Brown University Analysis Seminar

Eigenvalue Statistics for Ensembles of Random Matrices (especially Toeplitz and Palindromic Toeplitz)

Steven J. Miller

Brown University

Providence, RI, September 15^{th}, 2004

http://www.math.brown.edu/∼sjmiller/math/talks/talks.html

Collaborators

Random Matrices

- Peter Sarnak
- Rebecca Lehman
- Yi-Kai Liu
- Chris Hammond
- John Ramey
- John Sinsheimer
- Jason Teich
- Inna Zakharevich

Random Graphs

- Peter Sarnak
- Yakov Sinai
- Lior Silberman
- Dustin Steinhauer
- Randy Qian
- Felice Kuan
- Leo Goldmakher

Fundamental Problem

General Formulation: Studying some system, observe events at

$$t_1 \leq t_2 \leq t_3 \leq \cdots$$

Question: what rules govern the distribution of events?

Often normalize by average spacing.

Examples

- Spacings Between Energy Levels of **Heavy** Nuclei.

- Spacings Between Eigenvalues of **Large** Matrices.

- Spacings Between **High** Zeros of L-Functions.

Origins of Random Matrix Theory

Classical Mechanics: 3 Body Problem Intractable.

Heavy nuclei like Uranium ($200+$ protons / neutrons) even worse!

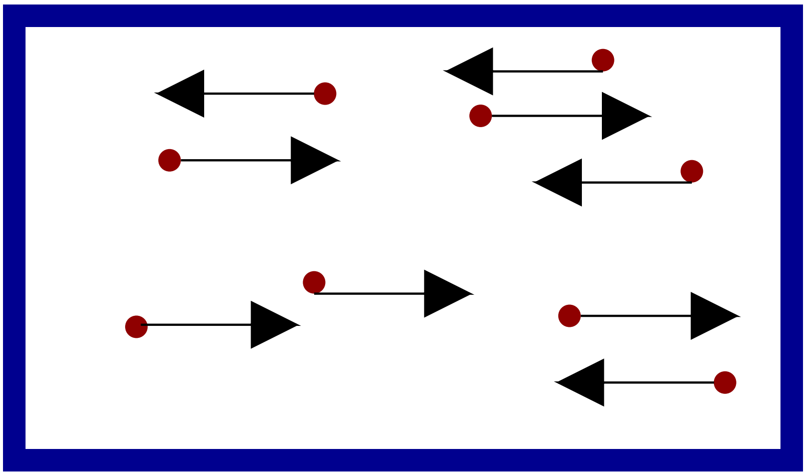
Fundamental Equation:

$$H\psi_n \; = \; E_n\psi_n$$

E_n are the energy levels

Approximate with finite matrix.

Statistical Mechanics



For each configuration, calculate quantity (say pressure).

Average over all configurations.

Random Matrix Ensembles

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1N} \\ a_{12} & a_{22} & a_{23} & \cdots & a_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1N} & a_{2N} & a_{3N} & \cdots & a_{NN} \end{pmatrix} = A^T$$

Let $p(x)$ be a probability density, often assume moments finite.

$$k^{th}\text{-moment} = \int_{\mathbb{R}} x^k p(x)dx.$$

Define

$$\text{Prob}(A)dA = \prod_{1 \leq i < j \leq N} p(a_{ij})da_{ij}.$$

Eigenvalue Questions

Density of Eigenvalues: How many eigenvalues lie in an interval $[a, b]$?

Spacings between Eigenvalues: How are the spacings between adjacent eigenvalues distributed?

Note: study *normalized* eigenvalues.

Eigenvalue Distribution

Key to Averaging:

$$\text{Trace}(A^k) \ = \ \sum_{i=1}^{N} \lambda_i^k(A).$$

By the Central Limit Theorem:

$$\begin{aligned}
\text{Trace}(A^2) \ &= \ \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} a_{ji} \\
&= \ \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}^2 \\
&\sim \ N^2 \cdot 1 \\
\sum_{i=1}^{N} \lambda_i^2(A) \ &\sim \ N^2
\end{aligned}$$

Gives $N \text{Ave}(\lambda_i^2(A)) \ \sim \ N^2$ or $\lambda_i(A) \sim \sqrt{N}$.

Eigenvalue Distribution (cont)

$\delta(x - x_0)$ is a unit point mass at x_0.

To each A, attach a probability measure:

$$\mu_{A,N}(x) = \frac{1}{N} \sum_{i=1}^{N} \delta\left(x - \frac{\lambda_i(A)}{2\sqrt{N}}\right)$$

Obtain:

$$\begin{aligned} k^{th}\text{-moment} &= \int x^k \mu_{A,N}(x) dx \\ &= \frac{1}{N} \sum_{i=1}^{N} \frac{\lambda_i^k(A)}{(2\sqrt{N})^k} \\ &= \frac{\text{Trace}(A^k)}{2^k N^{\frac{k}{2}+1}} \end{aligned}$$

Semi-Circle Law

$N \times N$ real symmetric matrices, entries i.i.d.r.v. from a fixed $p(x)$.

Semi-Circle Law: p mean 0, variance 1, other moments finite. With probability 1,

$$\mu_{A,N}(x) \longrightarrow \frac{2}{\pi}\sqrt{1-x^2} \quad \text{weakly}$$

Expected value of k^{th}-moment of $\mu_{A,N}(x)$ is

$$\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \frac{\text{Trace}(A^k)}{2^k N^{\frac{k}{2}+1}} \prod_{i<j} p(a_{ij}) da_{ij}$$

Method of Proof

With probability 1,

$$\mu_{A,N}(x) \longrightarrow \mu_{\text{Semi-Circle}}(x) \quad \text{weakly}.$$

Will do this by showing

$$\mathbb{E}[M_k(\mu_{A,N})] \rightarrow M_k(\text{SC})$$

and

$$\mathbb{E}[\left|M_k(\mu_{A,N}) - \mathbb{E}[M_k(\mu_{A,N})]\right|^2] \rightarrow 0.$$

Proof: 2^{nd}-Moment

$$\text{Trace}(A^2) \ = \ \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}a_{ji} \ = \ \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}^2.$$

Substituting:

$$\frac{1}{2^2N^2}\int_{\mathbb{R}}\cdots\int_{\mathbb{R}}\sum_{i,j=1}^{N} a_{ji}^2 \cdot p(a_{11})da_{11}\cdots p(a_{NN})da_{NN}$$

