Selfadjoint Polynomials in Independent Random Matrices

Roland Speicher
Universität des Saarlandes
Saarbrücken

We are interested in the limiting eigenvalue distribution of an

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

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 $G : \mathbb{C}^+ \to \mathbb{C}^-$ and we can recover μ from G by **Stieltjes inversion formula**

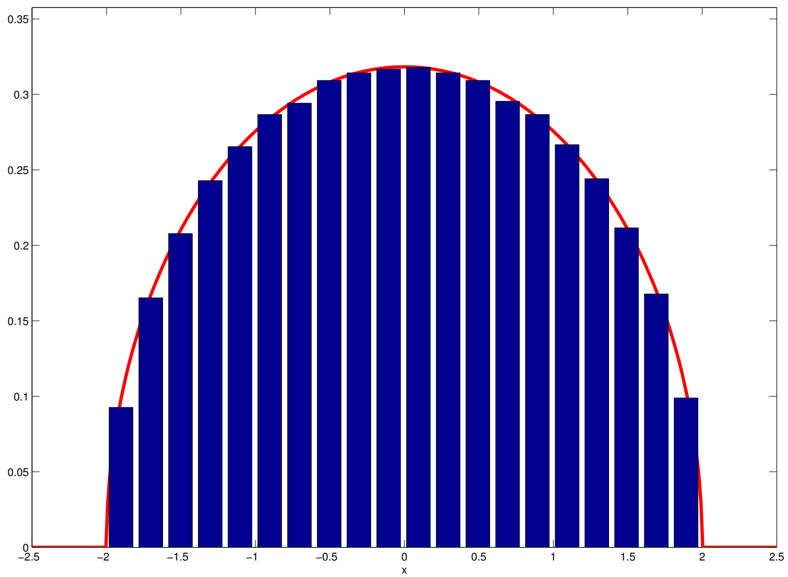
$$d\mu(t) = -\frac{1}{\pi} \lim_{\varepsilon \to 0} \Im G(t + i\varepsilon) dt$$

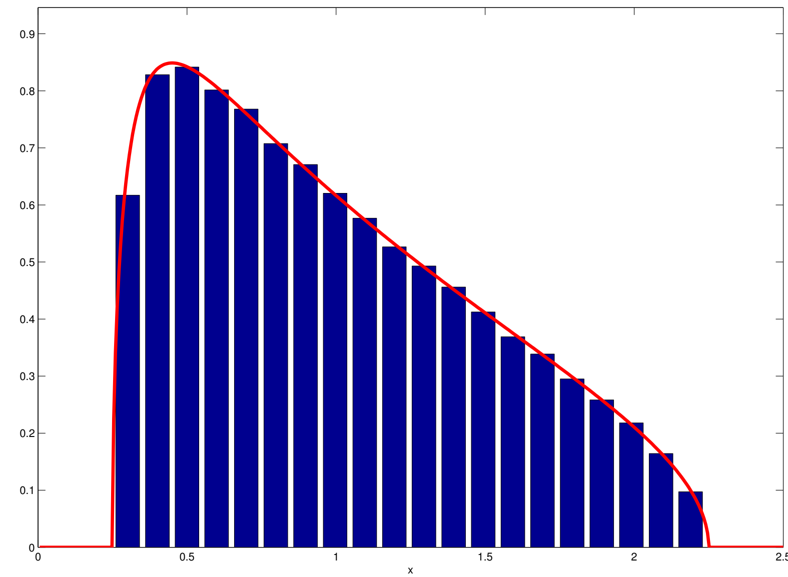
Wigner random matrix
and
Wigner's semicircle

