

Jeffery-Williams Lecture

**On the effectiveness of operator-valued
free probability theory**

Roland Speicher
Universität des Saarlandes
Saarbrücken, Germany

joint work with Serban Belinschi, Tobias Mai, John Treilhard,
Carlos Vargas

Once Upon a Time

... There Were Large Random Matrices

We are interested in the limiting eigenvalue distribution of an

$N \times N$ random matrix for $N \rightarrow \infty$.

We are interested in the limiting eigenvalue distribution of an

$N \times N$ random matrix for $N \rightarrow \infty$.

Typical phenomena for basic random matrix ensembles:

- almost sure convergence to a deterministic limit eigenvalue distribution
- this limit distribution can be effectively calculated

We are interested in the limiting eigenvalue distribution of an

$N \times N$ random matrix for $N \rightarrow \infty$.

Typical phenomena for basic random matrix ensembles:

- almost sure convergence to a deterministic limit eigenvalue distribution
- this limit distribution can be effectively calculated

Wigner Random Matrix

A Wigner random matrix

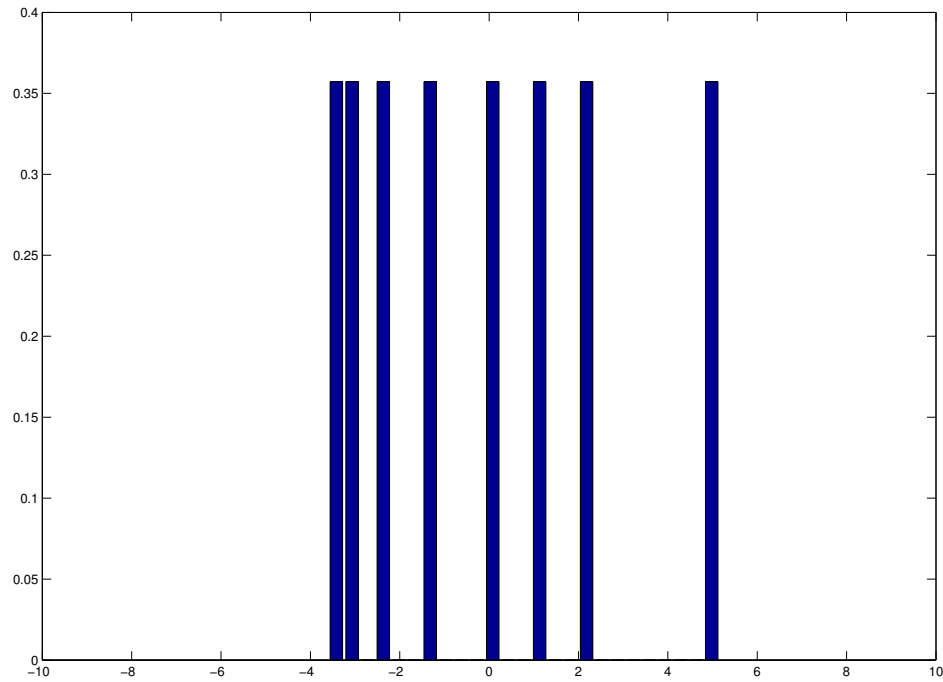
$$X = \left(x_{ij} \right)_{i,j=1}^N$$

- is symmetric:

$$X^* = X$$

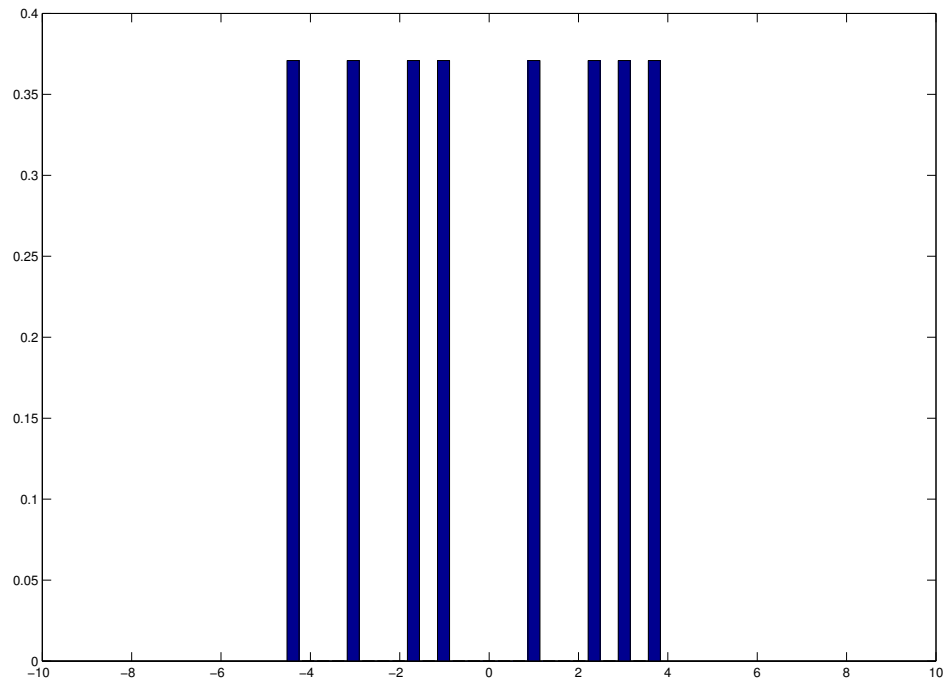
- $\{x_{ij} \mid 1 \leq i \leq j \leq N\}$ are independent and identically distributed

8 eigenvalues of an 8×8 matrix with random ± 1 entries



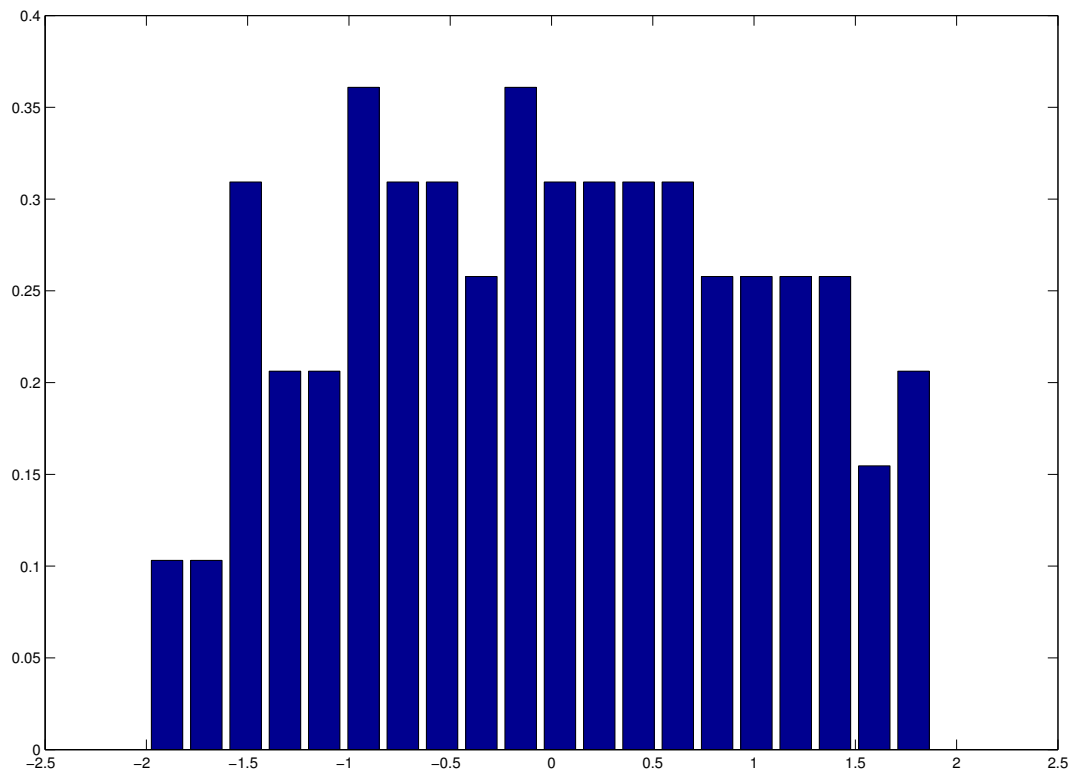
$$\begin{pmatrix} -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 \\ -1 & -1 & 1 & -1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 & -1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

8 eigenvalues of an 8×8 matrix with random ± 1 entries



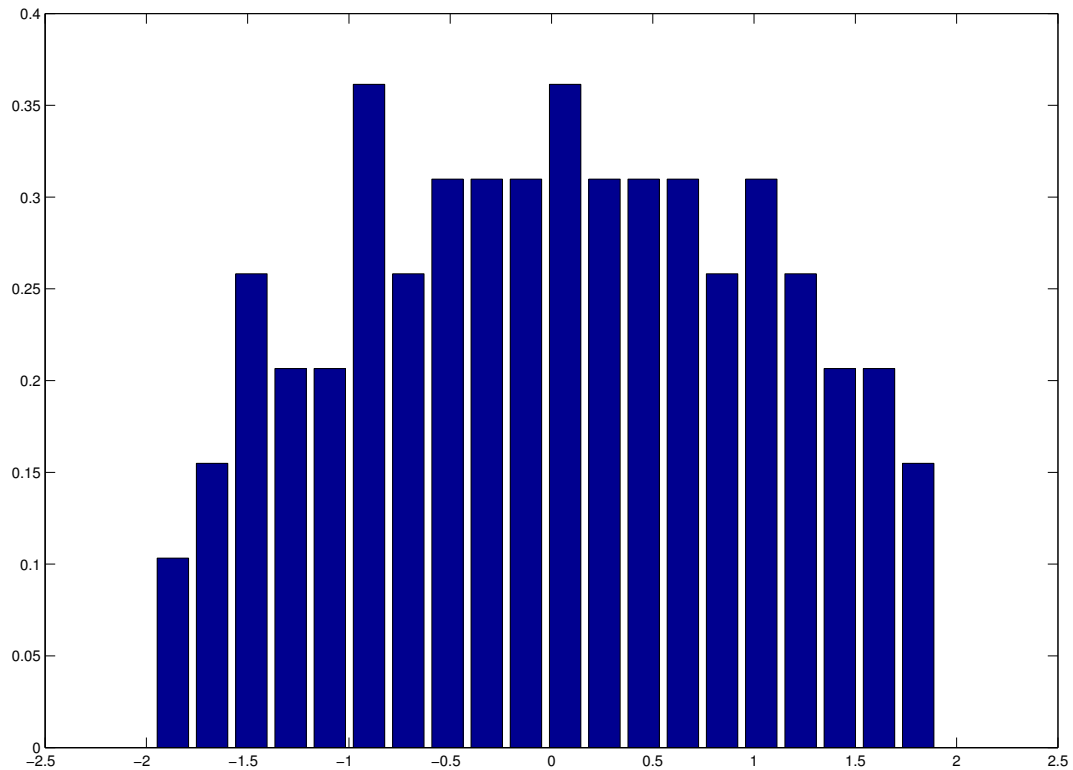
$$\begin{pmatrix} 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ -1 & 1 & -1 & 1 & -1 &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