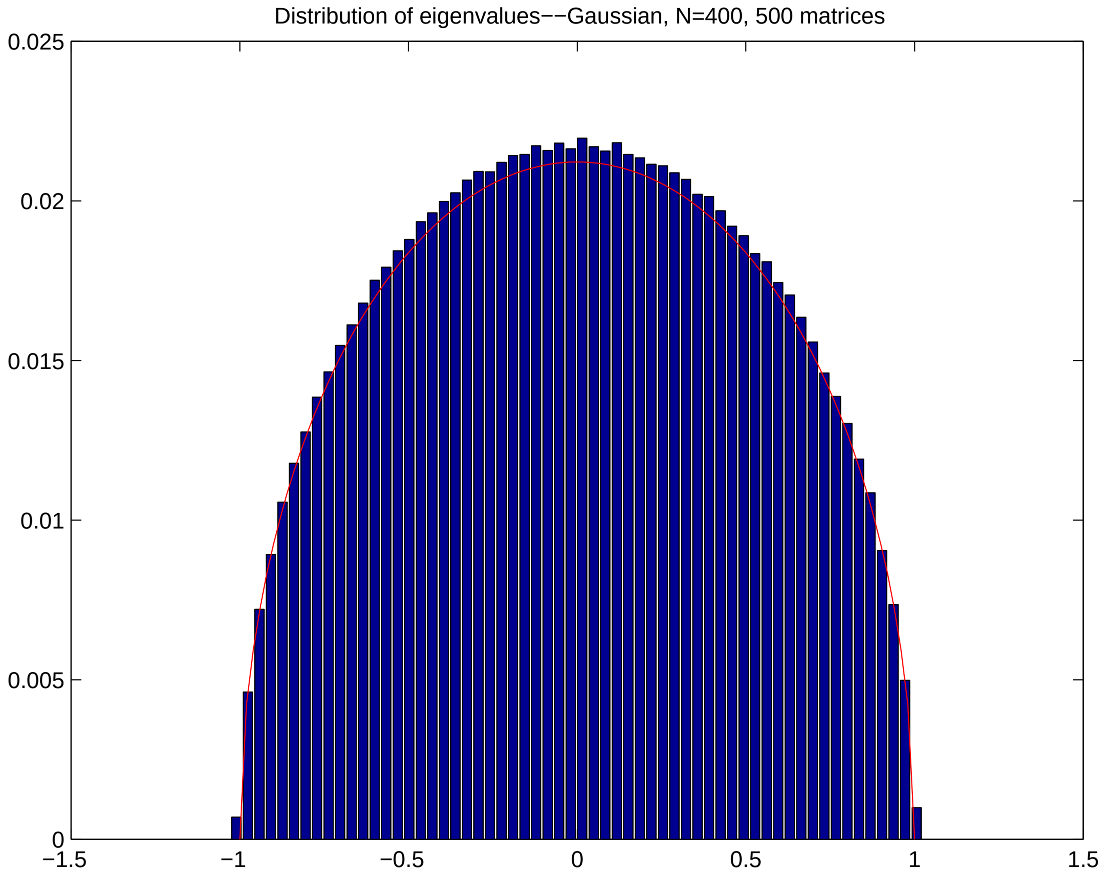
Integration factors as

$$\int_{a_{ij}\in\mathbb{R}} a_{ij}^2 p(a_{ij})da_{ij} \quad \cdot \quad \prod_{\substack{(k,l)\neq(ij)\\ k<l}} \int_{a_{kl}\in\mathbb{R}} p(a_{kl})da_{kl} \ = \ 1.$$

Have N^2 summands, answer is $\frac{1}{4}$.

Key: Trace and Averaging Formulas.

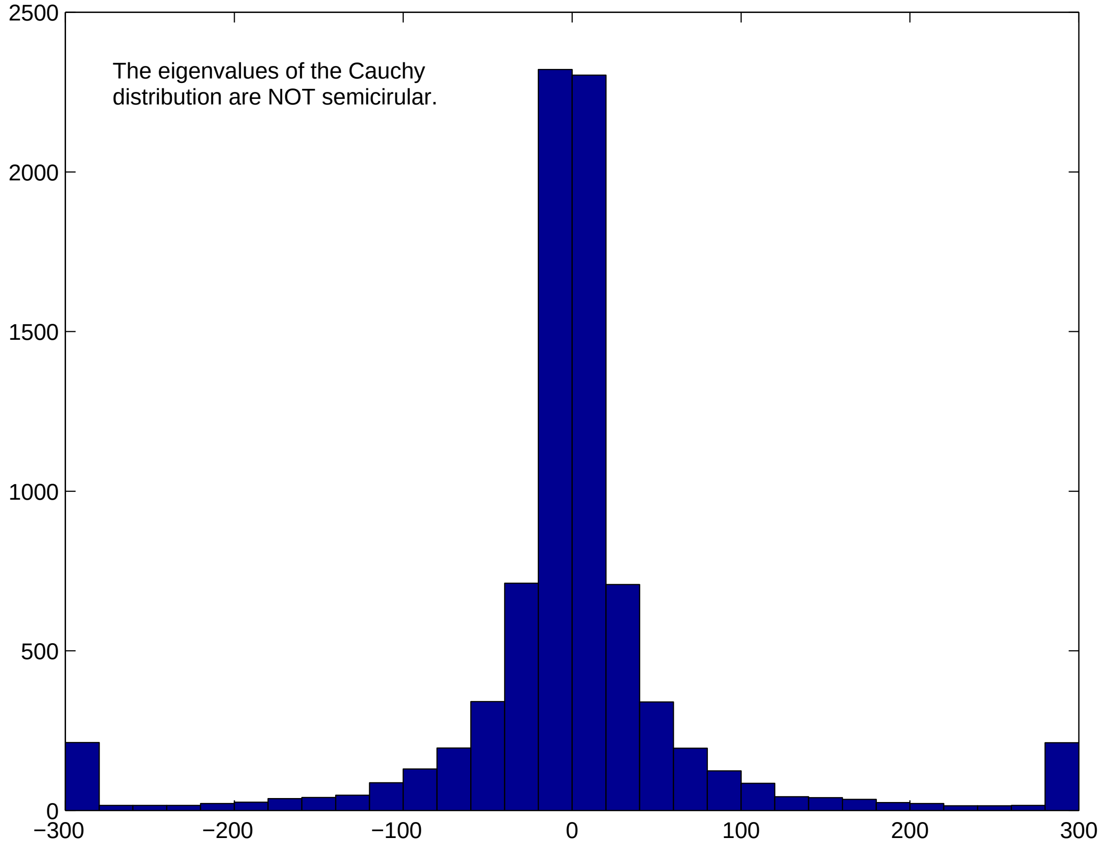
Random Matrix Theory: Semi-Circle Law



500 **Matrices: Gaussian** 400×400

$$p(x) \;=\; \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$$

Random Matrix Theory: Semi-Circle Law



Cauchy Distr: Not-Semicircular (Infinite Variance)

