Entry-wise application of non-linear functions on orthogonally invariant random matrices

Roland Speicher

Saarland University
Saarbrücken, Germany

joint work with Alexander Wendel
arXiv:2412.06943

supported by the DFG project "Free probability aspects of neural networks"

Question: What is the effect of applying a non-linear function to the entries of a random matrix?

Question: What is the effect of applying a non-linear function to the entries of a random matrix?

Given input

- random matrix $X_N = (x_{ij})_{i,j=1}^N$
- function $f : \mathbb{R} \to \mathbb{R}$ (for now, polynomial)

Question: What is the effect of applying a non-linear function to the entries of a random matrix?

Given input

- random matrix $X_N = (x_{ij})_{i,j=1}^N$
- function $f : \mathbb{R} \to \mathbb{R}$ (for now, polynomial)

Define the "non-linear" random matrix $Y_N = (y_{ij})$ by

$$y_{ij} = \begin{cases} \frac{1}{\sqrt{N}} f(\sqrt{N} x_{ij}), & i \neq j \\ 0, & i = j. \end{cases}$$

Question: What is the effect of applying a non-linear function to the entries of a random matrix?

Given input

- random matrix $X_N = (x_{ij})_{i,j=1}^N$
- function $f : \mathbb{R} \to \mathbb{R}$ (for now, polynomial)

Define the "non-linear" random matrix $Y_N = (y_{ij})$ by

$$y_{ij} = \begin{cases} \frac{1}{\sqrt{N}} f(\sqrt{N} x_{ij}), & i \neq j \\ 0, & i = j. \end{cases}$$

Question

How does this change the eigenvalue distribution, asymptotically, if $N \to \infty$?

What assumptions should we make about our random matrices X_N

previous work

- *kernel matrices:*
  El Karou 2010; Cheng and Singer 2013; Fan and Montanari 2019
- *covariance matrices:*
  Pennington and Worah 2017; Louart, Liao and Couillet 2018; Benigni and Péché 2021; Piccolo and Schröder 2021

What assumptions should we make about our random matrices X_N

previous work for quite specific random matrices

- *kernel matrices:*
  El Karou 2010; Cheng and Singer 2013; Fan and Montanari 2019
- *covariance matrices:*
  Pennington and Worah 2017; Louart, Liao and Couillet 2018; Benigni and Péché 2021; Piccolo and Schröder 2021

Our general setting

orthogonally invariant random matrices

Orthogonally invariant random matrices with limit distribution

Orthogonally invariant random matrices with limit distribution

Typical examples (say, symmetric real random matrices)

- given by density of the form

$$\exp\left(-N\mathsf{Tr}(q(X_N))\right) dX_N$$

- given as polynomials in independent GOEs

$$p(A_N, B_N, \dots, D_N)$$

Orthogonally invariant random matrices with limit distribution

Typical examples (say, symmetric real random matrices)

- given by density of the form

$$\exp\left(-N\mathsf{Tr}(q(X_N))\right) dX_N$$

- given as polynomials in independent GOEs

$$p(A_N, B_N, \dots, D_N)$$

Properties of this class of random matrices

- orthogonal invariance: entries of X_N have the same joint distribution as entries of OX_NO^T for any orthogonal O
- existence of scaled limit of correlations (c_n are classical cumulants)

$$\lim_{n\to\infty} N^{n-2} \cdot c_n(\mathsf{Tr}(X_N^{m_1}), \dots, \mathsf{Tr}(X_N^{m_n}))$$

Properties of this class of random matrices

- orthogonal invariance: entries of X_N have the same joint distribution as entries of OX_NO^T for any orthogonal O
- existence of scaled limit of correlations (c_n are classical cumulants)

$$\lim_{n\to\infty} N^{n-2} \cdot c_n(\mathrm{Tr}(X_N^{m_1}), \dots, \mathrm{Tr}(X_N^{m_n}))$$

Cyclic structure of the cumulants of entries

Properties of this class of random matrices

- orthogonal invariance: entries of X_N have the same joint distribution as entries of OX_NO^T for any orthogonal O
- existence of scaled limit of correlations (c_n are classical cumulants)

$$\lim_{n\to\infty} N^{n-2}\cdot c_n(\mathrm{Tr}(X_N^{m_1}),\dots,\mathrm{Tr}(X_N^{m_n}))$$

This implies for cumulants $c_n(x_{i_1j_1},x_{i_2j_2},\dots,x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero

Cyclic structure of the cumulants of entries

Properties of this class of random matrices

- orthogonal invariance: entries of X_N have the same joint distribution as entries of OX_NO^T for any orthogonal O
- existence of scaled limit of correlations (c_n are classical cumulants)

$$\lim_{n\to\infty} N^{n-2} \cdot c_n(\mathrm{Tr}(X_N^{m_1}), \dots, \mathrm{Tr}(X_N^{m_n}))$$

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

Cyclic structure of the cumulants of entries

Properties of this class of random matrices

- orthogonal invariance: entries of X_N have the same joint distribution as entries of OX_NO^T for any orthogonal O
- existence of scaled limit of correlations (c_n are classical cumulants)

$$\lim_{n\to\infty} N^{n-2} \cdot c_n(\text{Tr}(X_N^{m_1}), \dots, \text{Tr}(X_N^{m_n}))$$

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Examples

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \ldots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \ldots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Examples

- no cyclic structure: $c_4(x_{12}, x_{23}, x_{34}, x_{43}) = 0$

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Examples

- no cyclic structure: $c_4(x_{12}, x_{23}, x_{34}, x_{43}) = 0$
- one cycle: $c_4(x_{12}, x_{23}, x_{34}, x_{41}) \sim N^{-3}$

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1 j_1}, x_{i_2 j_2}, \dots, x_{i_n j_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1 i_2}, x_{i_2 i_3}, \dots, x_{i_n i_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Examples

- no cyclic structure: $c_4(x_{12}, x_{23}, x_{34}, x_{43}) = 0$
- one cycle: $c_4(x_{12}, x_{23}, x_{34}, x_{41}) \sim N^{-3}$
- two cycles: $c_4(x_{12}, x_{21}, x_{34}, x_{43}) \sim N^{-4}$

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Examples

- no cyclic structure: $c_4(x_{12}, x_{23}, x_{34}, x_{43}) = 0$
- one cycle: $c_4(x_{12}, x_{23}, x_{34}, x_{41}) \sim N^{-3}$
- two cycles: $c_4(x_{12}, x_{21}, x_{34}, x_{43}) \sim N^{-4}$
- one cycle with subcycles: $c_4(x_{12}, x_{21}, x_{12}, x_{21}) \sim N^{-3}$

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1 j_1}, x_{i_2 j_2}, \dots, x_{i_n j_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1 i_2}, x_{i_2 i_3}, \dots, x_{i_n i_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Examples

- no cyclic structure: $c_4(x_{12}, x_{23}, x_{34}, x_{43}) = 0$
- one cycle: $c_4(x_{12}, x_{23}, x_{34}, x_{41}) \sim N^{-3}$
- two cycles: $c_4(x_{12}, x_{21}, x_{34}, x_{43}) \sim N^{-4}$
- one cycle with subcycles: $c_4(x_{12}, x_{21}, x_{12}, x_{21}) \sim N^{-3}$
  has leading order N^{-3}, and a subleading contribution of order N^{-4}