100 eigenvalues of 100×100 matrix with random $\frac{\pm 1}{\sqrt{N}}$ entries



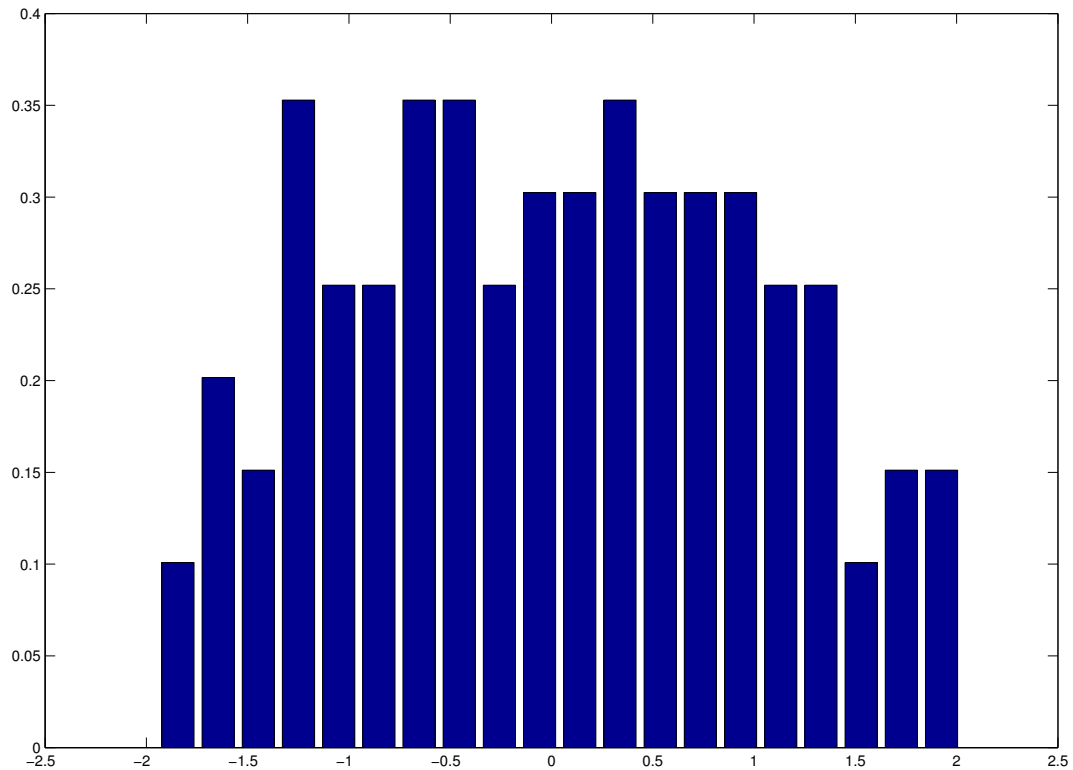
$N = 100$ realisation 1

100 eigenvalues of 100×100 matrix with random $\frac{\pm 1}{\sqrt{N}}$ entries



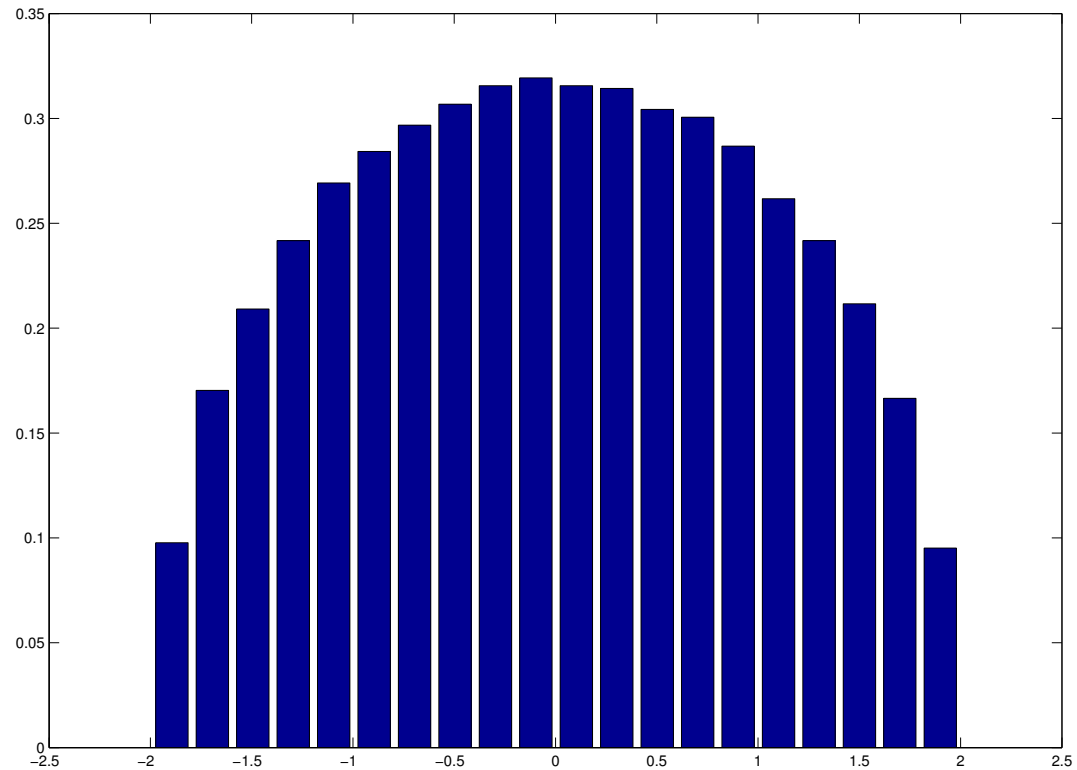
$N = 100$ realisation 2

100 eigenvalues of 100×100 matrix with random $\frac{\pm 1}{\sqrt{N}}$ entries



$N = 100$ realisation 3

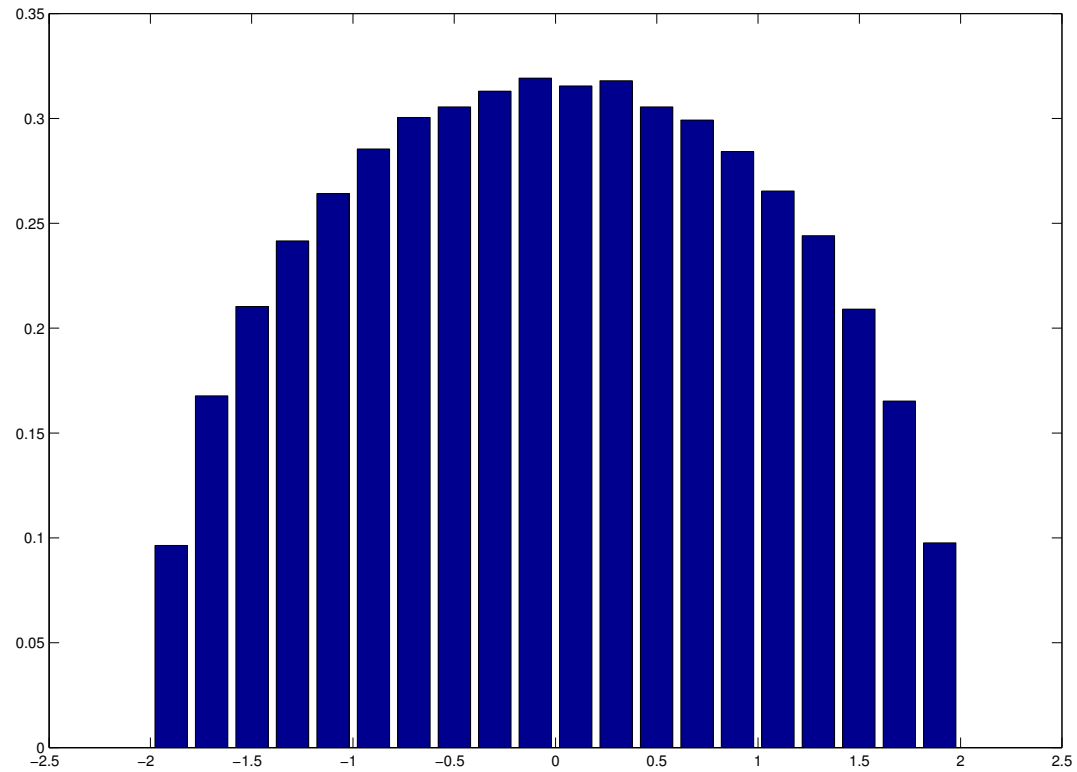
4000 eigenvalues of 4000×4000 matrix with random $\frac{\pm 1}{\sqrt{N}}$ entries