$$p(x) = \frac{1}{\pi(1+x^2)}$$

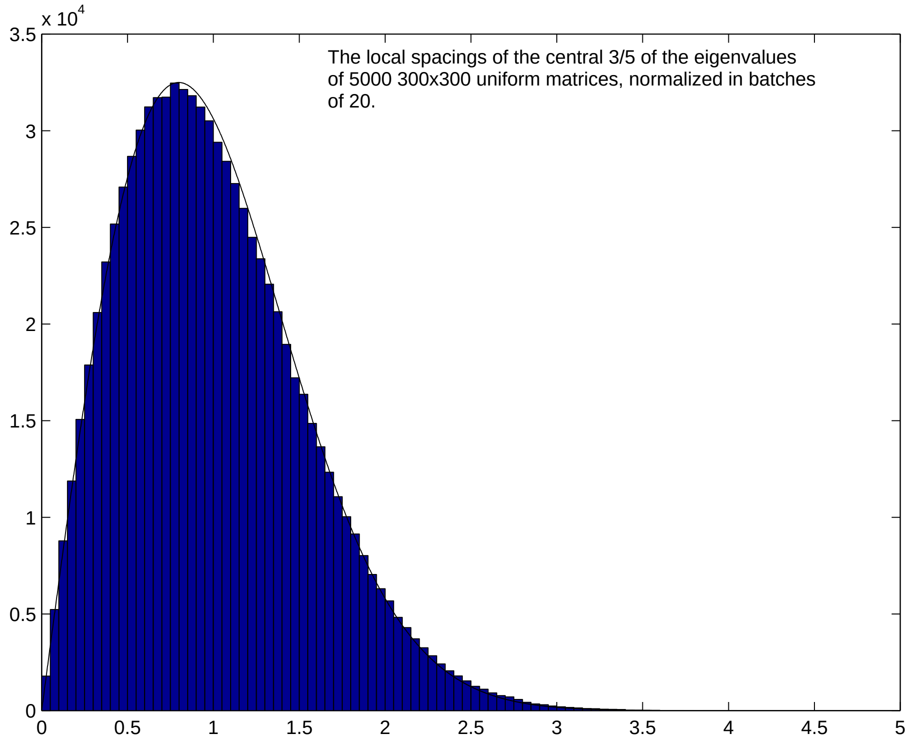
GOE Conjecture

GOE Conjecture: $N \times N$ Real Symmetric, entries iidrv. As $N \to \infty$, the probability density of the distance between two consecutive, normalized eigenvalues universal.

Only known if entries chosen from Gaussian.

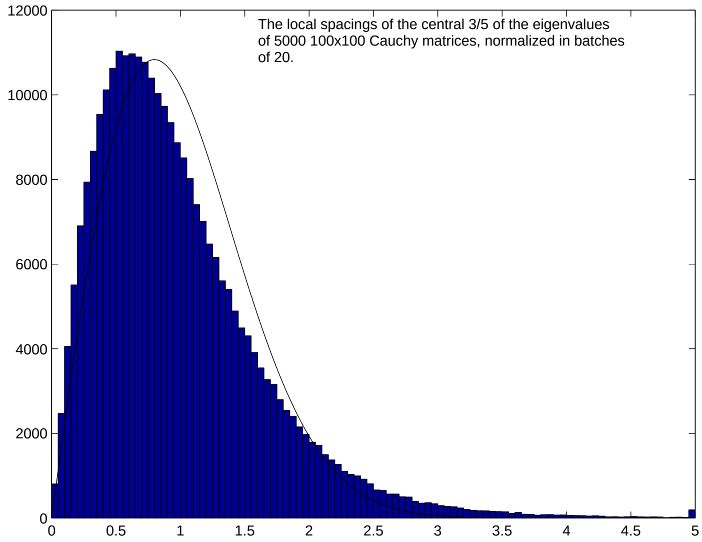
Consecutive spacings well approximated by Axe^{-Bx^2}.

Uniform Distribution: $p(x) = \frac{1}{2}$

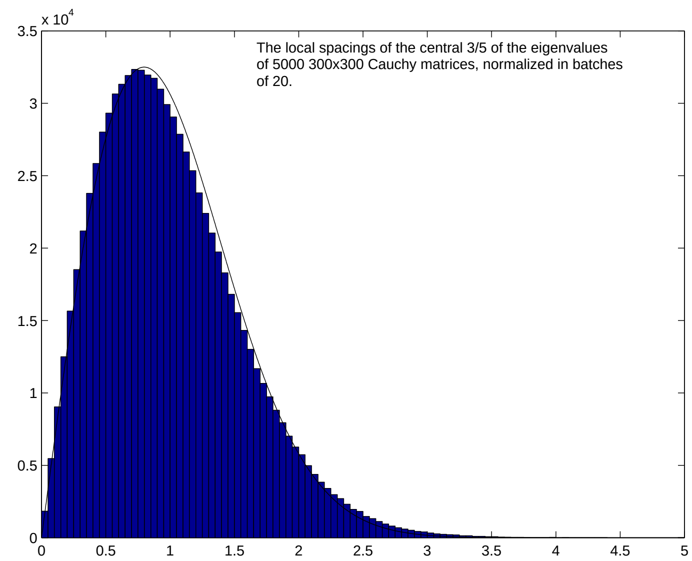


5000: 300×300 uniform on $[-1, 1]$

Cauchy Distribution: $p(x) = \frac{1}{\pi(1+x^2)}$



5000: 100×100 Cauchy



5000: 300×300 Cauchy

Fat Thin Families

Need a family ***F A T*** enough to do averaging.

Need a family ***THIN*** enough so that everything isn't averaged out.

Real Symmetric Matrices have $\frac{N(N+1)}{2}$ independent entries.

Examples of thin sub-families:

- Band Matrices
- Random Graphs
- Special Matrices (Toeplitz)

Band Matrices

Example of a Band 1 Matrix:

$$\begin{pmatrix} a_{11} & a_{12} & 0 & 0 & \cdots & 0 \\ a_{12} & a_{22} & a_{23} & 0 & \cdots & 0 \\ 0 & a_{23} & a_{33} & a_{24} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \cdots & a_{N-1,N} \\ 0 & 0 & 0 & \cdot\cdot & a_{N-1,N} & a_{NN} \end{pmatrix}$$

For Band 0 (Diagonal Matrices):

- Density of Eigenvalues: $p(x)$

- Spacings b/w eigenvalues: Poissonian.

Random Graphs

Degree of a vertex $=$ number of edges leaving the vertex.

Adjacency matrix: a_{ij} $=$ number edges from vertex i to vertex j.