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1 j_1}, x_{i_2 j_2}, \dots, x_{i_n j_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1 i_2}, x_{i_2 i_3}, \dots, x_{i_n i_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Meaning of cycle structure

Note that there is some freedom in arranging the entries of the cumulants

- arguments of classical cumulants can be permuted
- for our symmetric matrices X_N we have: $x_{ij} = x_{ji}$

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \ldots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \ldots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Meaning of cycle structure

Note that there is some freedom in arranging the entries of the cumulants

- arguments of classical cumulants can be permuted
- for our symmetric matrices X_N we have: $x_{ij} = x_{ji}$

e.g., $c_4(x_{12}, x_{34}, x_{23}, x_{14})$

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Meaning of cycle structure

Note that there is some freedom in arranging the entries of the cumulants

- arguments of classical cumulants can be permuted
- for our symmetric matrices X_N we have: $x_{ij} = x_{ji}$

e.g., $c_4(x_{12}, x_{34}, x_{23}, x_{14}) = c_4(x_{12}, x_{23}, x_{34}, x_{14})$

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Meaning of cycle structure

Note that there is some freedom in arranging the entries of the cumulants

- arguments of classical cumulants can be permuted
- for our symmetric matrices X_N we have: $x_{ij} = x_{ji}$

e.g., $c_4(x_{12}, x_{34}, x_{23}, x_{14}) = c_4(x_{12}, x_{23}, x_{34}, x_{14}) = c_4(x_{12}, x_{23}, x_{34}, x_{41})$

has also cyclic structure.

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Where to find such statements in the literature?

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1 j_1}, x_{i_2 j_2}, \ldots, x_{i_n j_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1 i_2}, x_{i_2 i_3}, \ldots, x_{i_n i_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Where to find such statements in the literature?

- Zinn-Justin 1999; Collins 2003; Guionnet, Maida 2005; Collins, Mingo, Sniady, Speicher 2007

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Where to find such statements in the literature?

- Zinn-Justin 1999; Collins 2003; Guionnet, Maida 2005; Collins, Mingo, Sniady, Speicher 2007
- specifically in the orthogonal case: Capitaine, Casalis 2007

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Where to find such statements in the literature?

- Zinn-Justin 1999; Collins 2003; Guionnet, Maida 2005; Collins, Mingo, Sniady, Speicher 2007
- specifically in the orthogonal case: Capitaine, Casalis 2007
- probably in many physics papers on a non-rigorous level???

Cyclic structure of the cumulants of entries

This implies for cumulants $c_n(x_{i_1j_1}, x_{i_2j_2}, \dots, x_{i_nj_n})$ of entries

- only indices with cycle structure are non-zero
- leading order for one full cycle is

$$c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$$

- leading order for r cycles is N^{2-r-n}

Where to find such statements in the literature?

- Zinn-Justin 1999; Collins 2003; Guionnet, Maida 2005; Collins, Mingo, Sniady, Speicher 2007
- specifically in the orthogonal case: Capitaine, Casalis 2007
- probably in many physics papers on a non-rigorous level???
- in particular: are there precise statements in the literature for the orthogonal case for the case of several cycles?

- only indices with cycle structure are non-zero
- leading order for one full cycle is $c_n(x_{i_1 i_2}, x_{i_2 i_3}, \dots, x_{i_n i_1}) \sim N^{1-n}$
- leading order for r cycles is N^{2-r-n}

Relation between free and classical cumulants

- only indices with cycle structure are non-zero
- leading order for one full cycle is $c_n(x_{i_1 i_2}, x_{i_2 i_3}, \dots, x_{i_n i_1}) \sim N^{1-n}$
- leading order for r cycles is N^{2-r-n}

Put $\kappa_n := \lim_{N\to\infty} N^{n-1} c_n(x_{i_1 i_2}, x_{i_2 i_3}, \dots, x_{i_n i_1})$

Relation between free and classical cumulants

- only indices with cycle structure are non-zero
- leading order for one full cycle is $c_n(x_{i_1 i_2}, x_{i_2 i_3}, \dots, x_{i_n i_1}) \sim N^{1-n}$
- leading order for r cycles is N^{2-r-n}

Put $\kappa_n := \lim_{N\to\infty} N^{n-1} c_n(x_{i_1 i_2}, x_{i_2 i_3}, \dots, x_{i_n i_1})$

Then
$$\mathbb{E}[\mathrm{tr}(X_N^n)] = \frac{1}{N} \sum_{i_1,\dots,i_n}^{N} \mathbb{E}[x_{i_1 i_2} x_{i_2 i_3} \cdots x_{i_n i_1}]$$
$$= \frac{1}{N} \sum_{i_1,\dots,i_n}^{N} \sum_{\pi \in P(n)} c_\pi[x_{i_1 i_2}, x_{i_2 i_3}, \dots, x_{i_n i_1}]$$

goes in leading order over to
$$\varphi(X^n) = \sum_{\pi \in NC(n)} \kappa_\pi;$$

Relation between free and classical cumulants

- only indices with cycle structure are non-zero
- leading order for one full cycle is $c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}) \sim N^{1-n}$
- leading order for r cycles is N^{2-r-n}

Put $\kappa_n := \lim_{N\to\infty} N^{n-1} c_n(x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1})$

Then
$$\mathbb{E}[\mathrm{tr}(X_N^n)] = \frac{1}{N} \sum_{i_1,\dots,i_n}^{N} \mathbb{E}[x_{i_1i_2}x_{i_2i_3}\cdots x_{i_ni_1}]$$
$$= \frac{1}{N} \sum_{i_1,\dots,i_n}^{N} \sum_{\pi\in P(n)} c_\pi[x_{i_1i_2}, x_{i_2i_3}, \dots, x_{i_ni_1}]$$

goes in leading order over to
$$\varphi(X^n) = \sum_{\pi\in NC(n)} \kappa_\pi;$$

and the κ_n are the free cumulants of the asymptotic eigenvalue distribution of X_N.

Question: What is the effect of applying a non-linear function to the entries of a random matrix?

Given input

- orthogonally invariant random matrix $X_N = (x_{ij})_{i,j=1}^N$
- function $f : \mathbb{R} \to \mathbb{R}$ (for now, polynomial)

Define the "non-linear" random matrix $Y_N = (y_{ij})$ by

$$y_{ij} = \begin{cases} \frac{1}{\sqrt{N}} f(\sqrt{N} x_{ij}), & i \neq j \\ 0, & i = j. \end{cases}$$

Question

How does this change the eigenvalue distribution, asymptotically, if $N \to \infty$?

Theorem

1. *The only cumulants which make asymptotically a contribution in the calculation of the moments of* $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ *are those with a cyclic index structure,*

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

   Those κ_n^f *are thus the free cumulants of the asymptotic eigenvalue distribution of the random matrix* Y_N*.*

2. *We have the following relation between the free cumulants* κ_n *of the asymptotic eigenvalue distribution of* X_N *and the free cumulants* κ_n^f *of the asymptotic eigenvalue distribution of the random matrix* Y_N*:*
   1. $\kappa_1^f = 0$;
   2. $\kappa_2^f = c_2(f(g), f(g)) = \mathbb{E}[f(g)^2] - \mathbb{E}[f(g)]^2$;
   3. *for* $n \geq 3$,

$$\kappa_n^f = \kappa_n \cdot \mathbb{E}[f'(g)]^n;$$

   where g *is a Gaussian random variable with mean zero and variance* κ_2*.*

1. The only cumulants which make asymptotically a contribution in the calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \ldots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

1. The only cumulants which make asymptotically a contribution in the calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Note that the cumulants of the y_{ij} are more complicated than the ones for the x_{ij};

1. The only cumulants which make asymptotically a contribution in the calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \ldots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Note that the cumulants of the y_{ij} are more complicated than the ones for the x_{ij}; in particular,

- non-cyclic cumulants can also be different from zero
- the subleading orders and subcycles are getting more complicated

1. The only cumulants which make asymptotically a contribution in the calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \ldots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Note that the cumulants of the y_{ij} are more complicated than the ones for the x_{ij}; in particular,

- non-cyclic cumulants can also be different from zero
- the subleading orders and subcycles are getting more complicated

But still, in leading order it's just the

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \ldots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct}$$

which contribute.