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

Wishart random matrix
and
Marchenko-Pastur distribution

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





We are now interested in the limiting eigenvalue distribution of

general selfadjoint polynomials $p(X_1, \dots, X_k)$

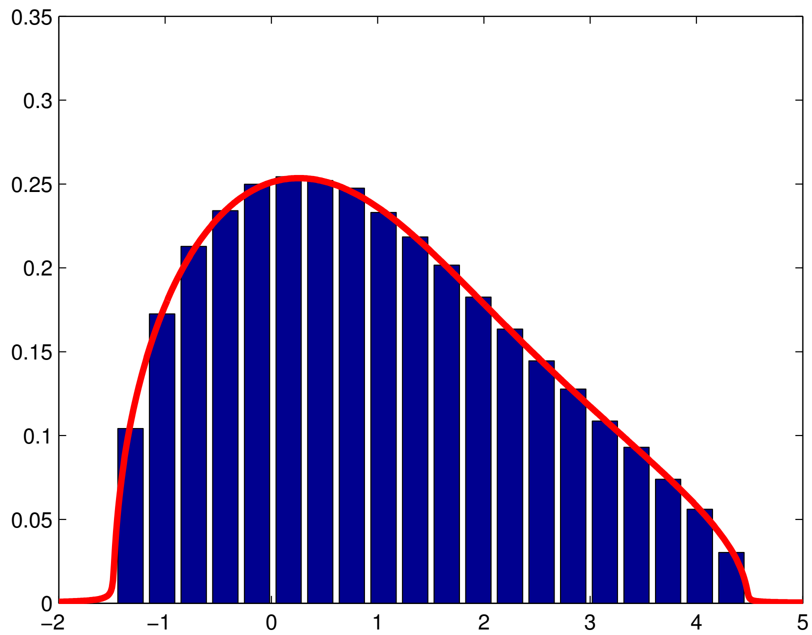
of **several** independent $N \times N$ random matrices $X_1, \dots, X_k$

Typical phenomena:

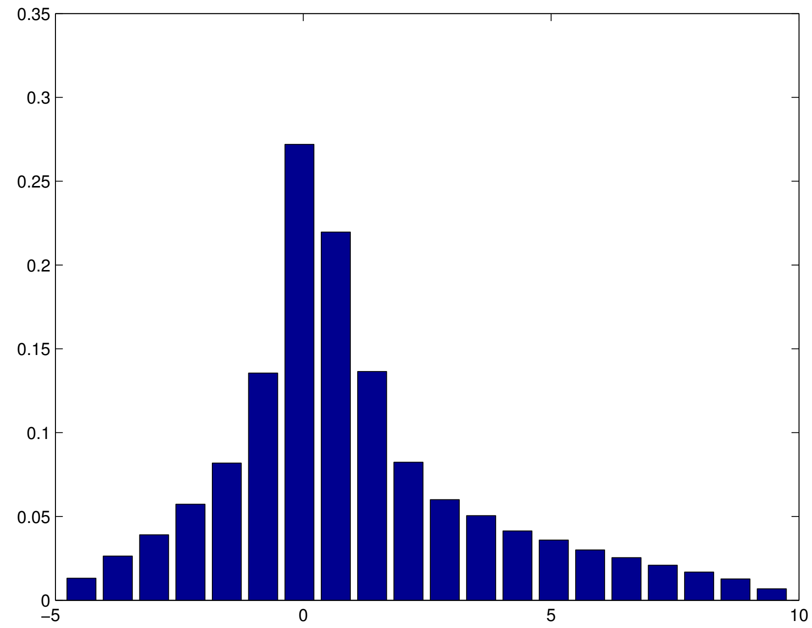
- almost sure convergence to a deterministic limit eigenvalue distribution

- this limit distribution can be effectively calculated **only in very simple situations** for X Wigner, Y Wishart

$$p(X,Y) = X + Y$$
$$G(z) = G_{\text{Wishart}}(z - G(z))$$



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



Existing 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):

$$RADA^*R^* \qquad \text{or} \qquad \sum_i R_iA_iD_iA_i^*R_i^*$$

Asymptotic Freeness of Random Matrices

Basic result of Voiculescu (1991):

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

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)

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

- 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 ...

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

...but there is a possible way around this: linearize the problem!!!

The Linearization Philosophy:

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 (based on Schur complement)

Consider a polynomial p in non-commuting variables x and y. A **linearization** of p is an $N \times N$ matrix (with $N \in \mathbb{N}$) of the form

$$\hat{p} = \begin{pmatrix} 0 & u \\ v & Q \end{pmatrix},$$

where

- u, v, Q are matrices of the following sizes: u is $1 \times (N-1)$; v is $(N-1) \times N$; and Q is $(N-1) \times (N-1)$

- each entry of u, v, Q is a polynomial in x and y, each of degree ≤ 1

- Q is invertible and we have

$$p = -uQ^{-1}v$$

Let

$$\widehat{p} = \begin{pmatrix} 0 & u \\ v & Q \end{pmatrix} \qquad \text{be a linearization of } p.$$