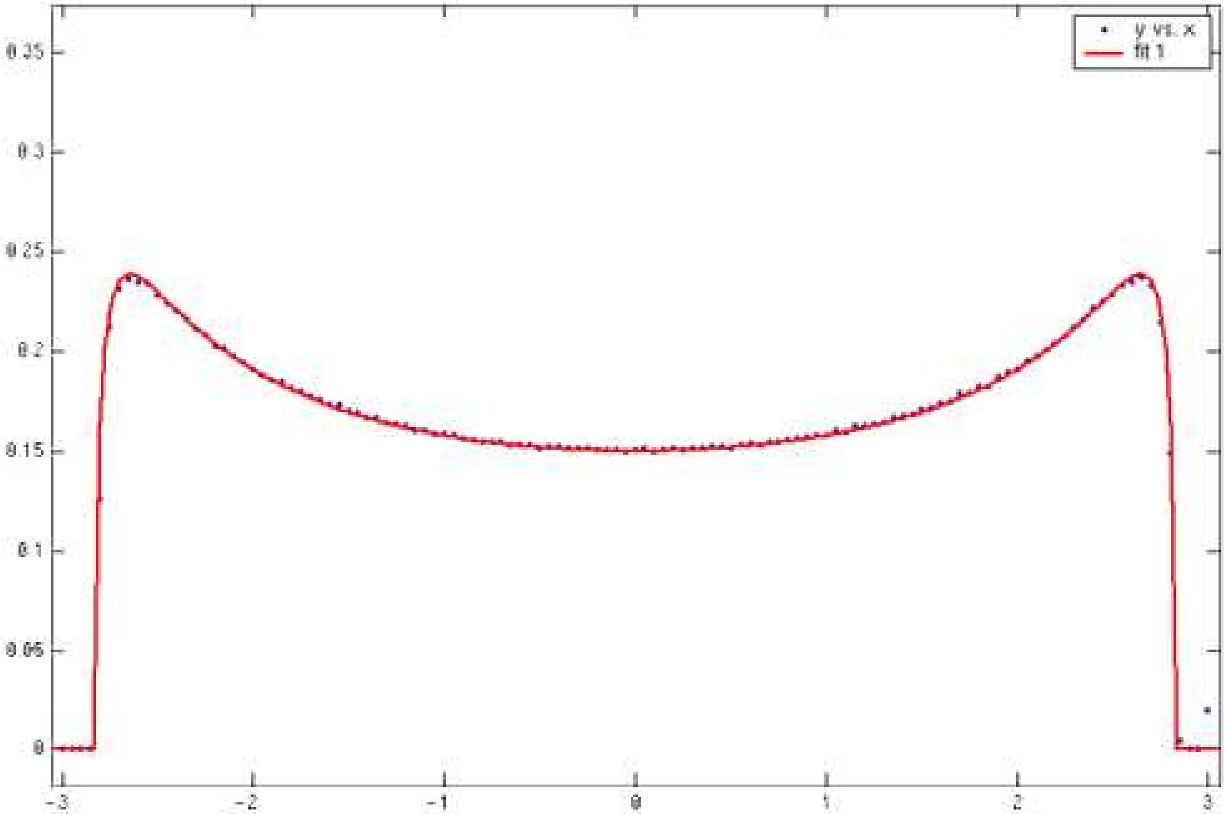
$$A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 2 \\ 1 & 0 & 2 & 0 \end{pmatrix}$$

These are Real Symmetric Matrices.

McKay's Law (Kesten Measure)

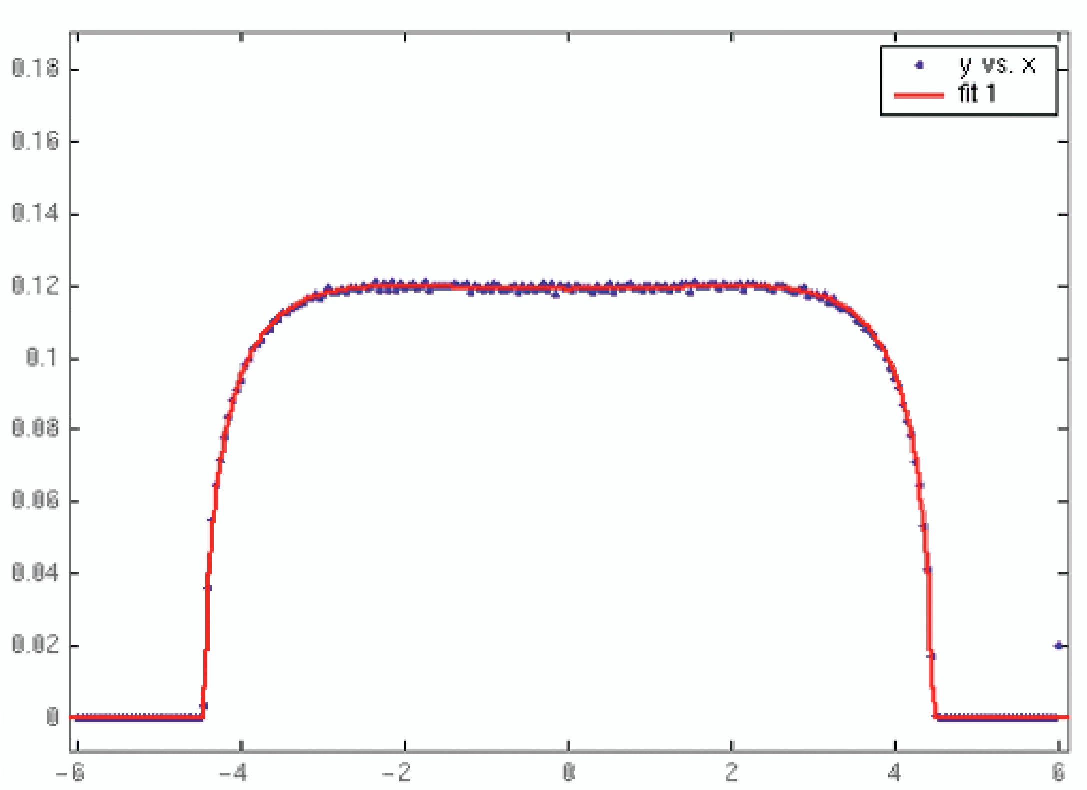
Density of States for d-regular graphs

$$f(x) = \begin{cases} \frac{d}{2\pi(d^2-x^2)}\sqrt{4(d-1)-x^2} & |x| \le 2\sqrt{d-1} \\ 0 & \text{otherwise} \end{cases}$$



$$d = 3.$$

McKay's Law (Kesten Measure)

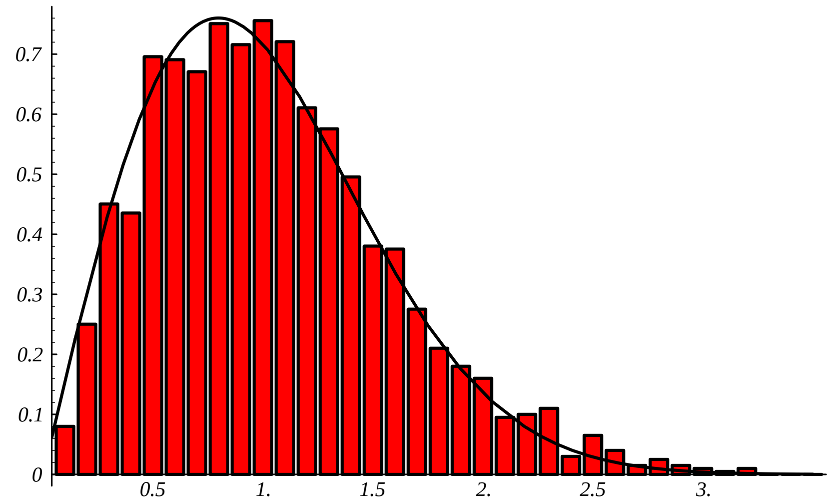


$$d = 6.$$

Idea of proof: Trace lemma, combinatorics and counting, locally a tree.

Fat Thin: fat enough to average, thin enough to get something different than Semi-circle.

d-Regular and GOE



3-Regular, 2000 Vertices: Graph courtesy of D. Jakobson, S. D. Miller, Z. Rudnick, R. Rivin

Toeplitz Ensembles

A Toeplitz matrix is of the form

$$\begin{pmatrix} b_0 & b_1 & b_2 & \cdots & b_{N-1} \\ b_{-1} & b_0 & b_1 & \cdots & b_{N-2} \\ b_{-2} & b_{-1} & b_0 & \cdots & b_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{1-N} & b_{2-N} & b_{3-N} & \cdots & b_0 \end{pmatrix}$$

- Will consider Symmetric Toeplitz.
- Main diagonal zero, N independent parameters.
- Normalize Eigenvalues by $\sqrt{N}$.