1. The only cumulants which make asymptotically a contribution in the calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1i_2}, y_{i_2i_3}, \dots, y_{i_ni_1}) \qquad \text{all } i_k \text{ distinct.}$$

Note that the cumulants of the y_{ij} are more complicated than the ones for the x_{ij}; in particular,

- non-cyclic cumulants can also be different from zero
- the subleading orders and subcycles are getting more complicated

But still, in leading order it's just the

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1i_2}, y_{i_2i_3}, \dots, y_{i_ni_1}) \qquad \text{all } i_k \text{ distinct}$$

which contribute.

Those κ_n^f are thus the free cumulants of the asymptotic eigenvalue distribution of the random matrix Y_N.

1. The only cumulants which make asymptotically a contribution in the calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Those κ_n^f are thus the free cumulants of the asymptotic eigenvalue distribution of the random matrix Y_N.

1. The only cumulants which make asymptotically a contribution in the calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Those κ_n^f are thus the free cumulants of the asymptotic eigenvalue distribution of the random matrix Y_N.

Consider $f(x) = x^3$. Then

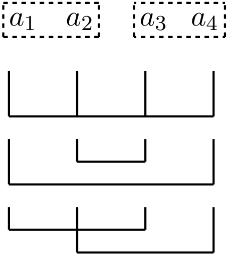
$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{23}x_{23}x_{23}, x_{34}x_{34}x_{34}, x_{41}x_{41}x_{41})$$

Theorem (Formula of Leonov and Shiryaev)

$$c_m(a_1 \cdots a_{i(1)}, \dots, a_{i(m-1)+1} \cdots a_{i(m)}) = \sum_{\substack{\pi \in P(n) \\ \pi \vee \tau = 1_n}} c_\pi[a_1, \dots, a_n]$$

where

$$\tau = \{(1, \dots, i(1)), \dots, (i(m-1)+1, \dots, i(m))\} \in \mathcal{P}(n).$$



assume $c_1(a_i) = 0$; then

$$c_2(a_1a_2, a_3a_4) = c_4(a_1, a_2, a_3, a_4) + c_2(a_1, a_4)c_2(a_2, a_3) + c_2(a_1, a_3)c_2(a_2, a_4)$$

1. Only cumulants which make asymptotically contribution in calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Those κ_n^f are thus the free cumulants of the asymptotic eigenvalue distribution of the random matrix Y_N.

Consider $f(x) = x^3$. Then

$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{23}x_{23}x_{23}, x_{34}x_{34}x_{34}, x_{41}x_{41}x_{41})$$

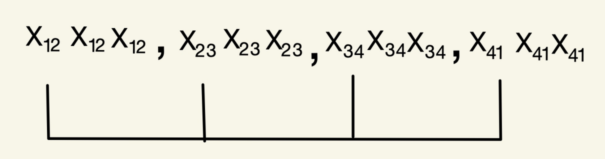
1. Only cumulants which make asymptotically contribution in calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Those κ_n^f are thus the free cumulants of the asymptotic eigenvalue distribution of the random matrix Y_N.

Consider $f(x) = x^3$. Then

$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{23}x_{23}x_{23}, x_{34}x_{34}x_{34}, x_{41}x_{41}x_{41})$$



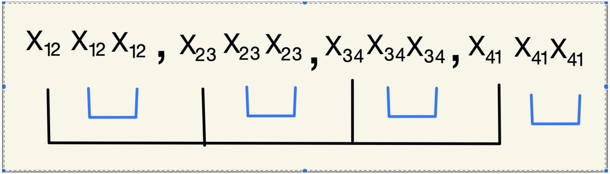
1. Only cumulants which make asymptotically contribution in calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Those κ_n^f are thus the free cumulants of the asymptotic eigenvalue distribution of the random matrix Y_N.

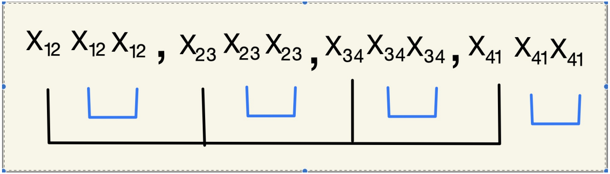
Consider $f(x) = x^3$. Then

$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{23}x_{23}x_{23}, x_{34}x_{34}x_{34}, x_{41}x_{41}x_{41})$$



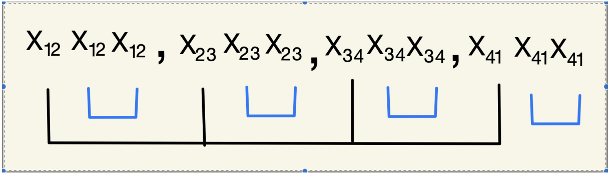
$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) \sim N^4 N^{-3} (N^{-1})^4 = N^{-3}$$

$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{23}x_{23}x_{23}, x_{34}x_{34}x_{34}, x_{41}x_{41}x_{41})$$



$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) \sim N^4 N^{-3}(N^{-1})^4 = N^{-3}$$

$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{23}x_{23}x_{23}, x_{34}x_{34}x_{34}, x_{41}x_{41}x_{41})$$

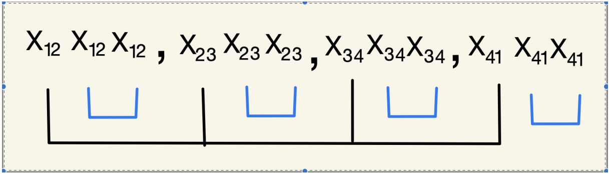


$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) \sim N^4 N^{-3} (N^{-1})^4 = N^{-3}$$

However, in the case with subcycles

$$c_4(y_{12}, y_{21}, y_{12}, y_{21}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{21}x_{21}x_{21}, x_{12}x_{12}x_{12}, x_{21}x_{21}x_{21})$$

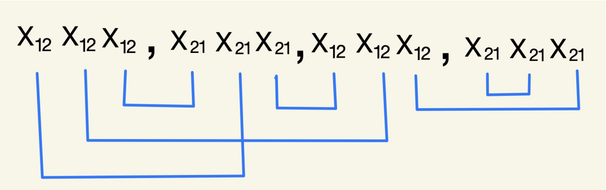
$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{23}x_{23}x_{23}, x_{34}x_{34}x_{34}, x_{41}x_{41}x_{41})$$



$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) \sim N^4 N^{-3} (N^{-1})^4 = N^{-3}$$