$N = 4000$

realisation 1

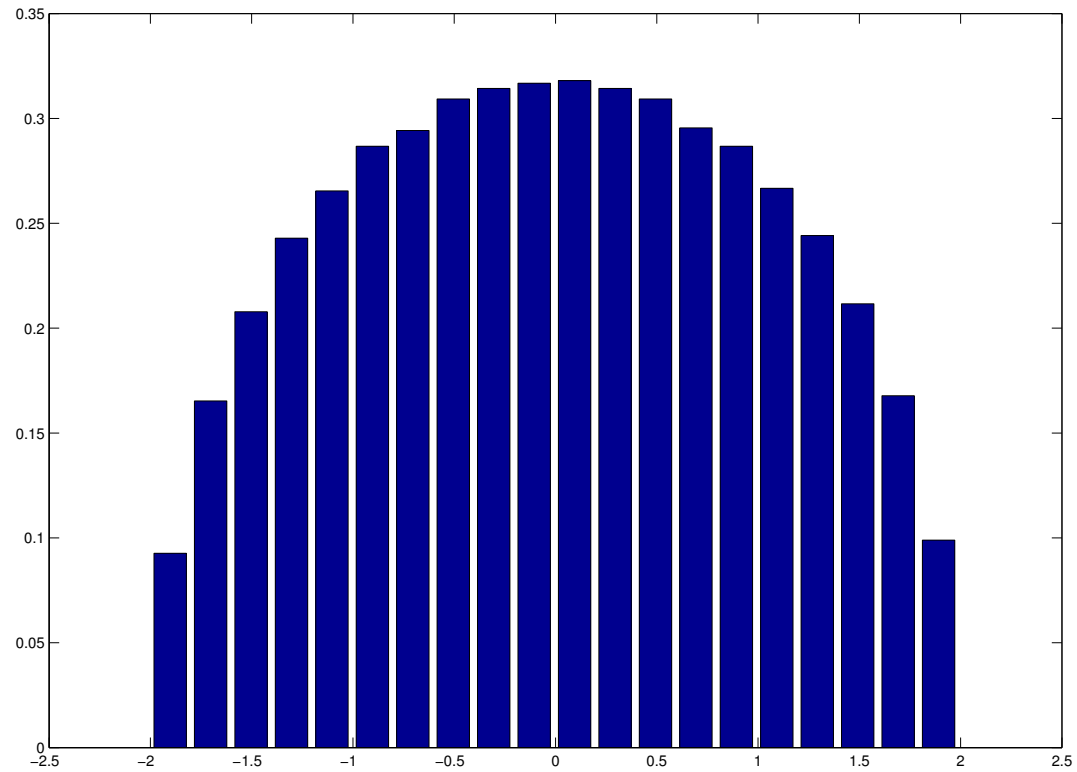
4000 eigenvalues of 4000×4000 matrix with random $\frac{\pm 1}{\sqrt{N}}$ entries



$N = 4000$

realisation 2

4000 eigenvalues of 4000×4000 matrix with random $\frac{\pm 1}{\sqrt{N}}$ entries



$N = 4000$

realisation 3

Almost Sure Convergence to a Deterministic Limit Eigenvalue Distribution

For large N , the eigenvalue distribution of X is with very high probability very close to a deterministic “limit distribution”.

Wishart Random Matrix

A Wishart random matrix X is of the form $X = AA^*$ where

- A is an $N \times M$ matrix

$$A = \left(a_{ij} \right)_{\substack{i=1,\dots,N \\ j=1,\dots,M}}$$

- where all entries are independent and identically distributed:

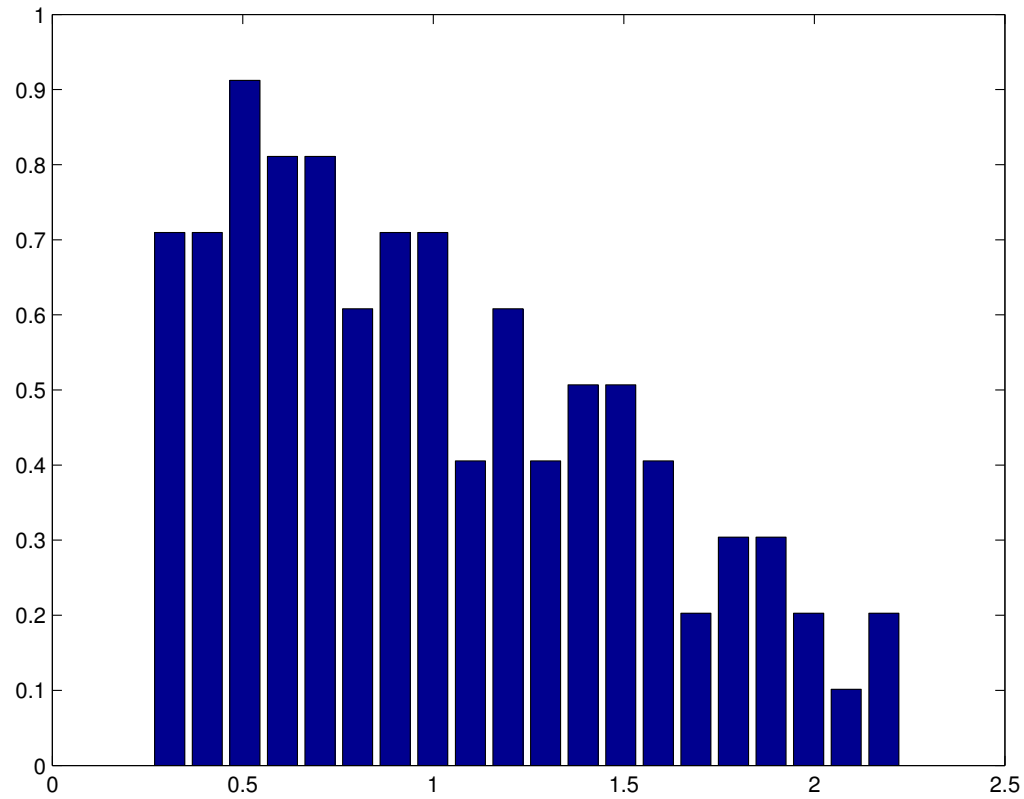
$$\{a_{ij} \mid 1 \leq i \leq N, 1 \leq j \leq M\} \text{ are iid}$$

For $N \rightarrow \infty$, one keeps the ratio

$$\lambda := \frac{N}{M}$$

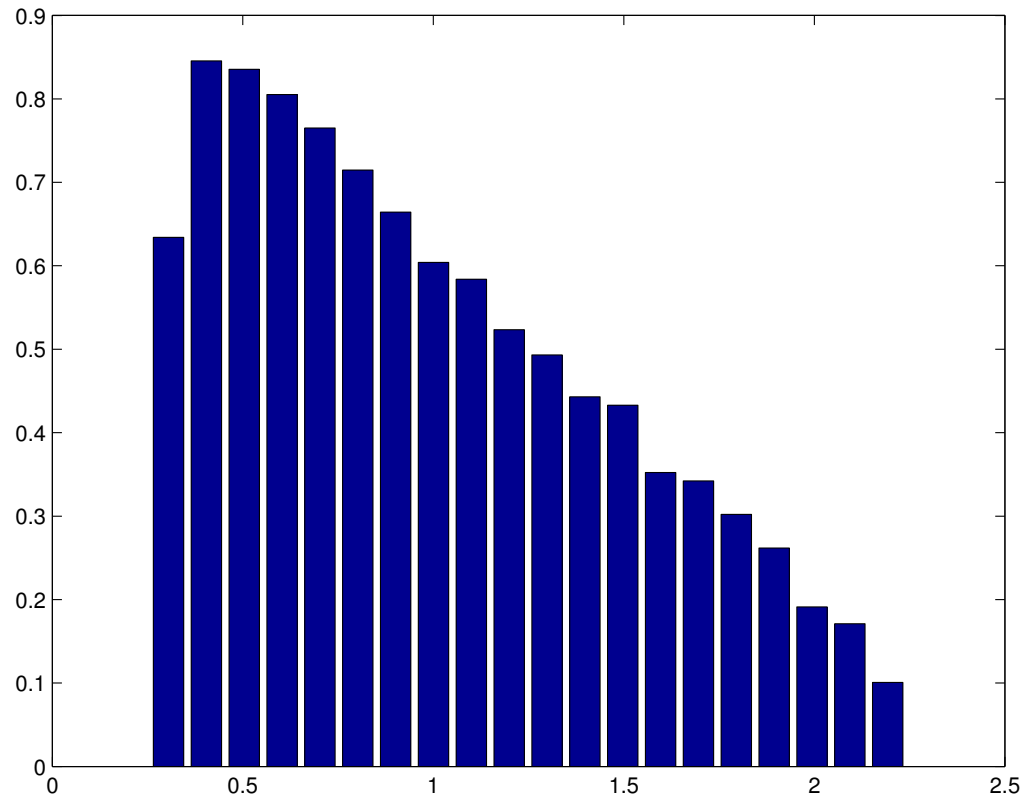
fixed.