Eigenvalue Density Measure

$$\mu_{A,N}(x)dx \;=\; \frac{1}{N}\sum_{i=1}^{N}\delta\left(x-\frac{\lambda_i(A)}{\sqrt{N}}\right)dx.$$

The k^{th} moment of $\mu_{A,N}(x)$ is

$$M_k(A,N) \;=\; \frac{1}{N^{\frac{k}{2}+1}}\sum_{i=1}^{N}\lambda_i^k(A).$$

Let

$$M_k(N) \;=\; \lim_{N\to\infty} M_k(A,N).$$

$k = 0, 2$ and k odd

$\forall N$, $M_0(N) = 1$.

For $k = 2$: as $a_{ij} = b_{|i-j|}$:

$$\begin{aligned} M_2(N) &= \frac{1}{N^2} \sum_{1 \le i_1, i_2 \le N} \mathbb{E}(b_{|i_1 - i_2|} b_{|i_2 - i_1|}) \\ &= \frac{1}{N^2} \sum_{1 \le i_1, i_2 \le N} \mathbb{E}(b_{|i_1 - i_2|}^2). \end{aligned}$$

$N^2 - N$ times get 1, N times 0.

Therefore $M_2(N) = 1 - \frac{1}{N}$.

Trivial counting: odd moments $\longrightarrow 0$.

Even Moments

$$M_{2k}(N) \;=\; \frac{1}{N^{k+1}} \sum_{1\le i_1,\cdots,i_{2k}\le N} \mathbb{E}(b_{|i_1-i_2|}b_{|i_2-i_3|}\cdots b_{|i_{2k}-i_1|}).$$

Main Term: b_js matched in pairs, say

$$b_{|i_m-i_{m+1}|} \;=\; b_{|i_n-i_{n+1}|}, \quad x_m \;=\; |i_m-i_{m+1}| \;=\; |i_n-i_{n+1}|.$$

Two possibilities:

$$i_m - i_{m+1} \;=\; i_n - i_{n+1} \quad \text{or} \quad i_m - i_{m+1} \;=\; -(i_n - i_{n+1}).$$

$(2k-1)!!$ ways to pair, 2^k choices of sign.

Main Term: All Signs Neg

$$M_{2k}(N) \;=\; \frac{1}{N^{k+1}} \sum_{1\le i_1,\cdots,i_{2k}\le N} \mathbb{E}(b_{|i_1-i_2|}b_{|i_2-i_3|}\cdots b_{|i_{2k}-i_1|}).$$

Let $x_1,\dots,x_k$ be the values of the $|i_j - i_{j+1}|$s.

Let $\epsilon_1,\dots,\epsilon_k$ be the choices of sign.

Define $\widetilde{x}_1 = i_1 - i_2$, $\widetilde{x}_2 = i_2 - i_3, \dots$

$$\begin{aligned} i_2 &= i_1 - \widetilde{x}_1 \\ i_3 &= i_1 - \widetilde{x}_1 - \widetilde{x}_2 \\ &\vdots \\ i_1 &= i_1 - \widetilde{x}_1 - \cdots - \widetilde{x}_{2k}. \end{aligned}$$

Therefore

$$\widetilde{x}_1 + \cdots + \widetilde{x}_{2k} \;=\; \sum_{j=1}^{k}(1+\epsilon_j)\eta_j x_j \;=\; 0, \quad \eta_j \;=\; \pm 1.$$

Even Moments: Summary

Main Term: paired, all signs negative.

$$M_{2k}(N) \ \leq \ (2k-1)!! + O_k\left(\frac{1}{N}\right).$$

Bounded by Gaussian.

The Fourth Moment

$$M_4(N) = \frac{1}{N^3} \sum_{1 \le i_1, i_2, i_3, i_4 \le N} \mathbb{E}(b_{|i_1 - i_2|} b_{|i_2 - i_3|} b_{|i_3 - i_4|} b_{|i_4 - i_1|})$$

Let $x_j = |i_j - i_{j+1}|$.

Case One: $x_1 = x_2$**,** $x_3 = x_4$**:**

$$i_1 - i_2 \;=\; -(i_2 - i_3) \quad \text{and} \quad i_3 - i_4 \;=\; -(i_4 - i_1).$$

$i_1 = i_3$, i_2 and i_4 are arbitrary.

Left with $\mathbb{E}[b_{x_1}^2 b_{x_3}^2]$.

$N^3 - N$ times get 1, N times get $p_4 = \mathbb{E}[b_{x_1}^4]$.

Diophantine Obstruction

$$M_4(N) = \frac{1}{N^3} \sum_{1 \le i_1, i_2, i_3, i_4 \le N} \mathbb{E}(b_{|i_1 - i_2|} b_{|i_2 - i_3|} b_{|i_3 - i_4|} b_{|i_4 - i_1|})$$

Case Two: $x_1 = x_3$ and $x_2 = x_4$.

$$i_1 - i_2 \;=\; -(i_3 - i_4) \quad \text{and} \quad i_2 - i_3 \;=\; -(i_4 - i_1).$$

This yields

$$i_1 \;=\; i_2 + i_4 - i_3, \quad i_1, i_2, i_3, i_4 \in \{1, \dots, N\}.$$

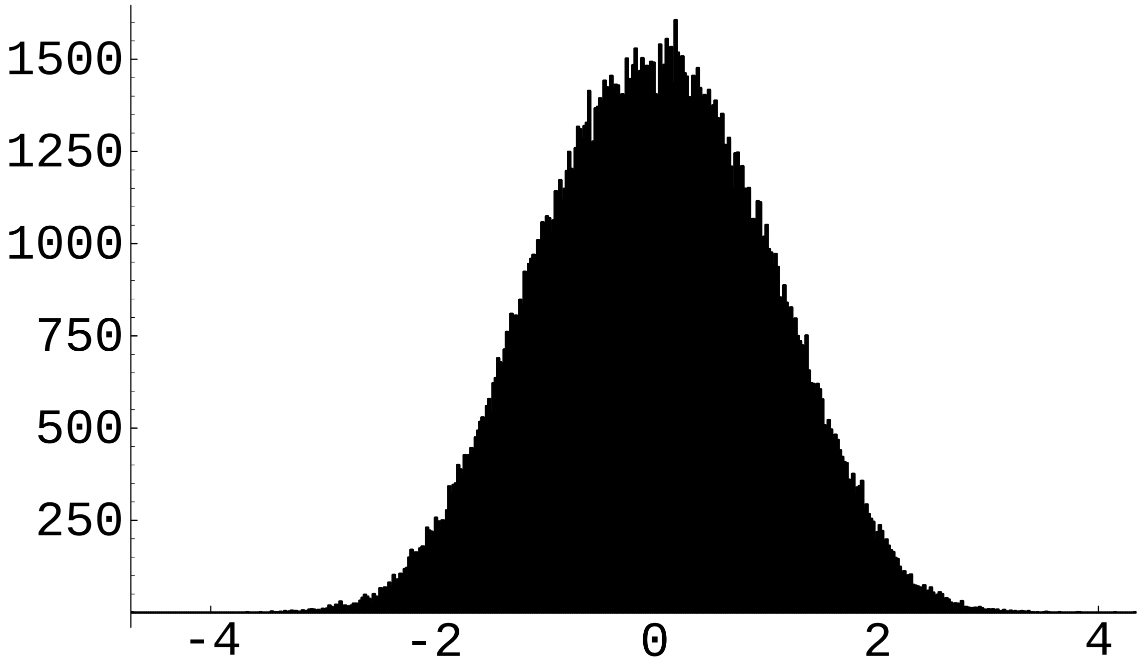
If $i_2, i_4 \geq \frac{2N}{3}$ and $i_3 < \frac{N}{3}$, $i_1 > N$: at most $(1 - \frac{1}{27})N^3$ valid choices.