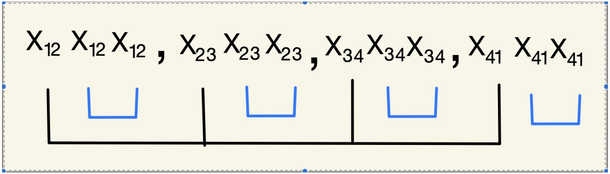
However, in the case with subcycles

$$c_4(y_{12}, y_{21}, y_{12}, y_{21}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{21}x_{21}x_{21}, x_{12}x_{12}x_{12}, x_{21}x_{21}x_{21})$$



$$c_4(y_{12}, y_{21}, y_{12}, y_{21}) \sim N^4 (N^{-1})^6 = N^{-2}$$

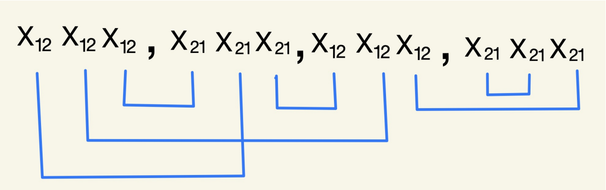
$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{23}x_{23}x_{23}, x_{34}x_{34}x_{34}, x_{41}x_{41}x_{41})$$



$$c_4(y_{12}, y_{23}, y_{34}, y_{41}) \sim N^4 N^{-3}(N^{-1})^4 = N^{-3} \qquad \longrightarrow \qquad \text{4 free indices}$$

However, in the case with subcycles

$$c_4(y_{12}, y_{21}, y_{12}, y_{21}) = N^4 c_4(x_{12}x_{12}x_{12}, x_{21}x_{21}x_{21}, x_{12}x_{12}x_{12}, x_{21}x_{21}x_{21})$$



$$c_4(y_{12}, y_{21}, y_{12}, y_{21}) \sim N^4 (N^{-1})^6 = N^{-2} \qquad \longrightarrow \qquad \text{only 2 free indices}$$

1. Only cumulants which make asymptotically contribution in calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Note: Non-cyclic cumulants can be non-zero for the y_{ij}

1. Only cumulants which make asymptotically contribution in calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Note: Non-cyclic cumulants can be non-zero for the y_{ij}
Consider $f(x) = x^2$.

$$c_2(y_{12}, y_{23}) = N c_2(x_{12}x_{12}, x_{23}x_{23})$$

1. Only cumulants which make asymptotically contribution in calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \ldots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Note: Non-cyclic cumulants can be non-zero for the y_{ij}
Consider $f(x) = x^2$.

$$\begin{aligned} c_2(y_{12}, y_{23}) &= N c_2(x_{12}x_{12}, x_{23}x_{23}) \\ &= N c_2(x_{21}x_{12}, x_{23}x_{32}) \end{aligned}$$

1. Only cumulants which make asymptotically contribution in calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Note: Non-cyclic cumulants can be non-zero for the y_{ij}
Consider $f(x) = x^2$.

$$\begin{aligned} c_2(y_{12}, y_{23}) &= N c_2(x_{12}x_{12}, x_{23}x_{23}) \\ &= N c_2(x_{21}x_{12}, x_{23}x_{32}) \\ &= N c_4(x_{21}, x_{12}, x_{23}, x_{32}) \end{aligned}$$

1. Only cumulants which make asymptotically contribution in calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Note: Non-cyclic cumulants can be non-zero for the y_{ij}
Consider $f(x) = x^2$.

$$\begin{aligned} c_2(y_{12}, y_{23}) &= N c_2(x_{12}x_{12}, x_{23}x_{23}) \\ &= N c_2(x_{21}x_{12}, x_{23}x_{32}) \\ &= N c_4(x_{21}, x_{12}, x_{23}, x_{32}) \\ &\sim N N^{-3} \\ &\sim N^{-2} \end{aligned}$$

1. Only cumulants which make asymptotically contribution in calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Note: Non-cyclic cumulants can be non-zero for the y_{ij}
Consider $f(x) = x^2$.

$$\begin{aligned} c_2(y_{12}, y_{23}) &= N c_2(x_{12}x_{12}, x_{23}x_{23}) \\ &= N c_2(x_{21}x_{12}, x_{23}x_{32}) \\ &= N c_4(x_{21}, x_{12}, x_{23}, x_{32}) \\ &\sim N N^{-3} \\ &\sim N^{-2} \end{aligned}$$

But this is still small enough to not contribute to $\mathbb{E}[\mathrm{tr}(Y_N^n)]$.

- The only cumulants which make asymptotically a contribution in the calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

Those κ_n^f are thus the free cumulants of the asymptotic eigenvalue distribution of the random matrix Y_N.

- The only cumulants which make asymptotically a contribution in the calculation of the moments of $\mathbb{E}[\mathrm{tr}(Y_N^n)]$ are those with a cyclic index structure,

$$\kappa_n^f := \lim_{N\to\infty} N^{n-1} c_n(y_{i_1 i_2}, y_{i_2 i_3}, \dots, y_{i_n i_1}) \qquad \text{all } i_k \text{ distinct.}$$

  Those κ_n^f are thus the free cumulants of the asymptotic eigenvalue distribution of the random matrix Y_N.
- The κ_n^f are given by
  - for $n = 1$
$$\kappa_1^f = 0$$
  - for $n = 2$
$$\kappa_2^f = c_2(f(g), f(g))$$
  - for $n \geq 3$,
$$\kappa_n^f = \kappa_n \cdot \mathbb{E}[f'(g)]^n;$$

  where g is a Gaussian random variable with mean zero and variance κ_2.

- for $n \geq 3$: $\kappa_n^f = \kappa_n \cdot \mathbb{E}[f'(g)]^n$

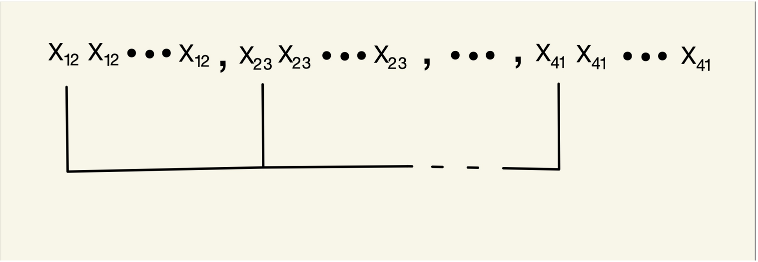
- for $n \geq 3$: $\kappa_n^f = \kappa_n \cdot \mathbb{E}[f'(g)]^n$

$$c_4(f_1(x_{12}), f_2(x_{23}), \dots, f_4(x_{41})) = c_4(x_{12}^{m_1}, x_{23}^{m_2}, \dots, x_{41}^{m_4})$$

$$X_{12}\, X_{12} \bullet\bullet\bullet X_{12}\, ,\; X_{23}\, X_{23} \bullet\bullet\bullet X_{23}\, ,\; \bullet\bullet\bullet\, ,\; X_{41}\, X_{41} \bullet\bullet\bullet X_{41}$$