100 eigenvalues of a Wishart matrix, with $\lambda = 0.25$



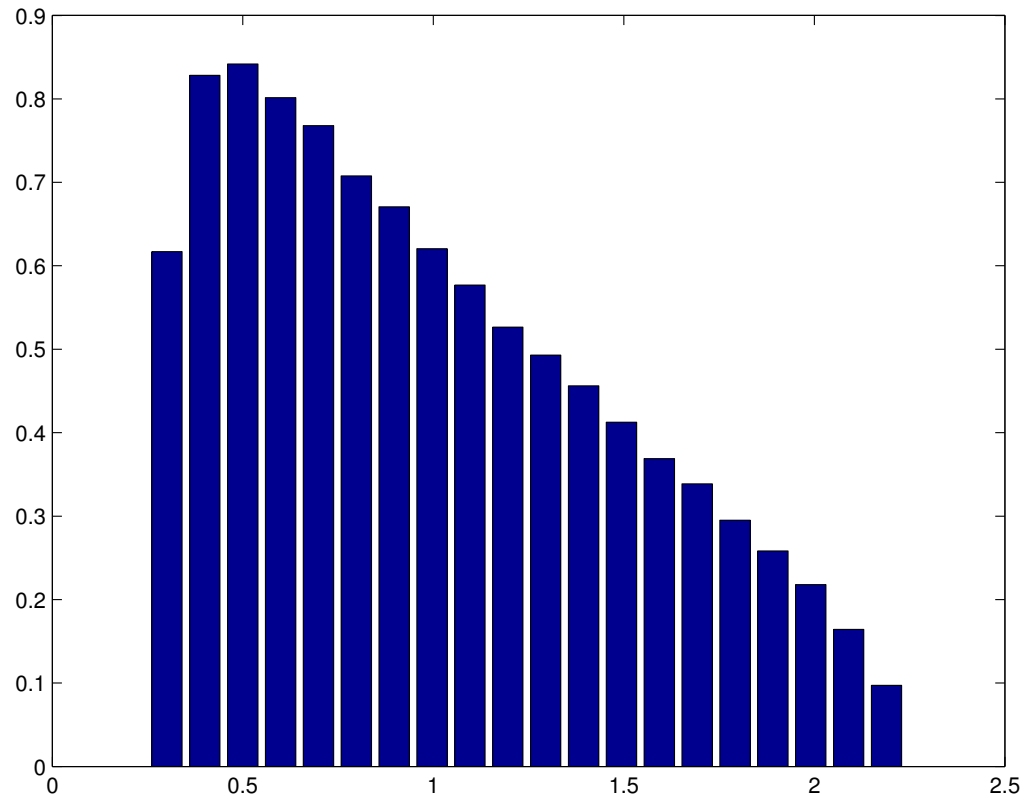
$N = 100$

1000 eigenvalues of a Wishart matrix, with $\lambda = 0.25$



$N = 1000$

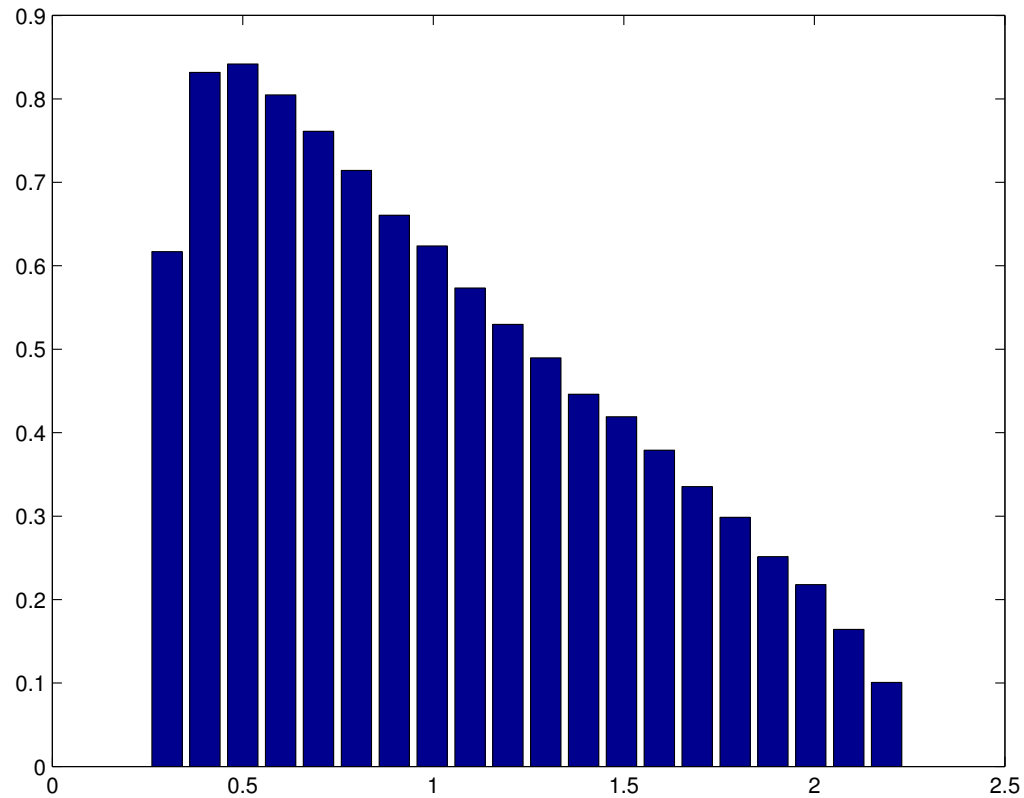
3000 eigenvalues of a Wishart matrix, with $\lambda = 0.25$



$N = 3000$

realisation 1

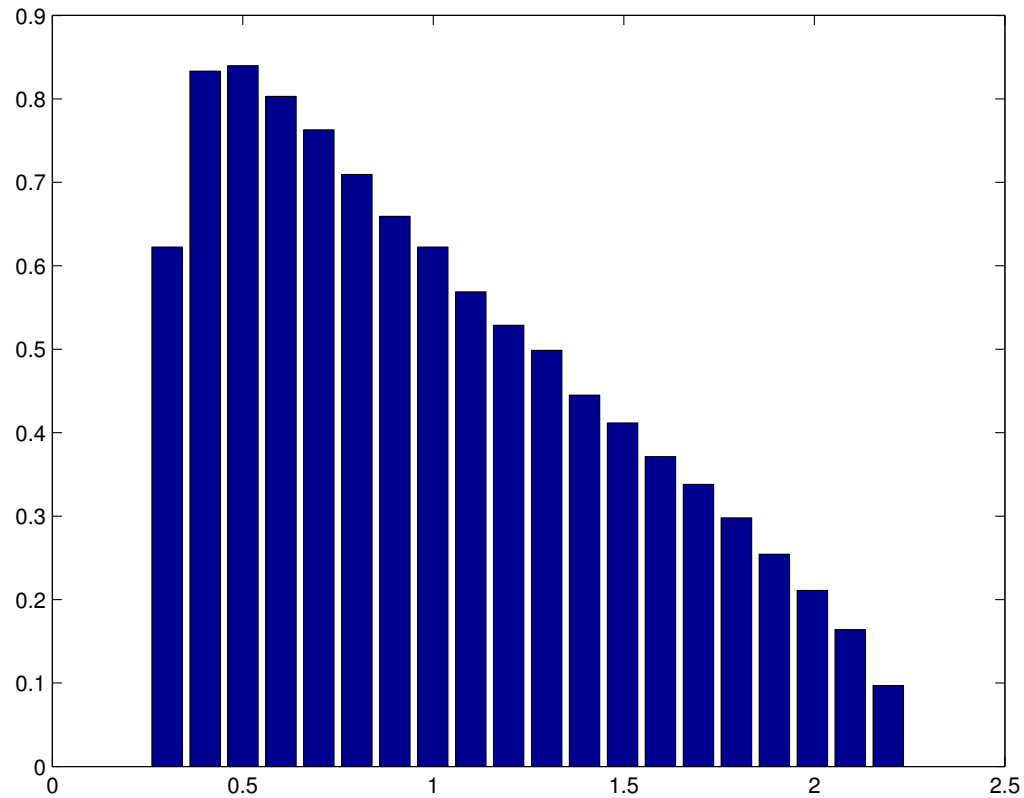
3000 eigenvalues of a Wishart matrix, with $\lambda = 0.25$



$N = 3000$

realisation 2

3000 eigenvalues of a Wishart matrix, with $\lambda = 0.25$



$N = 3000$

realisation 3

Almost Sure Convergence to a Deterministic Limit Eigenvalue Distribution

For large N , the eigenvalue distribution of X is with very high probability (for generic choices of X) very close to a deterministic “limit distribution”, which depends on λ .

We are interested in the limiting eigenvalue distribution of an

$N \times N$ random matrix for $N \rightarrow \infty$.

Typical phenomena for basic random matrix ensembles:

- almost sure convergence to a deterministic limit eigenvalue distribution
- this limit distribution can be effectively calculated

We are interested in the limiting eigenvalue distribution of an

$N \times N$ random matrix for $N \rightarrow \infty$.

Typical phenomena for basic random matrix ensembles:

- almost sure convergence to a deterministic limit eigenvalue distribution
- **this limit distribution can be effectively calculated**

The Cauchy (or Stieltjes) Transform

For any probability measure μ on \mathbb{R} we define its Cauchy transform

$$G(z) := \int_{\mathbb{R}} \frac{1}{z - t} d\mu(t)$$

This is an analytic function $G : \mathbb{C}^+ \rightarrow \mathbb{C}^-$ and we can recover μ from G by **Stieltjes inversion formula**

$$d\mu(t) = -\frac{1}{\pi} \lim_{\varepsilon \rightarrow 0} \Im G(t + i\varepsilon) dt$$

For our basic random matrix ensembles one can derive equations for the Cauchy transform of the limiting eigenvalue distribution, solve those equations and then get the density via Stieltjes inversion:

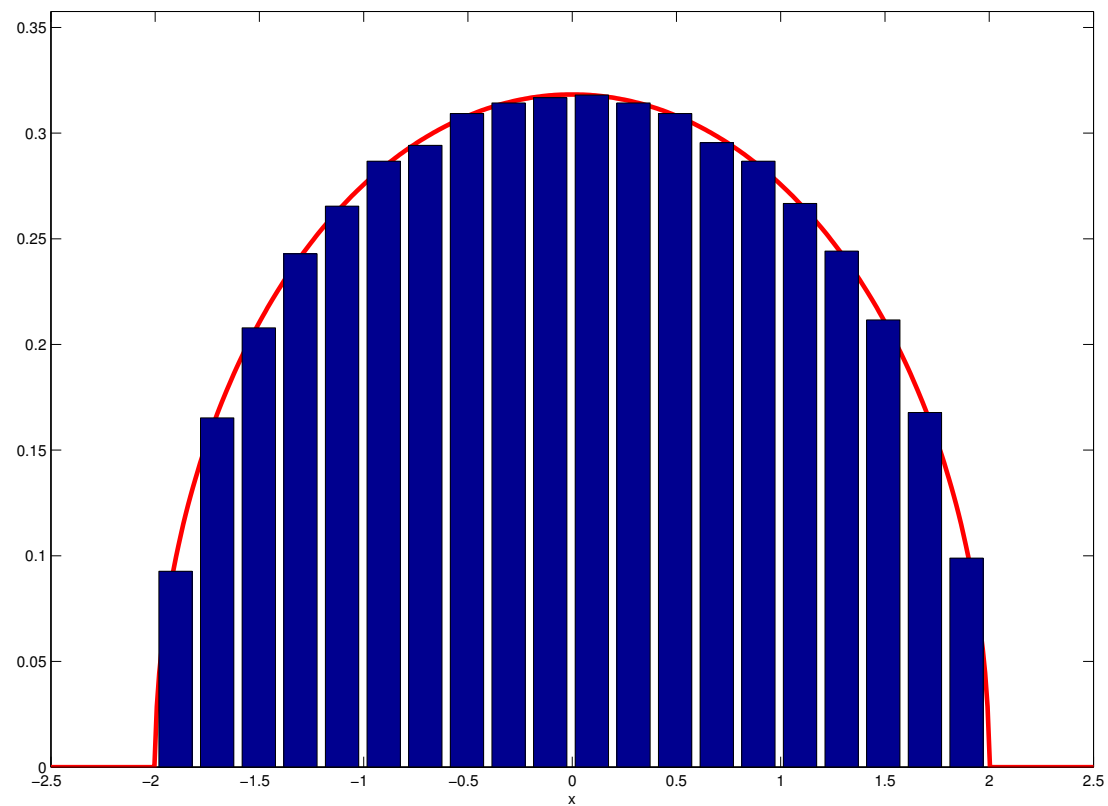
Wigner random matrix

$$G(z)^2 + 1 = zG(z),$$

which can be solved as

$$G(z) = \frac{z - \sqrt{z^2 - 4}}{2}, \quad \text{thus} \quad d\mu_s(t) = \frac{1}{2\pi} \sqrt{4 - t^2} dt$$

Wigner random matrix and Wigner's semicircle



For our basic random matrix ensembles one can derive equations for the Cauchy transform of the limiting eigenvalue distribution, solve those equations and then get the density via Stieltjes inversion:

Wishart random matrix

$$\frac{\lambda}{1 - G(z)} + \frac{1}{G(z)} = z$$

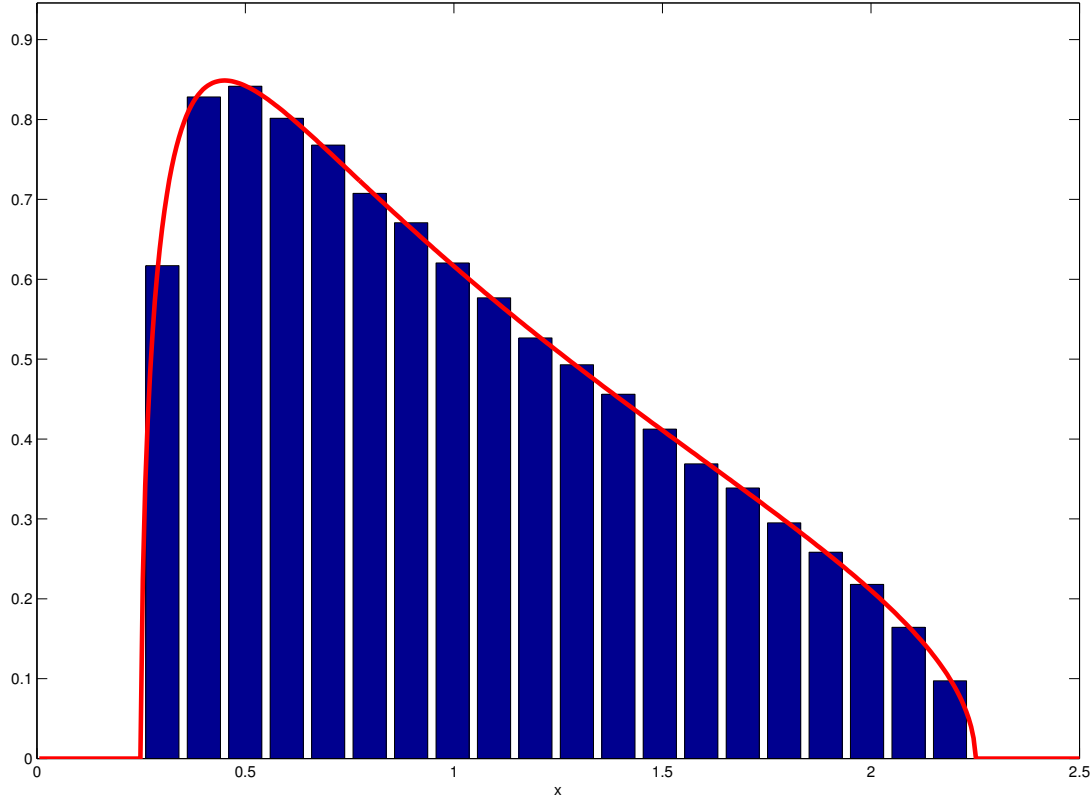
which can be solved as

$$G(z) = \frac{z + 1 - \lambda - \sqrt{(z - (1 + \lambda))^2 - 4\lambda}}{2z}$$

and thus

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

Wishart random matrix and Marchenko-Pastur distrib.



The Saga Begins ...

.... Consider Functions of **Several** Independent
Random Matrices

We are interested in the limiting eigenvalue distribution of

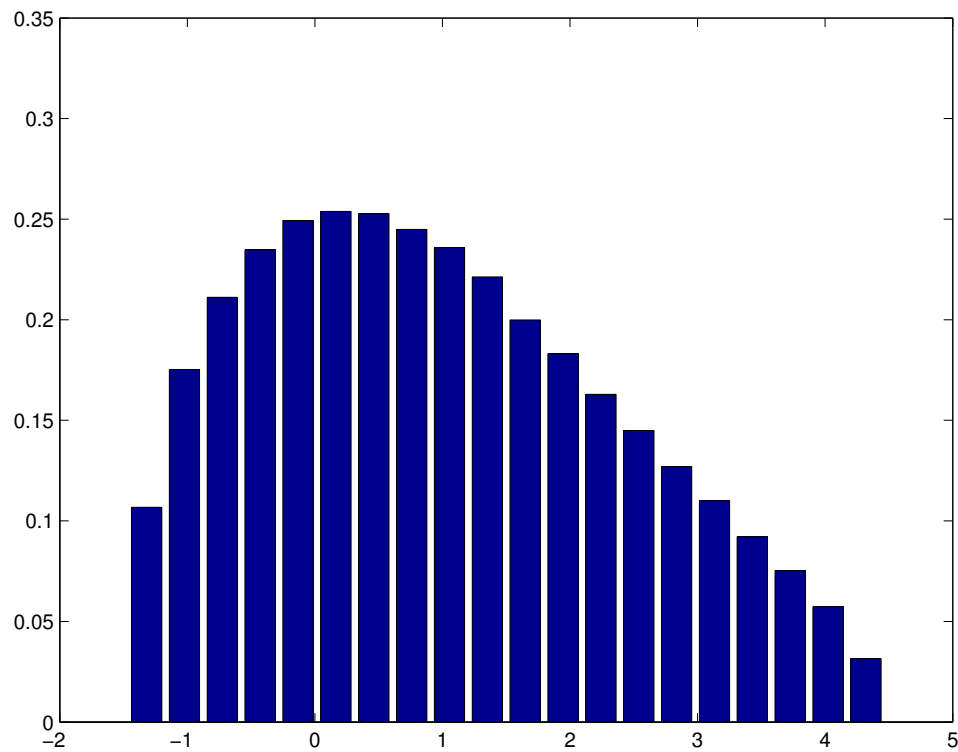
functions of **several** $N \times N$ random matrices for $N \rightarrow \infty$.

