Universality and delocalization of band matrix

UCLA
Jun Yin

Joint work with
H.T. Yau and F. Yang

Recall Wigner Random Matrix:

$$H = H^{\dagger}, \quad H_{ij} \quad i.i.d. \quad \mathbb{E}(H_{ij}) = 0, \quad \mathbb{E}(|H_{ij}|^2) = 1/N$$

$$1 \leq i, j \leq N$$

In the last 10 years, we have seen many variations.

* Not identical distribution
* Not mean zero
* Not uniform variance
* Not quite independent
* With spikes (like E. R graph)

In a Wigner random band matrix, we have the properties 1, 3, 5, but the most important is that it is a **Non-mean-field** model.

$$H_{xy} = 0, \quad a.s. \quad if \quad |x - y| \gg W$$

In a system with size length **L** , there is no interaction between x and y, if the distance between them is much greater than **W** (interaction scale).

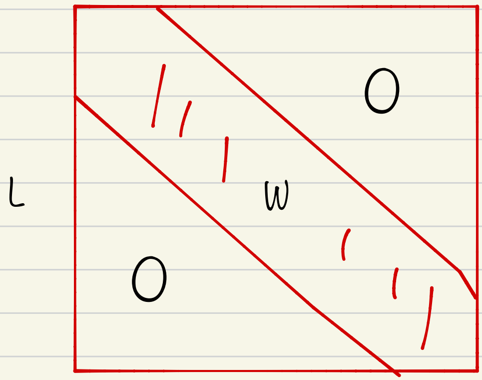
W is called band width

A more precise definition:

$$H = H^\dagger, \quad H_{ij} \quad i.d. \quad \mathbb{E}(H_{xy}) = 0, \quad \mathbb{E}(|H_{xy}|^2) = S_{xy}$$

$$1 \leq x, y \leq (\mathbb{Z}_N)^d$$

$$\sum_x S_{xy} = 1, \quad S_{xy} = O(1/W), \quad S_{xy} = 0 \;\; if \;\; \|x-y\| >> W$$

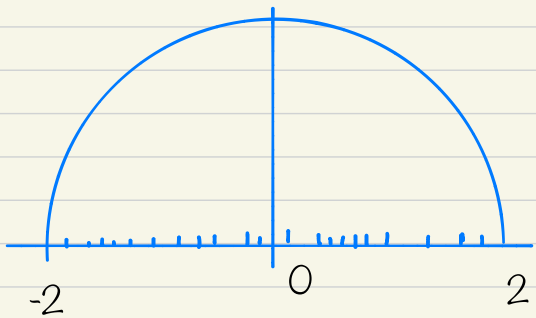


The interesting case is

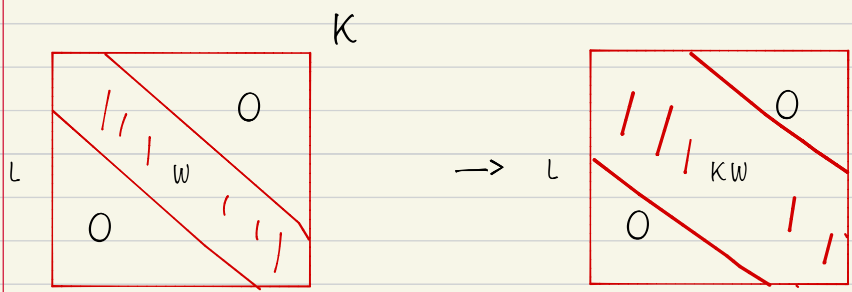
$$W = L^a, \quad 0 \leq a \leq 1$$

Global statistics: 2009 Erdos, Yau and Y:

Semi-Circle law holds up to scale W^{-d}.