- for $n \geq 3$: $\kappa_n^f = \kappa_n \cdot \mathbb{E}[f'(g)]^n$

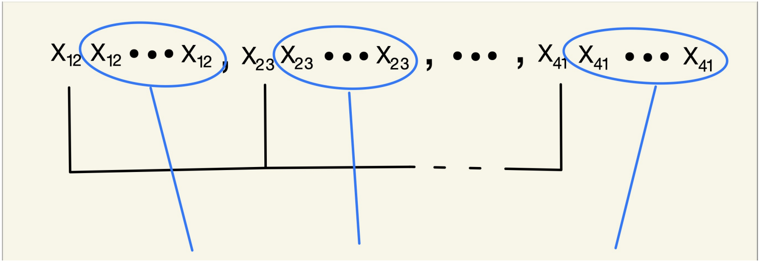
$$c_4(f_1(x_{12}), f_2(x_{23}), \ldots, f_4(x_{41})) = c_4(x_{12}^{m_1}, x_{23}^{m_2}, \ldots, x_{41}^{m_4})$$



κ_4 m_1 m_2 $\ldots$ m_4

- for $n \geq 3$: $\kappa_n^f = \kappa_n \cdot \mathbb{E}[f'(g)]^n$

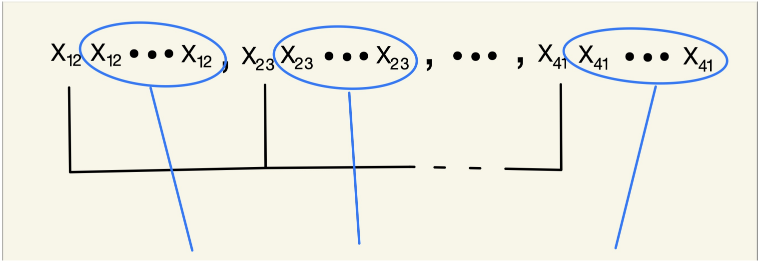
$$c_4(f_1(x_{12}), f_2(x_{23}), \dots, f_4(x_{41})) = c_4(x_{12}^{m_1}, x_{23}^{m_2}, \dots, x_{41}^{m_4})$$



κ_4 m_1 m_2 $\dots$ m_4

- for $n \geq 3$: $\kappa_n^f = \kappa_n \cdot \mathbb{E}[f'(g)]^n$

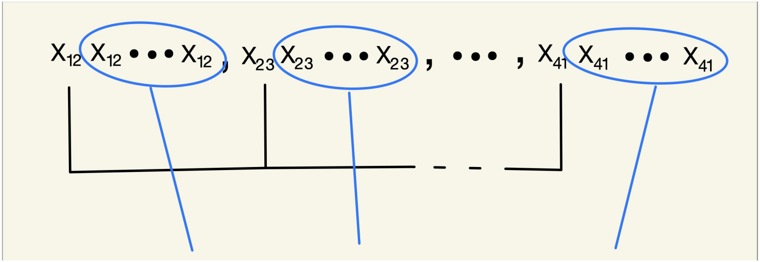
$$c_4(f_1(x_{12}), f_2(x_{23}), \dots, f_4(x_{41})) = c_4(x_{12}^{m_1}, x_{23}^{m_2}, \dots, x_{41}^{m_4})$$



$\kappa_4 \quad m_1 \quad \mathbb{E}[g^{m_1-1}] \quad m_2 \quad \mathbb{E}[g^{m_2-1}] \quad \dots \quad m_4 \quad \mathbb{E}[g^{m_4-1}]$

- for $n \geq 3$: $\kappa_n^f = \kappa_n \cdot \mathbb{E}[f'(g)]^n$

$$c_4(f_1(x_{12}), f_2(x_{23}), \dots, f_4(x_{41})) = c_4(x_{12}^{m_1}, x_{23}^{m_2}, \dots, x_{41}^{m_4})$$



$\kappa_4 \quad m_1 \quad \mathbb{E}[g^{m_1-1}] \quad m_2 \quad \mathbb{E}[g^{m_2-1}] \quad \dots \quad m_4 \quad \mathbb{E}[g^{m_4-1}]$

$\kappa_4 \qquad \mathbb{E}[f_1'(g)] \qquad \mathbb{E}[f_2'(g)] \qquad \cdots \qquad \mathbb{E}[f_4'(g)]$

where g is a Gaussian random variable with mean zero and variance κ_2.

- for $n = 2$: $\kappa_2^f = c_2(f(g), f(g))$

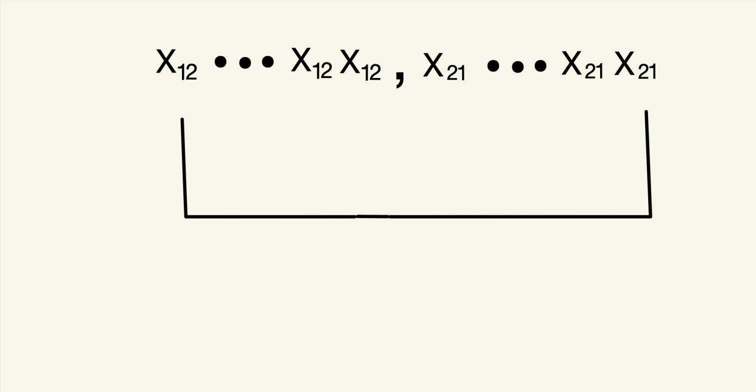
- for $n = 2$: $\kappa_2^f = c_2(f(g), f(g))$

$$c_2(f_1(x_{12}), f_2(x_{21})) = c_4(x_{12}^{m_1}, x_{21}^{m_2})$$

$$X_{12} \bullet\bullet\bullet X_{12} X_{12} \, , \, X_{21} \bullet\bullet\bullet X_{21} X_{21}$$

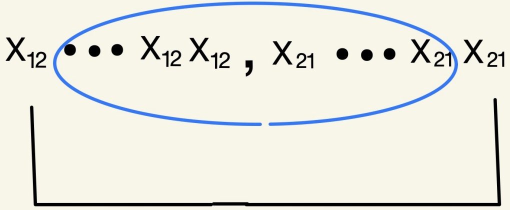
- for $n = 2$: $\kappa_2^f = c_2(f(g), f(g))$

$$c_2(f_1(x_{12}), f_2(x_{21})) = c_4(x_{12}^{m_1}, x_{21}^{m_2})$$



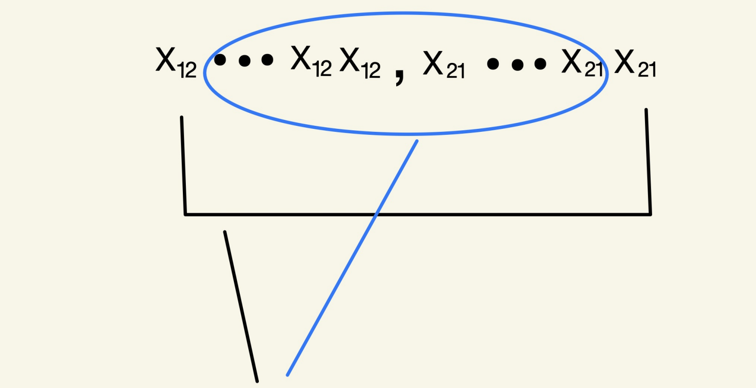
- for $n = 2$: $\kappa_2^f = c_2(f(g), f(g))$

$$c_2(f_1(x_{12}), f_2(x_{21})) = c_4(x_{12}^{m_1}, x_{21}^{m_2})$$



- for $n = 2$: $\kappa_2^f = c_2(f(g), f(g))$

$$c_2(f_1(x_{12}), f_2(x_{21})) = c_4(x_{12}^{m_1}, x_{21}^{m_2})$$



$c_2(f_1(g), f_2(g))$

where g is a Gaussian random variable with mean zero and variance κ_2.