For $z \in \mathbb{C}$ put $b = \begin{pmatrix} z & 0 \\ 0 & 0 \end{pmatrix}$

Then we have

$$b - \widehat{p} = \begin{pmatrix} z & -u \\ -v & -Q \end{pmatrix} = \begin{pmatrix} 1 & uQ^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z - p & 0 \\ 0 & -Q \end{pmatrix} \begin{pmatrix} 1 & 0 \\ Q^{-1}v & 1 \end{pmatrix}$$

hence

$$z - p \text{ invertible} \qquad \Longleftrightarrow \qquad b - \widehat{p} \text{ invertible}$$

Actually,

$$(b-\widehat{p})^{-1} = \begin{pmatrix} 1 & 0 \\ -Q^{-1}v & 1 \end{pmatrix} \begin{pmatrix} (z-p)^{-1} & 0 \\ 0 & -Q^{-1} \end{pmatrix} \begin{pmatrix} 1 & -uQ^{-1} \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} (z-p)^{-1} & * \\ * & * \end{pmatrix}$$

and we can get

$$G_p(z) = \varphi((z-p)^{-1})$$

as the (1,1)-entry of the operator-valued Cauchy-transform

$$G_{\widehat{p}}(b) = \mathrm{id} \otimes \varphi((b-\widehat{p})^{-1}) = \begin{pmatrix} \varphi((z-p)^{-1}) & \varphi(*) \\ \varphi(*) & \varphi(*) \end{pmatrix}$$

Theorem (Anderson 2012): One has

- for each p there exists a linearization $\widehat{p}$ (with an explicit algorithm for finding those)

- if p is selfadjoint, then this $\widehat{p}$ is also selfadjoint

The selfadjoint linearization of

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

because we have

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

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 $\hat{p}$:

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

where

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

is a linear polynomial, but with matrix-valued coefficients.

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

In order to understand $\hat{p}$, we have to calculate the free convolution of

$$\hat{x} = \begin{pmatrix} 0 & 1 & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix} \otimes x \qquad \text{and} \qquad \hat{y} := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \otimes y$$

with respect to $E = \mathrm{id} \otimes \varphi$

E is not an expectation, but a conditional expectation (e.g., partial trace).

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

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

is a **conditional expectation** if

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

and

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

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

Consider an operator-valued probability space $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$.

Random variables $x_i \in \mathcal{A}$ $(i \in I)$ are **free with respect to E** (or **free with amalgamation over $\mathcal{B}$)** 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

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

Consider $E : \mathcal{A} \to \mathcal{B}$.

Define **free cumulants**

$$\kappa_n^{\mathcal{B}} : \mathcal{A}^n \to \mathcal{B}$$

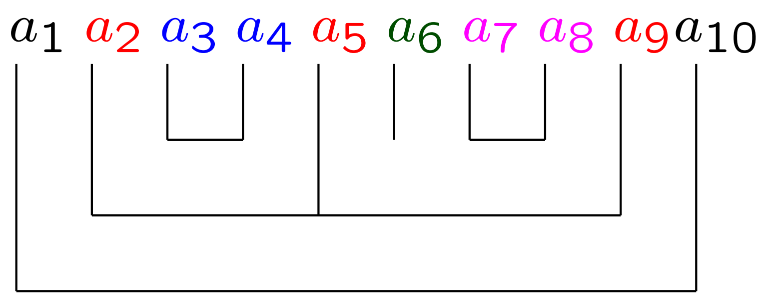
by

$$E[a_1 \cdots a_n] = \sum_{\pi \in NC(n)} \kappa_\pi^{\mathcal{B}}[a_1, \dots, a_n]$$

- arguments of $\kappa_\pi^{\mathcal{B}}$ are distributed according to blocks of π
- but now: cumulants are nested inside each other according to nesting of blocks of π

Example:

$$\pi = \big\{\{1,10\},\{2,5,9\},\{3,4\},\{6\},\{7,8\}\big\} \in NC(10),$$



$$\kappa_\pi^{\mathcal{B}}[a_1,\dots,a_{10}]$$

$$= \kappa_2^{\mathcal{B}}\Big(a_1 \cdot \kappa_3^{\mathcal{B}}\big(a_2 \cdot \kappa_2^{\mathcal{B}}(a_3,a_4), a_5 \cdot \kappa_1^{\mathcal{B}}(a_6) \cdot \kappa_2^{\mathcal{B}}(a_7,a_8), a_9\big), a_{10}\Big)$$