Typical phenomena:

- almost sure convergence to a deterministic limit eigenvalue distribution

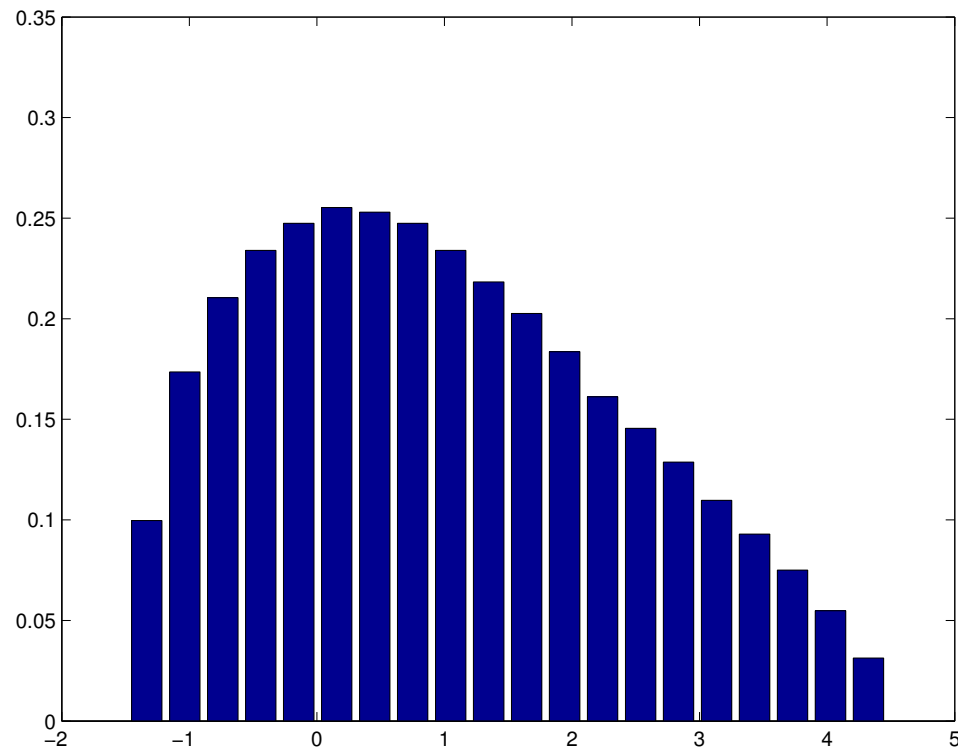


Wigner + Wishart random matrices, $N = 3000$



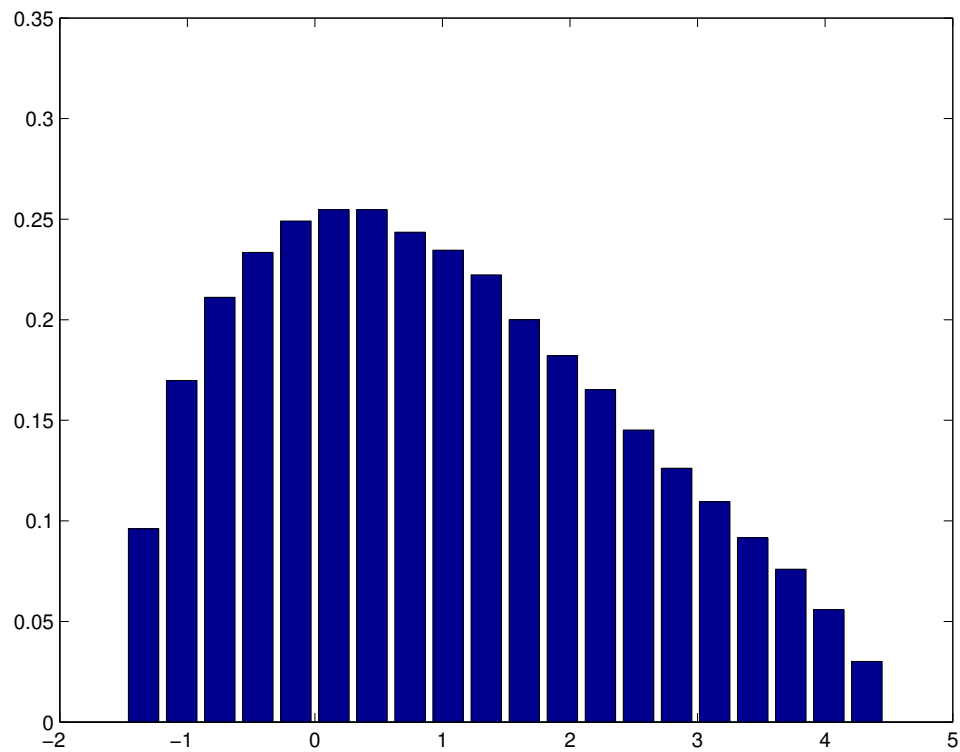
realization 1

Wigner + Wishart random matrices, $N = 3000$



realization 2

Wigner + Wishart random matrices, $N = 3000$



realization 3

We are interested in the limiting eigenvalue distribution of

functions of **several** $N \times N$ random matrices for $N \rightarrow \infty$.

Typical phenomena:

- almost sure convergence to a deterministic limit eigenvalue distribution
- this limit distribution can be effectively calculated **only in very simple situations**

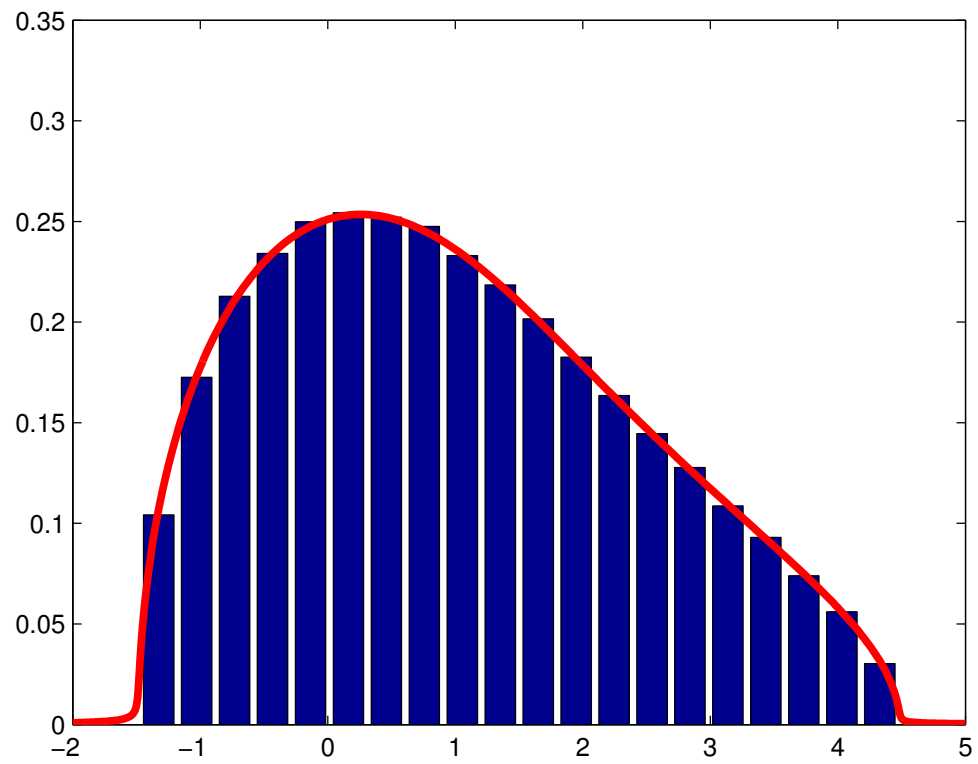
For simple situations one can derive equations for the Cauchy transform of the limiting eigenvalue distribution; those can usually not be solved explicitly; however, as fixed point equations they have a good analytic behaviour and can be solved numerically by iteration algorithms

Wigner + Wishart: For $G(z) := G_{\text{Wigner+Wishart}}(z)$ one finds the fixed point equation (in subordination form)

$$G(z) = G_{\text{Wishart}}(z - G(z)),$$

which can be easily solved by iteration.

Wigner + Wishart random matrices, $N = 3000$



Results for Calculations of the Limit Eigenvalue Distribution

- Marchenko, Pastur 1967: general Wishart matrices ADA^*
- Pastur 1972: deterministic + Wigner (deformed semicircle)
- Speicher, Nica 1998; Vasilchuk 2003: commutator or anti-commutator: $X_1X_2 \pm X_2X_1$
- more general models in wireless communications (Tulino, Verdu 2004; Couillet, Debbah, Silverstein 2011):

$$RADA^*R \quad \text{or} \quad \sum_i R_i A_i D_i A_i^* R_i$$

The Quest:

But What About More Complicated or Even General Selfadjoint Polynomials

.... something like

$$P(X, Y) = XY + YX + X^2$$

or

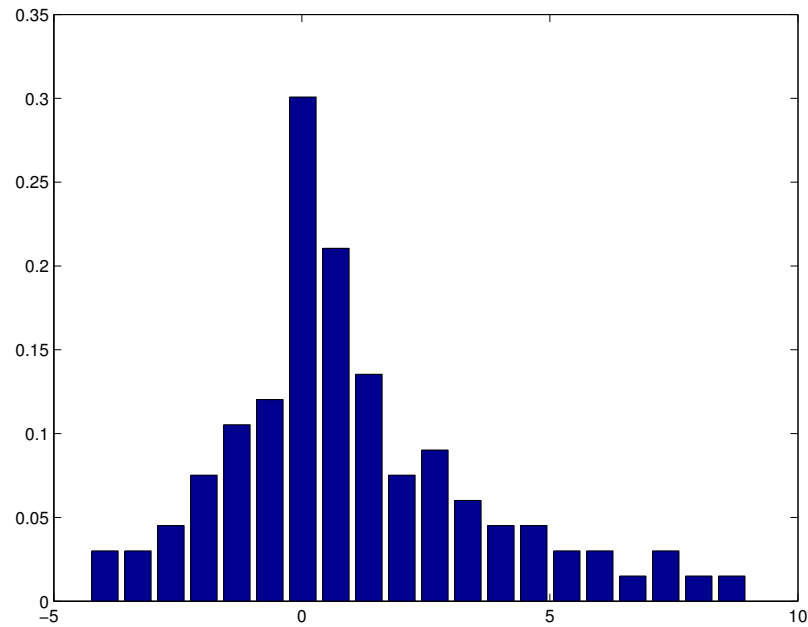
$$P(X_1, X_2, X_3) = X_1X_2X_1 + X_2X_3X_2 + X_3X_1X_3$$

or even just

$$P(X_1, \dots, X_k) \quad P \text{ selfadjoint polynomial}$$

$$P(X, Y) = XY + YX + X^2$$

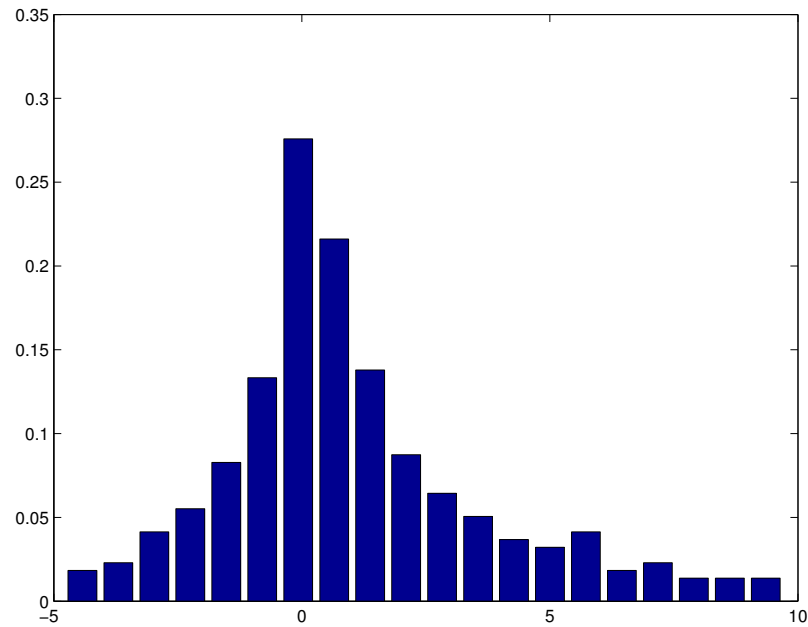
for independent X, Y ; X is Wigner, Y is Wishart