Gaussian equivalence principle

Corollary

Let X_N and Y_N be as above. Then the non-linear random matrix model Y_N has the same asymptotic eigenvalue distribution as the linear model

$$\hat{Y}_N := \theta_1 X_N + \theta_2 Z_N,$$

where Z_N is a symmetric standard GOE random matrix which is independent from X_N and where

$$\theta_1 := \mathbb{E}[f'(g)]$$

and

$$\theta_2 : = \sqrt{\mathbb{E}[f(g)^2] - \mathbb{E}[f(g)]^2 - \kappa_2 \mathbb{E}[f'(g)]^2},$$

where g is a Gaussian random variable with mean zero and variance κ_2.

Example: The ReLU function

The function

$$f(x) = \mathsf{ReLU}(x) := \max(0, x).$$

Example: The ReLU function

The function

$$f(x) = \mathsf{ReLU}(x) := \max(0, x).$$

The orthogonal invariant random matrix

$$X_N = A_N^2 + A_N + B_N A_N + A_N B_N + B_N,$$

where A_N and B_N are independent standard GOE.

Example: The ReLU function

The function

$$f(x) = \mathsf{ReLU}(x) := \max(0, x).$$

The orthogonal invariant random matrix

$$X_N = A_N^2 + A_N + B_N A_N + A_N B_N + B_N,$$

where A_N and B_N are independent standard GOE.

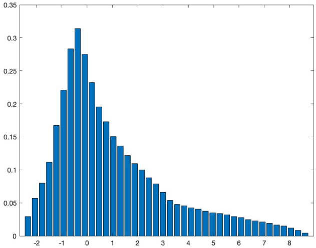
We want to compare

non-linear random matrix $\qquad Y_N = \mathsf{ReLU}(X_N)$

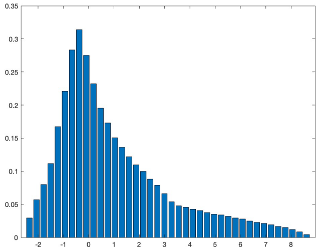
versus

linear Gaussian equivalent $\qquad \hat{Y}_N := \frac{1}{2} X_N + \frac{1}{2}\sqrt{5(1 - 2/\pi)} Z_N$

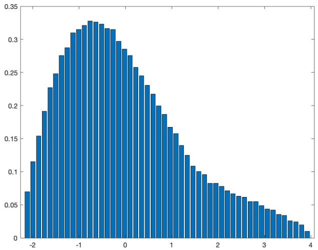
$$X_N = A_N^2 + A_N + B_N A_N + A_N B_N + B_N$$



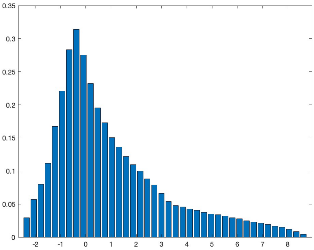
$$X_N = A_N^2 + A_N + B_N A_N + A_N B_N + B_N$$



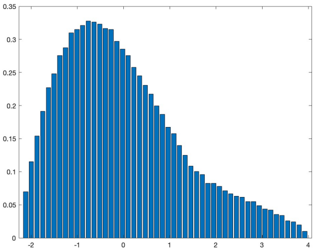
Y_N, after applying ReLU entrywise to X_N



$$X_N = A_N^2 + A_N + B_N A_N + A_N B_N + B_N$$

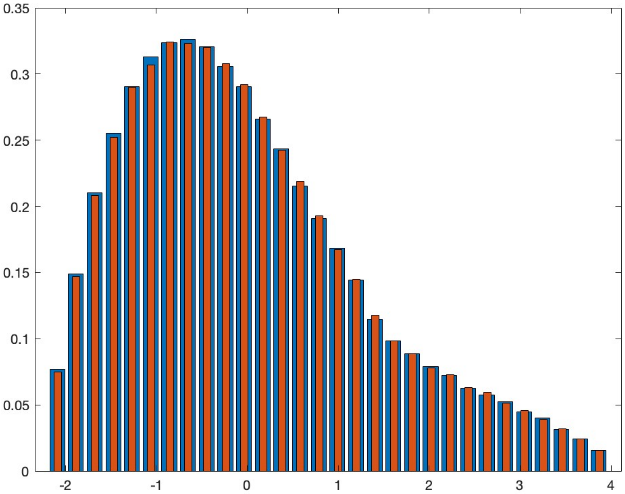


Y_N, after applying ReLU entrywise to X_N



Gaussian equivalent $\hat{Y}_N := \frac{1}{2} X_N + \frac{1}{2}\sqrt{5(1-2/\pi)}\, Z_N$

Superposition of the eigenvalues of the non-linear matrix Y_N and of its Gaussian equivalent $\hat{Y}_N$



Extension to multivariate case

Consider 2 independent matrices

$$X_N^{(1)} = (x_{ij}^{(1)})_{i,j=1}^N \qquad \text{and} \qquad X_N^{(2)} = (x_{ij}^{(2)})_{i,j=1}^N.$$

and a function (for now, a polynomial)

$$f : \mathbb{R}^2 \to \mathbb{R}$$

and define the non-linear random matrix $Y_N = (y_{ij})_{i,j=1}^N$ by

$$y_{ij} = \begin{cases} \frac{1}{\sqrt{N}} f\left(\sqrt{N} x_{ij}^{(1)}, \sqrt{N} x_{ij}^{(2)}\right), & i \neq j \\ 0, & i = j. \end{cases}$$

Corollary

Let $X_N^{(1)}$, $X_N^{(2)}$ *and* Y_N *be as above. Then the non-linear random matrix model* Y_N *has the same asymptotic eigenvalue distribution as the linear model*

$$\hat{Y}_N := \theta_1 X_N^{(1)} + \theta_2 X_N^{(2)} + \theta Z_N,$$

where Z_N *is a symmetric standard* GOE *random matrix which is independent from* $X_N^{(1)}$ *and* $X_N^{(2)}$ *and where*

$$\theta_1 := \mathbb{E}[\partial_1 f(g_1, g_2)], \qquad \theta_2 := \mathbb{E}[\partial_2 f(g_1, g_2)]$$

and

$$\theta := \sqrt{c_2(f(g_1, g_2), f(g_1, g_2)) - \kappa_2^{(1)} \theta_1^2 - \kappa_2^{(2)} \theta_2^2},$$

where g_1 *and* g_2 *are independent Gaussian random variables with mean zero and variance* $\kappa_2^{(1)}$ *and* $\kappa_2^{(2)}$*, respectively.*

Example: the maximum function

The function

$$f(x_1, x_2) = \max(x_1, x_2) = \frac{1}{2}(x_1 + x_2 + |x_1 - x_2|).$$

Example: the maximum function

The function

$$f(x_1, x_2) = \max(x_1, x_2) = \frac{1}{2}(x_1 + x_2 + |x_1 - x_2|).$$

The orthogonally invariant random matrices

$$X_N^{(1)} = \frac{-A_N^2 + B_N}{\sqrt{2}}, \qquad X_N^{(2)} = \frac{C_N^4 + C_N D_N + D_N C_N}{\sqrt{12}},$$

where A_N, B_N, C_N, D_N are independent standard GOE.