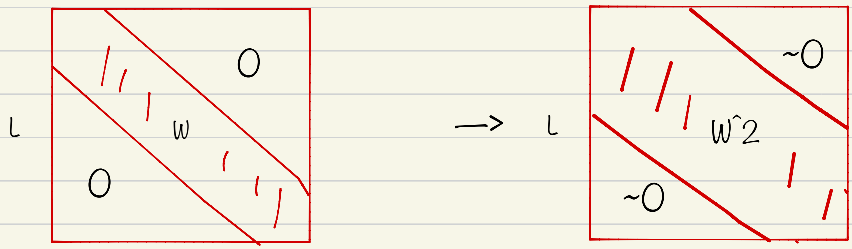
Interest property :



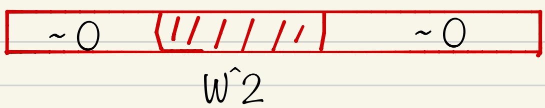
But it could be still a band matrix, even if the power K is very very large.

d=1

K->infinity



The reason is that in this case most eigenvectors of H are localized in the scale of W^2:



Recall:

$$H^k = \sum_\alpha \lambda_\alpha^k v_\alpha^\dagger v_\alpha$$

On the other hand, physicists want to know why sometimes it is **delocalized**

30 years ago, numerical calculations (P. R.L) and non-rigorous super symmetry (P.R.L) arguments showed that there is a phase transition between localization and delocalization.

Random Band Matrix Conjecture (bulk)

$$d = 1 \qquad W \sim L^{1/2}$$

$$d = 2 \qquad W \sim (\log L)^{1/2}$$

$$d \geq 3 \qquad W \gg 1 \quad (W \geq 3)$$

It is also the threshold for local statistics:

Small W	Poisson distribution
Large W	RM distribution (Sine-kernel)

The main new work for this talk is: 2020, Yau, Yang and Y

Delocalization of eigenvectors holds if

$$d \geq 10, \quad W = L^a, \quad a > 0$$

A quick review on previous results

Universality part:

$W \sim L$ Bourgade, Erdos, Yau and Y 2015

$W \gg L^{3/4}$ Bourgade, Yau, Yang and Y 2017-18

Localization part: for d=1 Gaussian band matrices

$W \ll L^{1/8}$ J. Schenker

$W \ll L^{1/7}$ R. Peled, J. Schenker, M. Shamis and S. Sodin

For delocalization part: $d=1$

$W \gg L^{6/7}$	Erdos and Knowles
$W \gg L^{4/5}$	Erdos, Knowles, Yau and Y.
$W \gg L^{7/9}$	Y. He and M. Marcozzi
$W \gg L^{3/4}$	Bourgade, Yau, Yang and Y

For Edge eigenvector:

$W \sim L^{5/6}$, S. Sodin (moment method)

For supersymmetry:

When the entries are Gaussian with some specific covariance profile, one can apply supersymmetry techniques

$W \sim L$	T. Shcherbina 2014
$W \sim L^{(6/7)}$	Bao and Erdos 2017
$W \sim L^{(1/2)}$	T. Shcherbina and M. Shcherbina 2017, 2019
$W \sim L^{(1/2)}$	M. Disertoril, M. Lohmann, S. Sodin 2018

For delocalization, $d>1$, $W = L^a$

$$a > \frac{6}{d+6}$$ L. Erdos and A. Knowles

$$a > \frac{d+2}{2d+2}$$ L. Erdos, A. Knowles, Yau and Y

$$a > \frac{d+1}{2d+1}$$ Y. He and M. Marcozzi

$$a > \frac{2}{2d+2}$$ Yang and Y. 2019-2020

Main new result of this talk

$a > 0$, $d >= 10$ Yau, Yang and Y 2020

Hope to finish the proof for $d >= 6$ in the next year.

Resolvent Tool

$$G = (H - z)^{-1}, \quad z \in \mathbb{C}$$

Delocalization of eigenvectors is implied by the following sense of delocalization of resolvent:

$$\exists\, \eta > 0, \quad s.t. \quad \forall\, E \in (-2, 2), \quad \forall\, \ell \ll L$$

$$\max_x \frac{\sum_{|y-x| \le \ell} |G_{xy}(z)|^2}{\sum |G_{xy}(z)|^2} \ll 1, \qquad z = E_0 + i\eta$$

Usually people choose $\eta \sim W^2/L^2$

When eta is larger, the above inequality does not hold.

But actually resolvent polynomially decays

(slower than power of d)

$$|G_{xy}|^2 \sim L/W^2, \qquad d = 1$$

$$|G_{xy}|^2 \sim 1/W^2, \qquad d = 2$$

$$|G_{xy}|^2 \sim \frac{1}{W^d + W^2|x-y|^{d-2}}, \qquad d \geq 3$$

$$\max \sim W^{-d}, \qquad \min \sim \frac{1}{W^2 L^{d-2}}$$

The main contribution comes from |x-y| ~ L (min part)

Therefore, really we need to prove this profile for delocalization.