Fourth Moment: Answer

Theorem: Fourth Moment: Let p_4 be the fourth moment of p. Then

$$M_4(N) \;=\; 2\frac{2}{3} + O_{p_4}\left(\frac{1}{N}\right).$$

500 Toeplitz Matrices, 400×400.

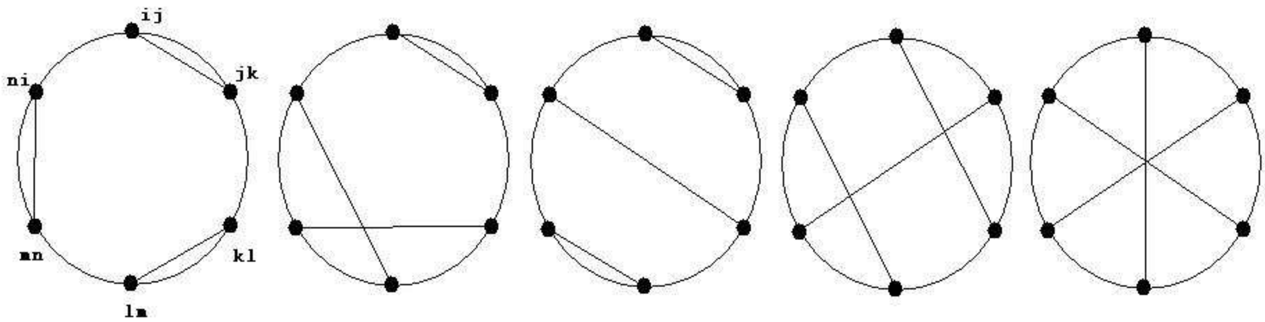


Higher Moments

$M_6(N) = 11$ (Gaussian's is 15).

$M_8(N) = 64\frac{4}{15}$ (Gaussian's is 105).

For sixth moment, five configurations:



Occur (respectively) $2, 6, 3, 3$ and 1 time.

Lemma: For $2k \geq 4$, $\lim_{N\to\infty} M_{2k}(N) < (2k-1)!!$.

Lower Bound of High Moments

unbounded support: a lower bound L_{2k} such that $\lim_{k\to\infty} \sqrt[2k]{L_{2k}} = \infty$.

$$\frac{1}{N^{k+1}}\mathbb{E}\left[\sum_{i_1=1}^{N}\cdots\sum_{i_{2k}=1}^{N} b_{|i_1-i_2|}b_{|i_2-i_3|}\cdots b_{|i_{2k}-i_1|}\right].$$

Matched in k pairs, matchings must occur with negative signs.

Denote positive differences of $|i_n - i_{n+1}|$ by $x_1, \dots, x_k$.

Let $\widetilde{x}_j = i_j - i_{j+1}$; k are positive (negative)

$k+1$ degrees of freedom: k differences x_j, and any index, say i_1.

Lower Bound of High Moments (II)

$$\begin{aligned} i_2 &= i_1 - \widetilde{x}_1 \\ i_3 &= i_1 - \widetilde{x}_1 - \widetilde{x}_2 \\ &\vdots \\ i_{2k} &= i_1 - \widetilde{x}_1 - \cdots - \widetilde{x}_{2k}. \end{aligned}$$

Once specify i_1 and $\widetilde{x}_1$ through $\widetilde{x}_{2k}$, all indices are determined.

If matched in pairs and each $i_j \in \{1, \dots, N\}$, have a valid configuration, contributes $+1$.

Problem: a running sum $i_1 - \widetilde{x}_1 - \cdots - \widetilde{x}_m \notin \{1, \dots, N\}$

Lower Bound of High Moments (III)

$\alpha \in (\frac{1}{2}, 1)$, $I_A = \{1, \dots, A\}$ where $A = \frac{N}{k^\alpha}$.

Choose each difference x_j from I_A: A^k ways, to first order distinct.

Put half of positive x_js and and non-matching negatives in first half $(\widetilde{x}_1, \dots, \widetilde{x}_k)$.

Have not specified the order of the differences, just how many positive (negative) are in the first half / second half.

Two different k-tuples of differences x_j *cannot* give rise to the same configuration (if we assume the differences are distinct).

Lower Bound of High Moments (IV)

Assume have ordered the positive differences in first half. There are $\binom{k}{2}!$ ways to relatively order corresponding negatives in second half.

If ordered the negatives in first half, $\binom{k}{2}!$ ways to relatively order positives in second half.

Still have freedom on how to intersperse positives and negatives in second half.

"Most" configurations can be made valid by Central Limit Theorem.

Lower Bound of High Moments (V)

Regard the $\widetilde{x}_p$s ($\widetilde{x}_n$s) as iidrv from I_A ($-I_A$) with mean $\approx \frac{1}{2}A$ ($\approx -\frac{1}{2}A$) and standard deviation $\approx \frac{1}{2\sqrt{3}}A$.

CLT: sum of the $\frac{k}{2}$ positive (negative) $\widetilde{x}_p$s ($\widetilde{x}_n$s) in the first block converges to Gaussian, mean $\approx \frac{kA}{4}$ ($\approx -\frac{kA}{4}$) and standard deviation $\approx \sqrt{\frac{k}{2}} \cdot \frac{A}{2\sqrt{3}}$.

N, k large, at least $\frac{1}{4}A^k$ have positive sum in

$$\left[\frac{kA}{4} - \frac{\sqrt{k}A}{2\sqrt{6}}, \frac{kA}{4} + \frac{\sqrt{k}A}{2\sqrt{6}}\right]$$

(and similarly for negative sum).

Freedom to choose how to intersperse the positives and negatives in the first and second halves. Keep running sum small: can ensure at most

$$\max\left(2A, 2\frac{\sqrt{k}A}{2\sqrt{6}}\right).$$

Intersperse same way in second half.

Lower Bound of High Moments (VI)

If $i_1 \in \left[\frac{7}{8}\frac{N}{k^{\alpha-\frac{1}{2}}}, \ \frac{N}{k^{\alpha-\frac{1}{2}}}\right]$, all indices in $\{1, \dots, N\}$.