For $a \in \mathcal{A}$ define its **operator-valued Cauchy transform**

$$G_a(b) := E[\frac{1}{b-a}] = \sum_{n \geq 0} E[b^{-1}(ab^{-1})^n]$$

and **operator-valued R-transform**

$$\begin{aligned} R_a(b) : &= \sum_{n \geq 0} \kappa^{\mathcal{B}}_{n+1}(ab, ab, \dots, ab, a) \\ &= \kappa^{\mathcal{B}}_1(a) + \kappa^{\mathcal{B}}_2(ab, a) + \kappa^{\mathcal{B}}_3(ab, ab, a) + \cdots \end{aligned}$$

Then

$$bG(b) = 1 + R(G(b)) \cdot G(b) \qquad \text{or} \qquad G(b) = \frac{1}{b - R(G(b))}$$

If x and y are free over $\mathcal{B}$, then

- mixed $\mathcal{B}$-valued cumulants in x and y vanish

- $R_{x+y}(b) = R_x(b) + R_y(b)$

- we have the subordination $G_{x+y}(z) = G_x(\omega(z))$

Theorem (Belinschi, Mai, Speicher 2013): Let x and y be selfadjoint operator-valued random variables free over B. Then there exists a Fréchet analytic map $\omega\colon \mathbb{H}^+(B) \to \mathbb{H}^+(B)$ so that

$$G_{x+y}(b) = G_x(\omega(b)) \text{ for all } b \in \mathbb{H}^+(B).$$

Moreover, if $b \in \mathbb{H}^+(B)$, then $\omega(b)$ is the unique fixed point of the map

$$f_b\colon \mathbb{H}^+(B) \to \mathbb{H}^+(B), \qquad f_b(w) = h_y(h_x(w) + b) + b,$$

and

$$\omega(b) = \lim_{n\to\infty} f_b^{\circ n}(w) \qquad \text{for any } w \in \mathbb{H}^+(B).$$

where

$$\mathbb{H}^+(B) := \{b \in B \mid (b - b^*)/(2i) > 0\}, \qquad h(b) := \frac{1}{G(b)} - b$$

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

Problem: How do we deal with a polynomial p in free variables?

Idea: Linearize the polynomial and use operator-valued convolution for the linearization $\hat{p}$!

The linearization of $p = xy + yx + x^2$ is given by

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

This means that the Cauchy transform $G_p(z)$ is given as the (1,1)-entry of the operator-valued (3×3 matrix) Cauchy transform of $\hat{p}$:

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

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 $\widehat{p}$ is just the sum of two operator-valued variables

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

- where we understand the operator-valued distributions of $\widehat{x}$ and of $\widehat{y}$

- **and $\hat{x}$ and $\hat{y}$ are operator-valued freely independent!**

So we can use operator-valued free convolution to calculate the operator-valued Cauchy transform of $\widehat{x} + \widehat{y}$.

So we can use operator-valued free convolution to calculate the operator-valued Cauchy transform of $\widehat{p} = \widehat{x} + \widehat{y}$ in the subordination form

$$G_{\widehat{p}}(b) = G_{\widehat{x}}(\omega(b)),$$

where $\omega(b)$ is the unique fixed point in the upper half plane $\mathbb{H}_+(M_3(\mathbb{C})$ of the iteration

$$w \mapsto G_{\widehat{y}}(b + G_{\widehat{x}}(w)^{-1} - w)^{-1} - (G_{\widehat{x}}(w)^{-1} - w)$$

Input: $\boldsymbol{p(x,y),\ G_x(z), G_y(z)}$

$\downarrow$

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

$\downarrow$

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

$\downarrow$

Get $w(b)$ as the fixed point of the iteration

$$w \mapsto G_{\widehat{y}}(b + G_{\widehat{x}}(w)^{-1} - w)^{-1} - (G_{\widehat{x}}(w)^{-1} - w)$$