Previous tool / work

$$L_{\max}: \quad \max_{x\neq y} |G_{xy}| \sim W^{-d/2}$$ (highly non-trivial)

$$L_2: \quad \sum_y |G_{xy}|^2 \sim 1/\eta$$

In this work, we need to

$$G_{xy} \sim (W^{-d/2} - term) + (W^{-d} - term) + (W^{-3d/2} - term)$$

$$+ (\ldots) + (W^{-k\,d/2} - term) + +(\ldots)$$

$$+ (W^{-2}|x-y|^{-d+2} - term)$$

Actually, we do (term by term) prove these terms are zeros, except the last one.

Universal - Expansion

We expand the resolvent with
some resolvent identities
some conditional expectation tricks
and a well designed rule on expansion order,
(which will be the main component of the first paper in this series)

Why Universal:

There is a large universal expansion, for W = L^a , for fixed a>0,
we will only use a subpart of the whole expansion.

$$\mathcal{P}_a \subset \mathcal{P}_{a'}, \quad a \geq a$$

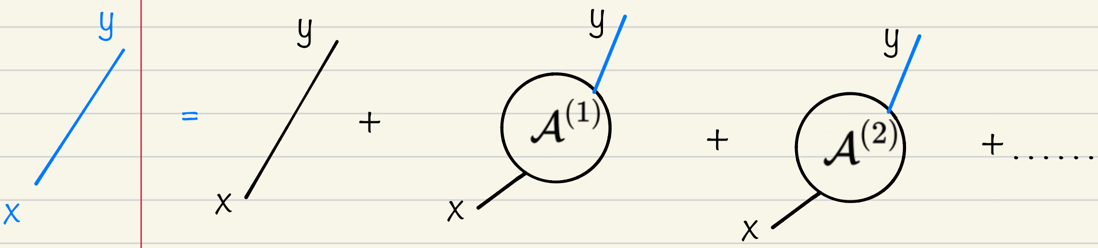
The limit expansion has a very clear fractal structure

$$\lim_{a \to 0} \mathcal{P}_a = \mathcal{P}_0$$

Different W cases share the same expansion is a very important property
for our proof. (You will see soon)

$$T_{xy} := \sum_{y':\ |y'-y|\sim W} |G_{xy'}|^2$$

$$T = B + B\cdot\mathcal{A}\cdot T, \quad B = S(I-S)^{-1}, \quad S_{xy} = \mathbb{E}|H_{xy}|^2$$



(B has the profile as above)

$$\|B\|_{L_1\to L_\infty} \sim \eta^{-1}$$

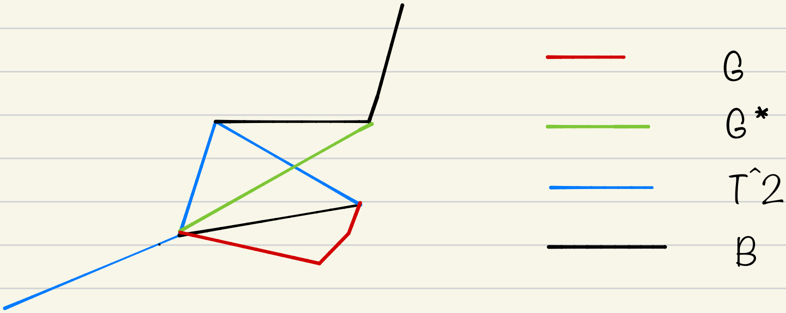
Here $\mathcal{A}^{(k)}$ is the sub graphs with $\sum_{z'} \mathcal{A}^{(k)}_{zz'} \sim (W^{-d/2})^k$

With some non-trivial cancellation (i.e., Z lemma idea), one can show A^1 term is zero.

Therefore, it would be an effective expansion if

$$\eta^{-1}W^{-d} \ll 1, \quad i.e., \quad L \ll W^{\frac{d+2}{2}}$$

Complexity of the expansion



Even for such simple case:

Each free index: K^d

Each G edge. $K^{(-d/2+1)}$ if $|x-y| \sim K$

Each B, T edge. $K^{(-d+2)}$

Expansion seems not a good idea, since one extra edge can not cancel the factor contributed by one free index.