Number of configurations giving 1 is at least

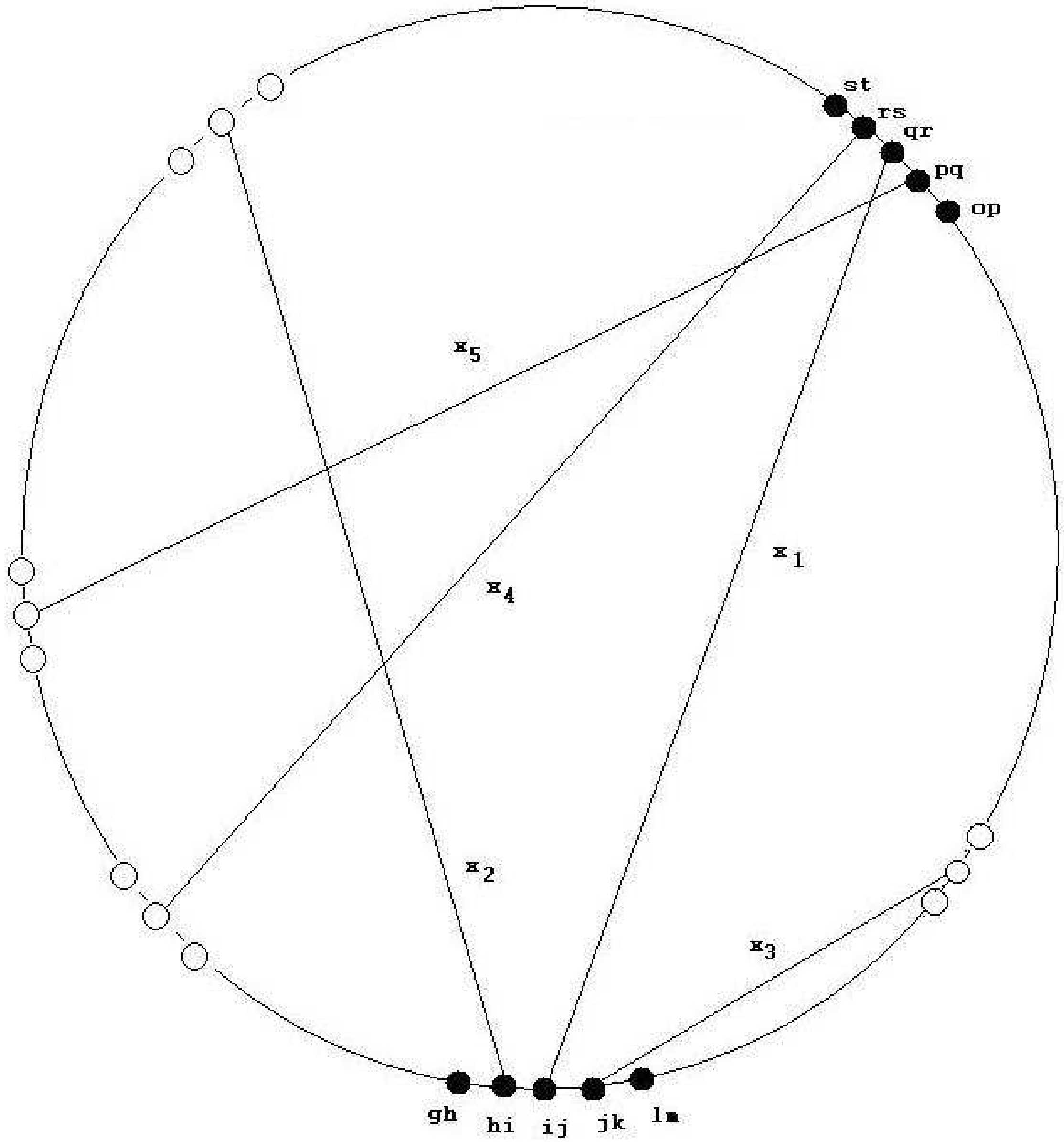
$$\left(\frac{1}{8}\frac{N}{k^{\alpha-\frac{1}{2}}}\right) \cdot \left(\frac{1}{4}A^k - \binom{k}{2}A^{k-1}\right) \cdot (k/2)!^2.$$

Divide by N^{k+1}, recall $A = \frac{N}{k^\alpha}$, main term:

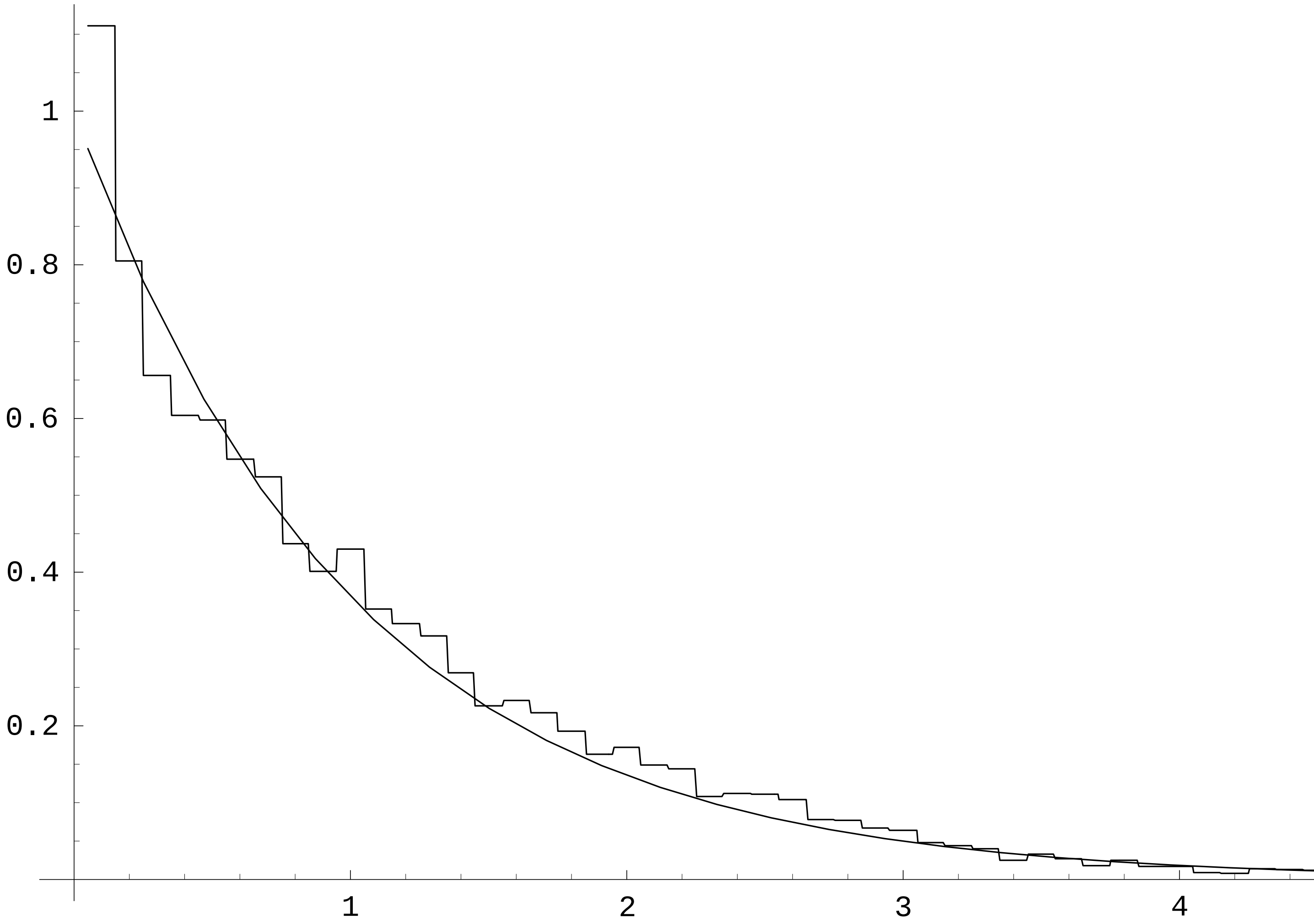
$$\frac{1}{32N^{k+1}} \frac{N^{k+1}}{k^{k\alpha-\frac{1}{2}}} \left(e^{\frac{k}{2}\log\frac{k}{2}-\frac{k}{2}}\sqrt{2\pi(k/2)}\right)^2 = \frac{\pi k^{\frac{3}{2}}}{16e^{(1+\log 2)k}} \cdot e^{(1-\alpha)k\log k}.$$

