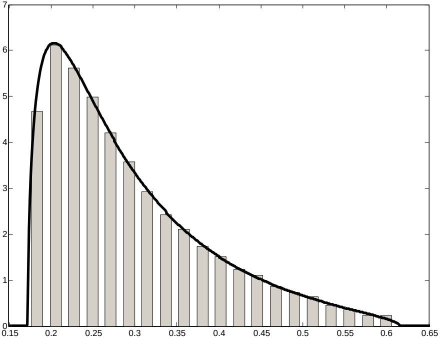
two copies of one-sided shift in different directions (creation and annihilation operators on full Fock space; Cuntz algebra)

$\varphi(a) = \langle \Omega, a\Omega \rangle$

- $r(X_N, Y_N) \to r(x,y)$ in distribution (Yin 2017)

Distribution of random matrices and their limit for

$r(x,y) = (4-x)^{-1} + (4-x)^{-1}y\big((4-x) - y(4-x)^{-1}y\big)^{-1}y(4-x)^{-1}$



X_N, Y_N independent Wigner

$x = l_1 + l_1^*$, $y = l_2 + l_2^*$

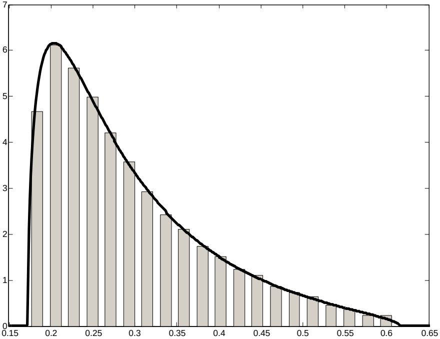
two copies of one-sided shift in different directions (creation and annihilation operators on full Fock space; Cuntz algebra)

$\varphi(a) = \langle \Omega, a\Omega \rangle$

- $r(X_N, Y_N) \to r(x,y)$ in distribution (Yin 2017)
- $\|r(X_N, Y_N)\| \to \|r(x,y)\|$ (Yin 2017)

Distribution of random matrices and their limit for

$$r(x,y) = (4-x)^{-1} + (4-x)^{-1}y\big((4-x) - y(4-x)^{-1}y\big)^{-1}y(4-x)^{-1}$$



X_N, Y_N independent Wigner

$x = l_1 + l_1^*$, $y = l_2 + l_2^*$

two copies of one-sided shift in different directions (creation and annihilation operators on full Fock space; Cuntz algebra)

$\varphi(a) = \langle \Omega, a\Omega \rangle$

- $r(X_N, Y_N) \to r(x,y)$ in distribution (Yin 2017)
- $\|r(X_N, Y_N)\| \to \|r(x,y)\|$ (Yin 2017)

Thank you!