The expansion does not produce enough edges .

The more expansion (with no correct plan),
the messier (out of control) the graphs will be.

The number of graphs

k-th order: $A^{(k)}$ has about $(2k)^{(3k)}$ graphs.

k=2, with some simplification,

there are about hundreds of graphs

k=3, with help of computer (Matlab),

it took iMac 1 hr on computation and me 1 month in coding

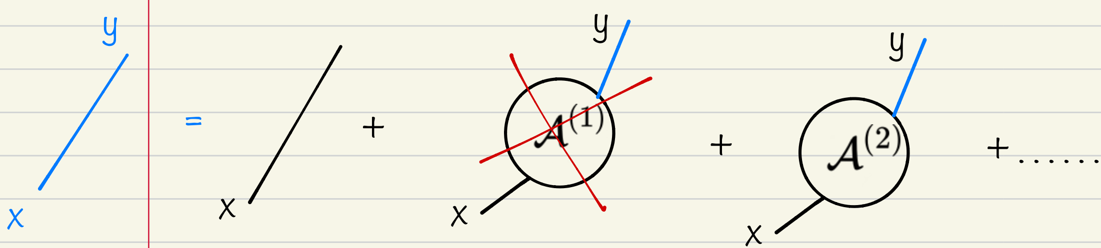
k=4, computer basically never stop

Due to shortage of time,

I will only focus on one of main novelty in the proof.

The whole proof is composed of 3 papers (or more).

SUM ZERO



Sum up x and y

$L^d\eta^{-1}$ $\qquad$ $L^d\eta^{-1}$ $\qquad$ $\eta^{-2}\sum_{zz'}\mathcal{A}^{(2)}_{zz'}$

It turns out if we choose

$$L^2/W^2 \cdot W^{-3d/2} \ll 1, \qquad L^2/W^2 \cdot W^{-d} \gg 1$$

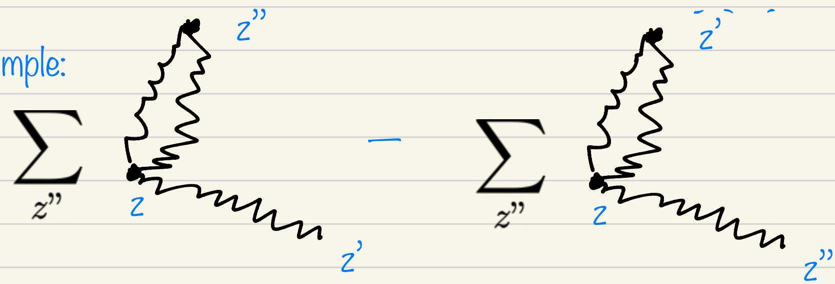
Then

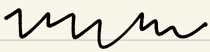
$$L^2/W^2 \sum_{z'} \underset{z}{\mathcal{A}^{(2)}}{}^{z'} \quad \ll 1$$

$$\Longrightarrow L^2/W^2 \sum_{z'} \mathcal{A}^{(2)}{}_{z}^{\;z'} = 0$$

SUM ZERO

Example:





S - edge

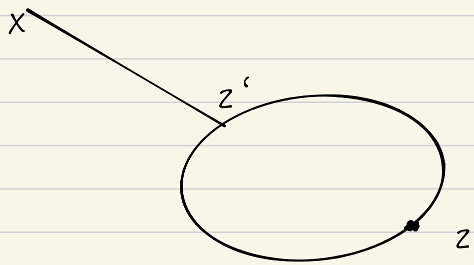
Why is Universal so important:

For certain L, we obtain the sum zero property of some A^(k)

From sum zero property, we know the cancellation between the graphs in A^(k)

Since for other (larger) L's, we use same graphs, so in those cases, A^(k) also has sum zero property for larger L case.

Why is it so important ?



$$\sum_{z'} B_{xz'} \mathcal{A}_{z'z} = \sum_{z'} \big(B_{xz'} - B_{xz} - \partial_z B_{xz}(z' - z)\big) \mathcal{A}_{z'z}$$

Thanks