Example: the maximum function

The function

$$f(x_1, x_2) = \max(x_1, x_2) = \frac{1}{2}(x_1 + x_2 + |x_1 - x_2|).$$

The orthogonally invariant random matrices

$$X_N^{(1)} = \frac{-A_N^2 + B_N}{\sqrt{2}}, \qquad X_N^{(2)} = \frac{C_N^4 + C_N D_N + D_N C_N}{\sqrt{12}},$$

where A_N, B_N, C_N, D_N are independent standard GOE.

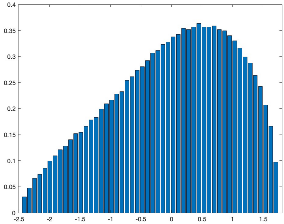
We want to compare

non-linear random matrix $\qquad Y_N = \max(X_1^{(1)}, X_2^{(2)})$

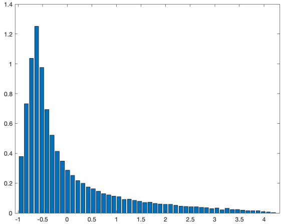
versus

Gaussian linear equivalent $\qquad \hat{Y}_N := \frac{1}{2} X_N^{(1)} + \frac{1}{2} X_N^{(2)} + \sqrt{\frac{1}{2} - \frac{1}{\pi}} Z_N$

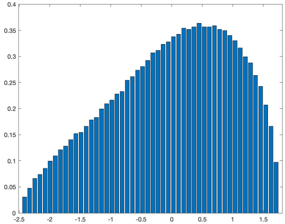
$$X_N^{(1)} = \frac{-A_N^2 + B_N}{\sqrt{2}}$$



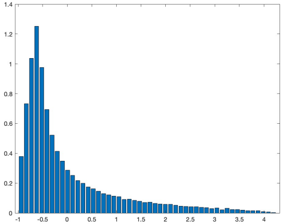
$$X_N^{(2)} = \frac{C_N^4 + C_N D_N + D_N C_N}{\sqrt{12}}$$



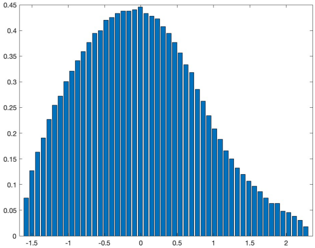
$$X_N^{(1)} = \frac{-A_N^2 + B_N}{\sqrt{2}}$$



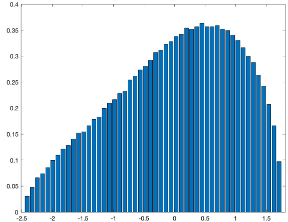
$$X_N^{(2)} = \frac{C_N^4 + C_N D_N + D_N C_N}{\sqrt{12}}$$



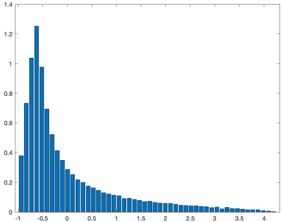
$$Y_N = \max(X_N^{(1)}, X_N^{(2)})$$



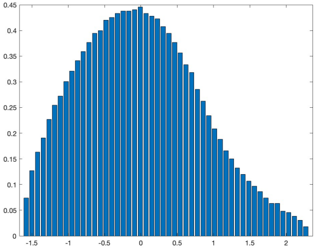
$$X_N^{(1)} = \frac{-A_N^2 + B_N}{\sqrt{2}}$$



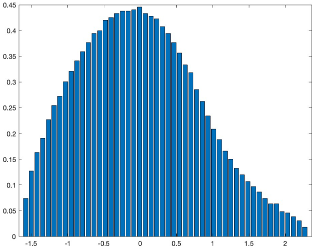
$$X_N^{(2)} = \frac{C_N^4 + C_N D_N + D_N C_N}{\sqrt{12}}$$



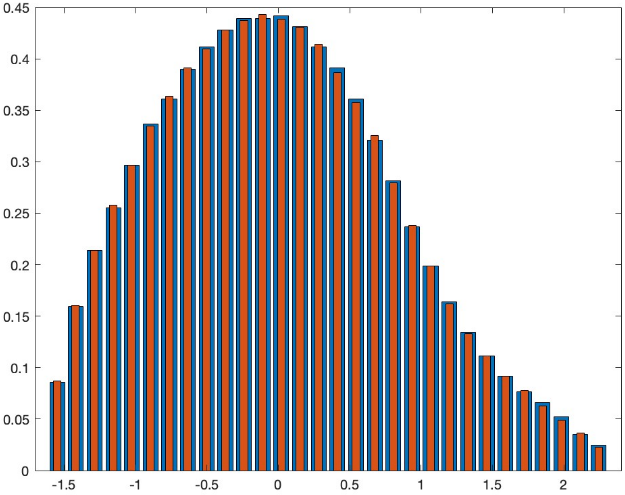
$$Y_N = \max(X_N^{(1)}, X_N^{(2)})$$



$$\hat{Y}_N := \frac{1}{2} X_N^{(1)} + \frac{1}{2} X_N^{(2)} + \sqrt{\frac{1}{2} - \frac{1}{\pi}} Z_N$$



Superposition of the eigenvalues of the non-linear matrix Y_N and of its Gaussian equivalent $\hat{Y}_N$



Example: the maximum function

The function

$$f(x_1, x_2) = \max(x_1, x_2) = \frac{1}{2}(x_1 + x_2 + |x_1 - x_2|).$$

The orthogonally invariant random matrices

$$X_N^{(1)} = \frac{-A_N^2 + B_N}{\sqrt{2}}, \qquad X_N^{(2)} = \frac{C_N^4 + C_N D_N + D_N C_N}{\sqrt{12}},$$

where A_N, B_N, C_N, D_N are independent standard GOE.

We want to compare

non-linear random matrix $\qquad Y_N = \max(X_1^{(1)}, X_2^{(2)})$

versus

Gaussian linear equivalent $\qquad \hat{Y}_N := \frac{1}{2} X_N^{(1)} + \frac{1}{2} X_N^{(2)} + \sqrt{\frac{1}{2} - \frac{1}{\pi}} Z_N$

Example: something more non-linear

The function

$$f(x_1, x_2) = \frac{1}{2} x_1^3 x_2^2 - 2 x_1^2 x_2$$

The orthogonally invariant random matrices

$$X_N^{(1)} = \frac{-A_N^2 + B_N}{\sqrt{2}}, \qquad X_N^{(2)} = \frac{C_N^4 + C_N D_N + D_N C_N}{\sqrt{12}},$$

where A_N, B_N, C_N, D_N are independent standard GOE.