$2k^{\text{th}}$ root looks like $\frac{e^{(1-\alpha)\log k}}{e^{1+\log 2}} > O(k^{1-\alpha})$, proving the support is unbounded; $2k^{\text{th}}$ root of Gaussian moment is $\frac{k}{e}$.

Decay of Higher Moments



Poissonian Behavior?



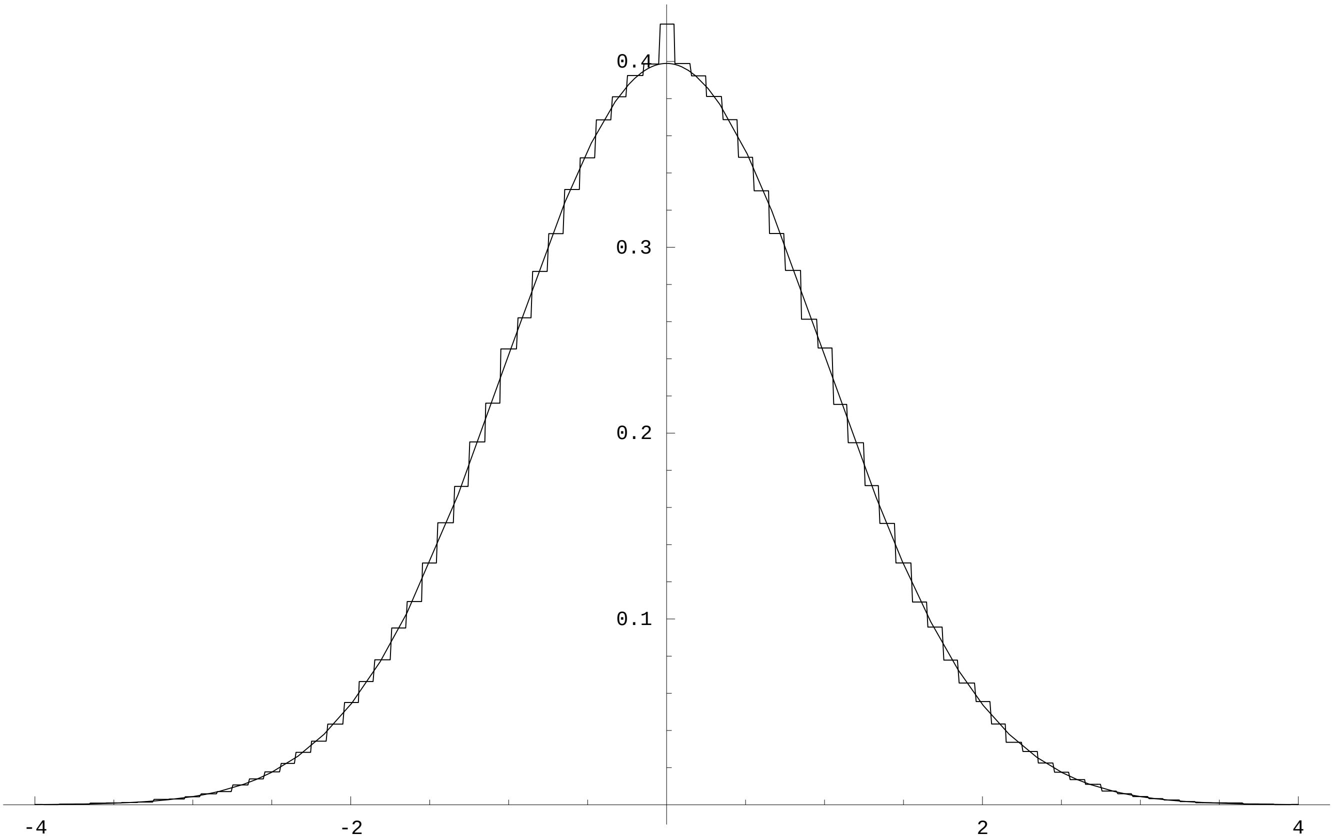
Not rescaled. Looking at middle 11 spacings, 1000 Toeplitz matrices (1000×1000), entries iidrv from the standard normal.

Real Symmetric Palindromic Toeplitz

$$\begin{pmatrix} b_0 & b_1 & b_2 & b_3 & \cdots & b_3 & b_2 & b_1 & b_0 \\ b_1 & b_0 & b_1 & b_2 & \cdots & b_4 & b_3 & b_2 & b_1 \\ b_2 & b_1 & b_0 & b_1 & \cdots & b_5 & b_4 & b_3 & b_2 \\ b_3 & b_2 & b_1 & b_0 & \cdots & b_6 & b_5 & b_4 & b_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ b_3 & b_4 & b_5 & b_6 & \cdots & b_0 & b_1 & b_2 & b_3 \\ b_2 & b_3 & b_4 & b_5 & \cdots & b_1 & b_0 & b_1 & b_2 \\ b_1 & b_2 & b_3 & b_4 & \cdots & b_2 & b_1 & b_0 & b_1 \\ b_0 & b_1 & b_2 & b_3 & \cdots & b_3 & b_2 & b_1 & b_0 \end{pmatrix}$$

- Extra symmetry seems to fix Diophantine Obstructions.
- Always have eigenvalue at 0.

Real Symmetric Palindromic Toeplitz (cont)



500 Real Symmetric Palindromic Toeplitz,
1000×1000

Note the bump at the zeroth bin is due to the forced eigenvalues at 0.

Summary

Ensemble	Degrees of Freedom	Density	Spacings
Real Symm	$O(N^2)$	Semi-Circle	GOE
Diagonal	$O(N)$	$p(x)$	Poisson
d-Regular	$O(dN)$	Kesten	GOE
Toeplitz	$O(N)$	Toeplitz	Poisson
Palindromic Toeplitz	$O(N)$	Gaussian	

Red is conjectured

Blue is new

Maroon: Partial Results (first 9 moments)

Summary (Toeplitz)

- Converges (in some sense) to new universal distribution, independent of p;

- Moments bounded by those of Gaussian;

- Ratio tends to 0, but unbounded support;

- Can interpret as Diophantine Obstructions (Hammond-Miller) or volumes of Euler solids (Bryc-Dembo-Jiang).

- Real Symmetric Palindromic Toeplitz looks Gaussian (Miller-Ramey-Sinsheimer-Teich).