$N = 100$

$$P(X, Y) = XY + YX + X^2$$

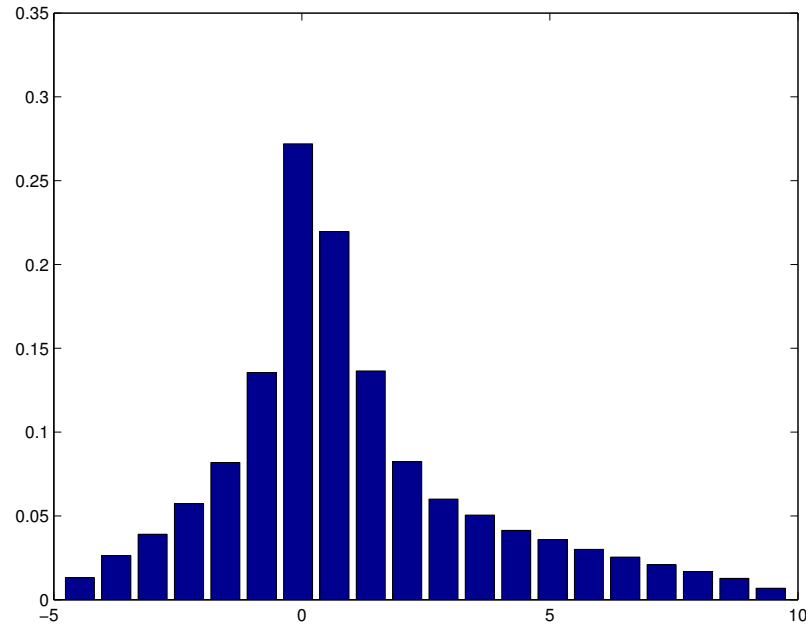
for independent X, Y ; X is Wigner, Y is Wishart



$N = 300$

$$P(X, Y) = XY + YX + X^2$$

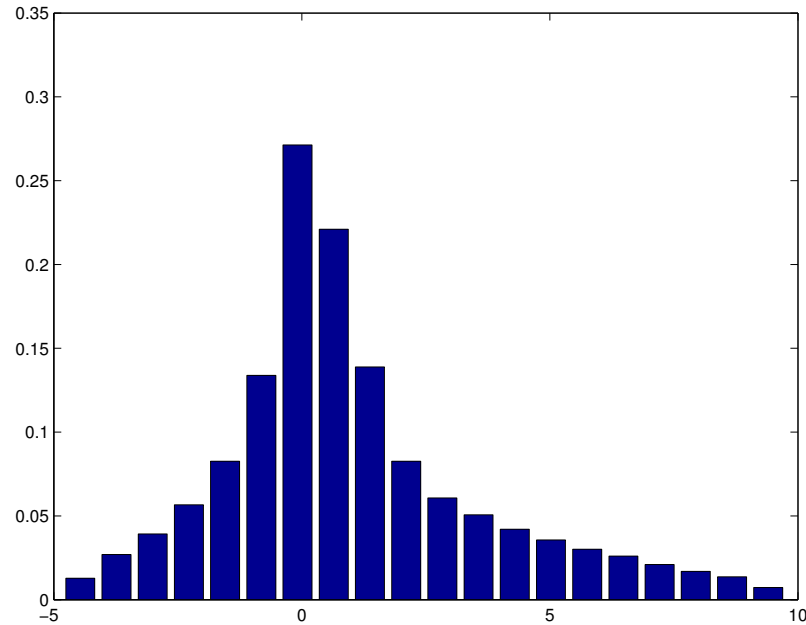
for independent X, Y ; X is Wigner, Y is Wishart



$N = 3000$, Realization 1

$$P(X, Y) = XY + YX + X^2$$

for independent X, Y ; X is Wigner, Y is Wishart



$N = 3000$, Realization 2

**The Hero:
Free Probability Theory**

Definition of Freeness (Voiculescu 1985)

Let (\mathcal{A}, φ) be **non-commutative probability space**, i.e., \mathcal{A} is a unital algebra and $\varphi : \mathcal{A} \rightarrow \mathbb{C}$ is unital linear functional (i.e., $\varphi(1) = 1$)

Unital subalgebras \mathcal{A}_i ($i \in I$) are **free** or **freely independent**, if $\varphi(a_1 \cdots a_n) = 0$ whenever

- $a_i \in \mathcal{A}_{j(i)}$, $j(i) \in I \quad \forall i$, $j(1) \neq j(2) \neq \cdots \neq j(n)$
- $\varphi(a_i) = 0 \quad \forall i$

Random variables $x_1, \dots, x_n \in \mathcal{A}$ are freely independent, if their generated unital subalgebras $\mathcal{A}_i := \text{algebra}(1, x_i)$ are so.

What Is Freeness?

Freeness between x and y is an infinite set of equations relating various moments in x and y :

$$\varphi\left(p_1(x)q_1(y)p_2(x)q_2(y)\cdots\right) = 0$$

Basic observation: free independence between x and y is actually a **rule for calculating mixed moments** in x and y from the moments of x and the moments of y :

$$\varphi\left(x^{m_1}y^{n_1}x^{m_2}y^{n_2}\cdots\right) = \text{polynomial}\left(\varphi(x^i), \varphi(y^j)\right)$$

If x and y are freely independent, then we have

$$\varphi(x^m y^n) = \varphi(x^m) \cdot \varphi(y^n)$$

$$\varphi(x^{m_1} y^n x^{m_2}) = \varphi(x^{m_1+m_2}) \cdot \varphi(y^n)$$

but also

$$\varphi(xyxy) = \varphi(x^2) \cdot \varphi(y)^2 + \varphi(x)^2 \cdot \varphi(y^2) - \varphi(x)^2 \cdot \varphi(y)^2$$

Free independence is a rule for calculating mixed moments, analogous to the concept of independence for random variables.

Note: free independence is a different rule from classical independence; free independence occurs typically for **non-commuting random variables**, like operators on Hilbert spaces

Consequence: Distribution of Polynomial in Freely Independent Variables Is Determined by Distributions of Their Variables

If x_1, \dots, x_k are freely independent, and p is a polynomial in k variables, then the distribution of $p(x_1, \dots, x_k)$ is determined by the moments of each of the x_i and by the fact that they are freely independent.

Where Does Free Independence Show Up?

- generators of the free group in the corresponding free group von Neumann algebras $L(\mathbb{F}_n)$
- creation and annihilation operators on full Fock spaces
-

Where Does Free Independence Show Up?

- generators of the free group in the corresponding free group von Neumann algebras $L(\mathbb{F}_n)$
- creation and annihilation operators on full Fock spaces
- **for many classes of random matrices**

Asymptotic Freeness of Random Matrices

Basic result of Voiculescu (1991):

Large classes of independent random matrices (like Wigner or Wishart matrices) become asymptotically freely independent, with respect to $\varphi = \frac{1}{N} \text{Tr}$, if $N \rightarrow \infty$.

Consequence: Reduction of Our random Matrix Problem to the Problem of Polynomial in Freely Independent Variables

If the random matrices X_1, \dots, X_k are asymptotically freely independent, then the distribution of a polynomial $p(X_1, \dots, X_k)$ is asymptotically given by the distribution of $p(x_1, \dots, x_k)$, where

- x_1, \dots, x_k are freely independent variables, and
- the distribution of x_i is the asymptotic distribution of X_i

Can We Actually Calculate Polynomials in Freely Independent Variables?

Free probability can deal effectively with simple polynomials

- the sum of variables (Voiculescu 1986, R -transform)

$$p(x, y) = x + y$$

- the product of variables (Voiculescu 1987, S -transform)

$$p(x, y) = xy \quad (= \sqrt{x}y\sqrt{x})$$

- the commutator of variables (Nica, Speicher 1998)

$$p(x, y) = xy - yx$$

There is no hope to calculate effectively more complicated or general polynomials in freely independent variables with usual free probability theory!

**The Superhero:
Operator-Valued Extension of Free
Probability**

Let $\mathcal{B} \subset \mathcal{A}$. A linear map

$$E : \mathcal{A} \rightarrow \mathcal{B}$$

is a **conditional expectation** if

$$E[b] = b \quad \forall b \in \mathcal{B}$$

and

$$E[b_1 a b_2] = b_1 E[a] b_2 \quad \forall a \in \mathcal{A}, \quad \forall b_1, b_2 \in \mathcal{B}$$

An **operator-valued probability space** consists of $\mathcal{B} \subset \mathcal{A}$ and a conditional expectation $E : \mathcal{A} \rightarrow \mathcal{B}$

Consider an operator-valued probability space $E : \mathcal{A} \rightarrow \mathcal{B}$.