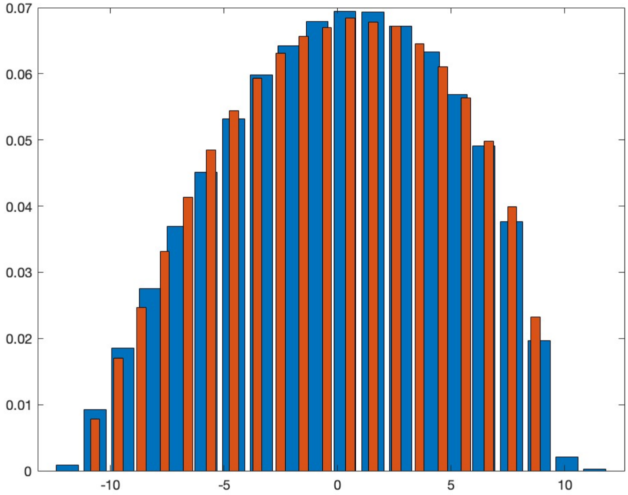
We want to compare

non-linear random matrix $\quad Y_N = \max(X_1^{(1)}, X_2^{(2)})$

versus

Gaussian linear equivalent $\quad \hat{Y}_N := \frac{3}{2} X_N^{(1)} - 2 X_N^{(2)} + \sqrt{\frac{93}{4} - \frac{9}{4} - 4} Z_N$

Superposition of the eigenvalues of the non-linear matrix Y_N and of its Gaussian equivalent $\hat{Y}_N$



Extension to multivariate case with correlation

Let $X_N^{(1)}$ and $X_N^{(2)}$ be jointly orthogonally invariant random matrices.

$$\kappa_n^{(r_1,\dots,r_n)} := \lim_{N\to\infty} N^{n-1} c_n(x_{i_1 i_2}^{(r_1)}, x_{i_2 i_3}^{(r_2)}, \dots, x_{i_n i_1}^{(r_n)}).$$

and define the non-linear random matrix $Y_N = f(X_N^{(1)}, X_N^{(2)}) = (y_{ij})_{i,j=1}^N$ by

$$y_{ij} = \begin{cases} \frac{1}{\sqrt{N}} f\left(\sqrt{N} x_{ij}^{(1)}, \sqrt{N} x_{ij}^{(2)}\right), & i \neq j \\ 0, & i = j. \end{cases}$$

Corollary

The non-linear random matrix model $Y_N = f(X_N^{(1)}, X^{(2)})$ *has the same asymptotic eigenvalue distribution as the linear model*

$$\hat{Y}_N := \theta_1 X_N^{(1)} + \theta_2 X_N^{(2)} + \theta Z_N,$$

where Z_N *is a symmetric standard* GOE *random matrix which is independent from* $X_N^{(1)}$ *and* $X_N^{(2)}$ *and where*

$$\theta_1 := \mathbb{E}[\partial_1 f(g_1, g_2)], \qquad \theta_2 := \mathbb{E}[\partial_2 f(g_1, g_2)]$$

$$\theta := \sqrt{c_2(f(g_1, g_2), f(g_1, g_2)) - \kappa_2^{(1,1)}\theta_1^2 - \kappa_2^{(2,2)}\theta_2^2 - 2\kappa_2^{(1,2)}\theta_1\theta_2},$$

where g_1 *and* g_2 *are a Gaussian family of random variables with mean zero and covariance*

$$c_2(g_1, g_1) = \kappa_2^{(1,1)}, \qquad c_2(g_2, g_2) = \kappa_2^{(2,2)}, \qquad c_2(g_1, g_2) = \kappa_2^{(1,2)}.$$

Example: the maximum function

The function

$$f(x_1, x_2) = \max(x_1, x_2) = \frac{1}{2}(x_1 + x_2 + |x_1 - x_2|).$$

Example: the maximum function

The function

$$f(x_1, x_2) = \max(x_1, x_2) = \frac{1}{2}(x_1 + x_2 + |x_1 - x_2|).$$

The orthogonally invariant random matrices

$$X_N^{(1)} = \frac{A_N^2 - \sqrt{2}B_N}{\sqrt{3}}, \qquad X_N^{(2)} = \frac{A_N^4 + C_N A_N + A_N C_N}{\sqrt{12}},$$

where A_N, B_N, C_N are independent standard GOE.

Example: the maximum function

The function

$$f(x_1, x_2) = \max(x_1, x_2) = \frac{1}{2}(x_1 + x_2 + |x_1 - x_2|).$$

The orthogonally invariant random matrices

$$X_N^{(1)} = \frac{A_N^2 - \sqrt{2}B_N}{\sqrt{3}}, \qquad X_N^{(2)} = \frac{A_N^4 + C_N A_N + A_N C_N}{\sqrt{12}},$$

where A_N, B_N, C_N are independent standard GOE.

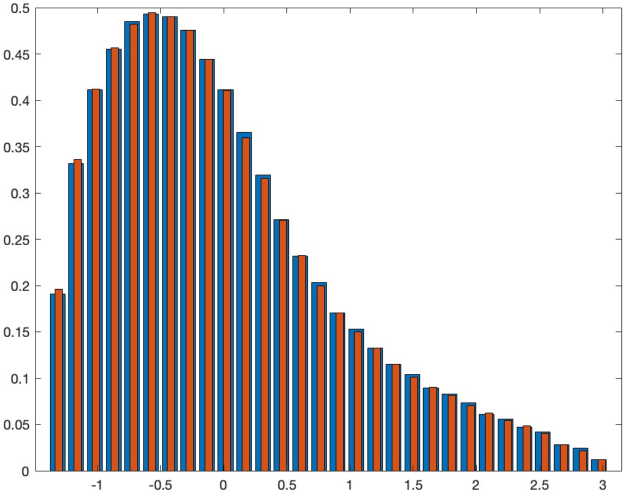
We want to compare

non-linear random matrix $\qquad Y_N = \max(X_1^{(1)}, X_2^{(2)})$

versus

Gaussian linear equivalent $\qquad \hat{Y}_N := \frac{1}{2}X_N^{(1)} + \frac{1}{2}X_N^{(2)} + \sqrt{\frac{1}{4} - \frac{1}{2\pi}}Z_N$

Superposition of the eigenvalues of the non-linear matrix Y_N and of its Gaussian equivalent $\hat{Y}_N$



What is still to do?

- extend proofs from polynomials to more general functions

What is still to do?

- extend proofs from polynomials to more general functions
- estimates for rate of convergence

What is still to do?

- extend proofs from polynomials to more general functions
- estimates for rate of convergence
- address question whether the Y_N are asymptotically free from deterministic matrices (as the X_N are)

What is still to do?

- extend proofs from polynomials to more general functions
- estimates for rate of convergence
- address question whether the Y_N are asymptotically free from deterministic matrices (as the X_N are)
- think about generalizations to random tensors

What is still to do?

- extend proofs from polynomials to more general functions
- estimates for rate of convergence
- address question whether the Y_N are asymptotically free from deterministic matrices (as the X_N are)
- think about generalizations to random tensors
- clarify relation with work in statistical physics around quantum symmetric simple exclusion process or the Eigenstate Thermalization Hypothesis (Bernard and Huzra; Fava, Kuchan, Pappalardi)
  $\longrightarrow$ generalization to operator-valued free probability

What is still to do?

- extend proofs from polynomials to more general functions
- estimates for rate of convergence
- address question whether the Y_N are asymptotically free from deterministic matrices (as the X_N are)
- think about generalizations to random tensors
- clarify relation with work in statistical physics around quantum symmetric simple exclusion process or the Eigenstate Thermalization Hypothesis (Bernard and Huzra; Fava, Kuchan, Pappalardi)
  $\longrightarrow$ generalization to operator-valued free probability

Thank you for your attention!

What is still to do?

- extend proofs from polynomials to more general functions
- estimates for rate of convergence
- address question whether the Y_N are asymptotically free from deterministic matrices (as the X_N are)
- think about generalizations to random tensors
- clarify relation with work in statistical physics around quantum symmetric simple exclusion process or the Eigenstate Thermalization Hypothesis (Bernard and Huzra; Fava, Kuchan, Pappalardi)
  $\longrightarrow$ generalization to operator-valued free probability

Thank you for your attention!