$\downarrow$

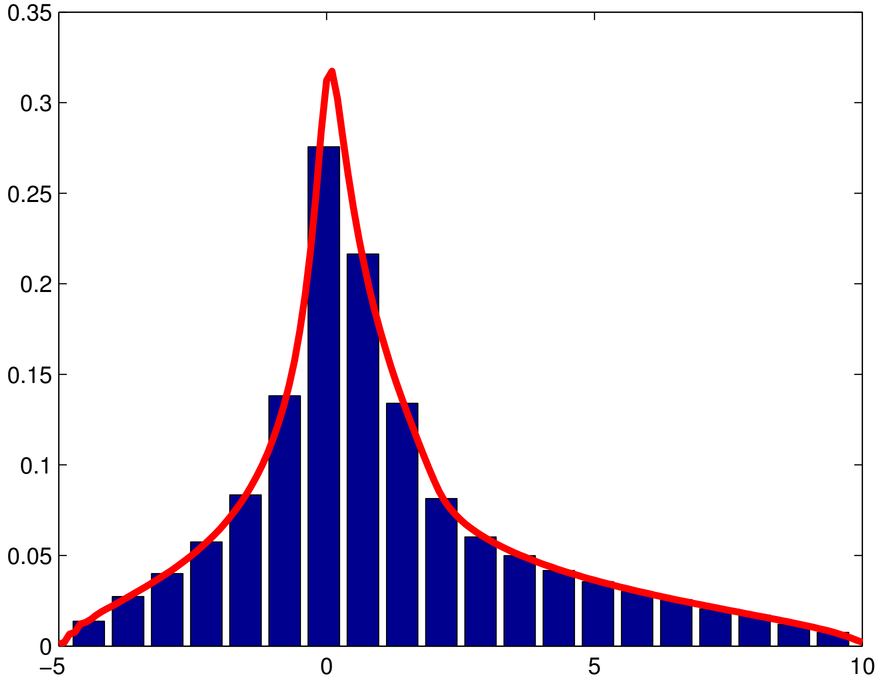
$$G_{\widehat{p}}(b) = G_{\widehat{x}}(\omega(b))$$

$\downarrow$

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

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

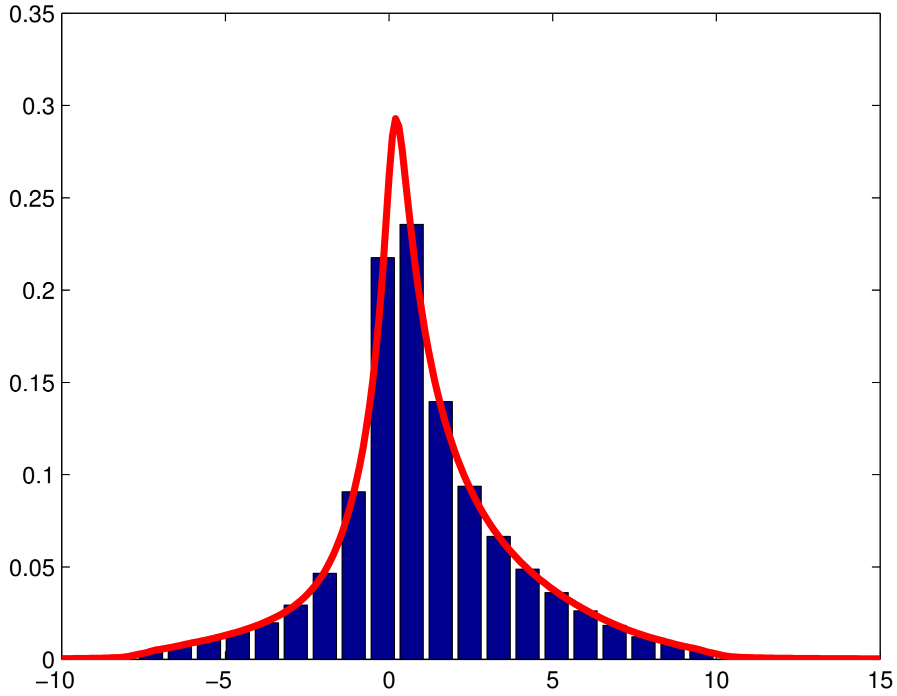
for independent X, Y; X is Wigner and Y is Wishart



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

for free x, y; x is semicircular and y is Marchenko-Pastur

$P(X_1, X_2, X_3) = X_1X_2X_1 + X_2X_3X_2 + X_3X_1X_3$
for independent X_1, X_2, X_3; X_1, X_2 Wigner, X_3 Wishart



$p(x_1, x_2, x_3) = x_1x_2x_1 + x_2x_3x_2 + x_3x_1x_3$
for free x_1, x_2, x_3; x_1, x_2 semicircular, x_3 Marchenko-Pastur

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.

Outlook: How about the case of non selfadjoint polynomials?

Drop in for Belinschi's talk on Friday!