Random variables $x_i \in \mathcal{A}$ ($i \in I$) are **freely independent with respect to E** (or operator-valued freely independent) if

$$E[a_1 \cdots a_n] = 0$$

whenever $a_i \in \mathcal{B}\langle x_{j(i)} \rangle$ are polynomials in some $x_{j(i)}$ with coefficients from \mathcal{B} and

$$E[a_i] = 0 \quad \forall i \quad \text{and} \quad j(1) \neq j(2) \neq \cdots \neq j(n).$$

Calculation Rule for Mixed Moments

For operator-valued freely independent variables, one has analogous formulas as in scalar-valued case, ...

The formula

$$\varphi(xyxy) = \varphi(xx)\varphi(y)\varphi(y) + \varphi(x)\varphi(x)\varphi(yy) - \varphi(x)\varphi(y)\varphi(x)\varphi(y)$$

has now to be written as

$$E[xyxy] = E[xE[y]x] \cdot E[y] + E[x] \cdot E[yE[x]y] - E[x]E[y]E[x]E[y]$$

Can We Actually Calculate Polynomials in Operator-Valued Freely Independent Variables?

Again, in principle all operator-valued polynomials in freely independent variables are determined, but effectively we can again only deal with simple polynomials:

- **the sum of variables**

Voiculescu 1995

Belinschi, Mai, Speicher 2012

- **the product of variables**

Voiculescu 1995; Dykema 2006

Belinschi, Speicher, Treilhard, Vargas 2012

**The Miracle:
The Linearization Trick**

Operator-Valued Polynomials Are Matrices of Polynomials

Operator-valued polynomials in variables x_1, \dots, x_k are matrices with entries given by polynomials in those random variables:

$$\begin{pmatrix} p_{11}(x_1, \dots, x_k) & \cdots & p_{1r}(x_1, \dots, x_k) \\ \vdots & \cdots & \vdots \\ p_{r1}(x_1, \dots, x_k) & \cdots & p_{rr}(x_1, \dots, x_k) \end{pmatrix}$$

The Linearization Philosophy:

In order to understand matrices of polynomials it suffices to understand (bigger) matrices of **linear** polynomials.

In particular, in order to understand polynomials in non-commuting variables, it suffices to understand matrices of **linear** polynomials in those variables.

- Voiculescu 1987: motivation
- Haagerup, Thorbjørnsen 2005: largest eigenvalue
- Anderson 2012: the selfadjoint version

The selfadjoint linearization of

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

This means: the Cauchy transform $G_p(z)$ of $p = xy + yx + x^2$ is given as the (1,1)-entry of the operator-valued (3×3 matrix) Cauchy transform of \hat{p} :

$$G_{\hat{p}}(b) = \text{id} \otimes \varphi \left[(b - \hat{p})^{-1} \right] = \begin{pmatrix} G_p(z) & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \quad \text{for} \quad b = \begin{pmatrix} z & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

But

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

with

$$\hat{x} = \begin{pmatrix} 0 & x & \frac{x}{2} \\ x & 0 & 0 \\ \frac{x}{2} & 0 & 0 \end{pmatrix} \quad \text{and} \quad \hat{y} = \begin{pmatrix} 0 & 0 & y \\ 0 & 0 & -1 \\ y & -1 & 0 \end{pmatrix}.$$

So \hat{p} is just the sum of two operator-valued variables

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

where we understand the operator-valued distributions of \hat{x} and of \hat{y} .

Are \hat{x} and \hat{y} freely independent?

Another Miracle

Matrices of Freely Independent Variables are matrix-valued Freely Independent

If x and y are freely independent with respect to φ , then for any polynomials p_{ij} in x and any polynomials q_{kl} in y one has:

$$\begin{pmatrix} p_{11}(x) & \dots & p_{1r}(x) \\ \vdots & \ddots & \vdots \\ p_{r1}(x) & \dots & p_{rr}(x) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} q_{11}(y) & \dots & q_{1r}(y) \\ \vdots & \ddots & \vdots \\ q_{r1}(y) & \dots & q_{rr}(y) \end{pmatrix}$$

are free with respect to

$$\text{id} \otimes \varphi$$

**The Final Battle:
Algorithm and Calculation for Arbitrary
Selfadjoint Polynomial in Freely
Independent Variables**

Input: $p(x, y)$, $G_x(z)$, $G_y(z)$



Linearize $p(x, y)$ to $\hat{p} = \hat{x} + \hat{y}$



$G_{\hat{x}}(b)$ out of $G_x(z)$ and $G_{\hat{y}}(b)$ out of $G_y(z)$



Get $w_1(b)$ as the fixed point of the iteration
 $w \mapsto G_{\hat{y}}(b + G_{\hat{x}}(w)^{-1} - w)^{-1} - (G_{\hat{x}}(w)^{-1} - w)$

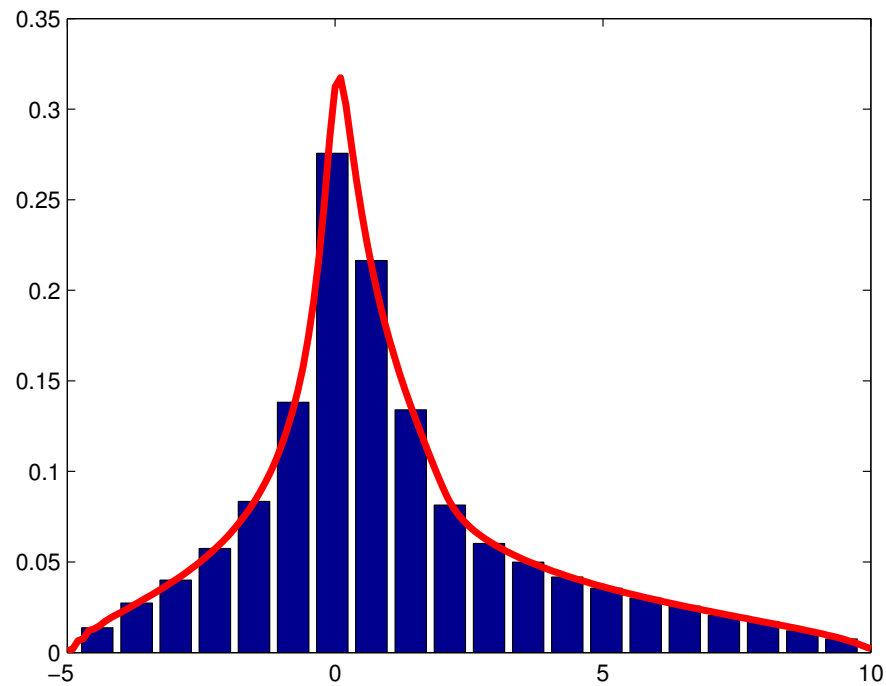


$G_{\hat{p}}(b) = G_{\hat{x}}(w_1(b))$



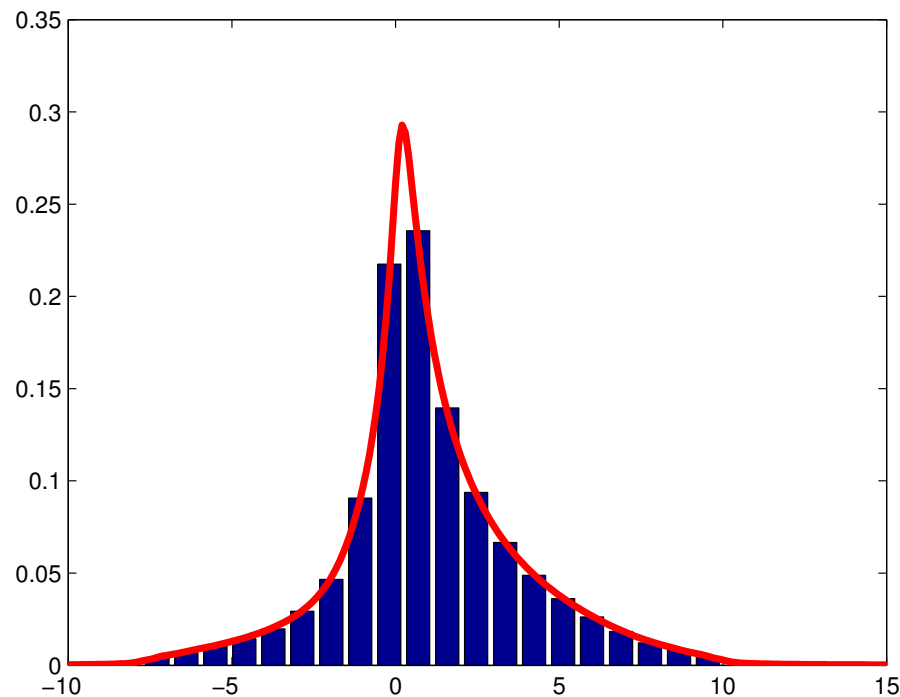
Recover $G_p(z)$ as one entry of $G_{\hat{p}}(b)$

$P(X, Y) = XY + YX + X^2$
for independent X, Y ; X is Wigner and Y is Wishart



$p(x, y) = xy + yx + x^2$
for free x, y ; x is semicircular and y is Marchenko-Pastur

$P(X_1, X_2, X_3) = X_1 X_2 X_1 + X_2 X_3 X_2 + X_3 X_1 X_3$
 for independent X_1, X_2, X_3 ; X_1, X_2 Wigner, X_3 Wishart



$p(x_1, x_2, x_3) = x_1 x_2 x_1 + x_2 x_3 x_2 + x_3 x_1 x_3$
 for free x_1, x_2, x_3 ; x_1, x_2 semicircular, x_3 Marchenko-Pastur

The Happy End

Theorem (Belinschi, Mai, Speicher 2012):

Combining the selfadjoint linearization trick with our new analysis of operator-valued free convolution we can provide an efficient and analytically controllable algorithm for calculating the asymptotic eigenvalue distribution of

- **any selfadjoint polynomial in asymptotically free random matrices.**