Random regular digraphs: singularity and spectrum

Nick Cook, UCLA

Probability Seminar, Stanford University
November 2, 2015
Talk outline

1. Universality for (global eigenvalue statistics of non-hermitian) random matrices

2. Random regular digraphs (adjacency matrices), and conjectured limiting spectral distributions

3. Two results:
   - Circular law for signed random regular digraphs
   - Bound on singularity probability for random regular digraphs
The circular law for i.i.d. matrices

**Definition (i.i.d. matrix)**

Let $x$ be a $\mathbb{C}$-valued random variable with

$$
\mathbb{E} x = 0, \quad \mathbb{E} |x|^2 = 1.
$$

For each $n$, let $X_n = (x_{ij})_{1 \leq i,j \leq n}$ have entries that are i.i.d. copies of $x$.


Let $\{\lambda_k(X_n)\}_{k=1}^n$ be the eigenvalues of $X_n$.

Define the (rescaled) empirical spectral distribution (ESD) of $X_n$:

$$
\mu_n := \frac{1}{n} \sum_{k=1}^n \delta_{\frac{1}{\sqrt{n}} \lambda_k(X_n)}.
$$

Almost surely, $\mu_n \Rightarrow \frac{1}{\pi} \mathbb{1}_{B_{\mathbb{C}}(0,1)} dx dy$. 

N. Cook
The circular law for i.i.d. matrices

Figure: Circular law universality class: eigenvalue plots for randomly generated $5000 \times 5000$ matrices using Bernoulli random variables (left) and Gaussian random variables (right). Figure by Philip Matchett Wood.
The circular law universality class
The circular law universality class

Circular law

- Bernoulli
The circular law universality class

Circular law

- Bernoulli
- Ginibre
The circular law universality class

- iid, finite 2nd moment (Tao-Vu '08)
  - Bernoulli
  - Ginibre

Circular law
The circular law universality class

iid, finite 2nd moment
(Tao-Vu '08)

- Bernoulli
- Ginibre
The circular law universality class

- iid, heavier tails
  - (Bordenave, Caputo, Chafaï '10)

- iid, finite 2nd moment
  - (Tao-Vu '08)

- Bernoulli
- Ginibre
The circular law universality class

- iid, heavier tails (Bordenave, Caputo, Chafaï '10)
- iid, finite 2nd moment (Tao-Vu '08)
  - Bernoulli
  - Ginibre

Dependent entries?
The circular law universality class

- iid, heavier tails (Bordenave, Caputo, Chafaï '10)
- iid, finite 2nd moment (Tao-Vu '08)
  - Bernoulli
  - Ginibre

Circular law

- Uniform doubly-stoch. (Nguyen '12)

Dependent entries?
The circular law universality class

- iid, heavier tails (Bordenave, Caputo, Chafaï '10)
- iid, finite 2nd moment (Tao-Vu '08)
  - Bernoulli
  - Ginibre
- Uniform doubly-stoch. (Nguyen '12)
- Unconditional log-concave (Adamczak, Chafaï '13)

Dependent entries?
The circular law universality class

- iid, heavier tails
  (Bordenave, Caputo, Chafaï '10)

- iid, finite 2nd moment
  (Tao-Vu '08)

- Bernoulli
- Ginibre

- Uniform doubly-stoch.
  (Nguyen '12)

- Unconditional log-concave
  (Adamczak, Chafaï '13)

Dependent entries?
The r.r.d. matrix ensemble

- $n$ large, $d \in [n]$
- $\mathcal{M}_{n,d} := \{ n \times n \text{ matrices, entries } \in \{0, 1\}, \text{ all row and column sums equal to } d \}$
  = \{ adjacency matrices of $d$-regular digraphs on $n$ vertices \}

\[
\begin{pmatrix}
1 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 \\
1 & 0 & 1 & 0
\end{pmatrix}
\]

Let $A \in \mathcal{M}_{n,d}$ uniform random.

“Random regular digraph (r.r.d.) matrix”
The circular law universality class

- iid, heavier tails (Bordenave, Caputo, Chafaï '10)
- iid, finite 2nd moment (Tao-Vu '08)
- Bernoulli
- Ginibre
- Uniform doubly-stoch. (Nguyen '12)
- Unconditional log-concave (Adamczak, Chafaï '13)

Dependent entries?
The circular law universality class

- iid, heavier tails
  (Bordenave, Caputo, Chafaï '10)

- iid, finite 2nd moment
  (Tao-Vu '08)

  - Bernoulli
  - Ginibre

- Exchangeable array + moment hypothesis
  (Adamczak, Chafaï, Wolff '14)

- Uniform doubly-stoch.
  (Nguyen '12)

- Unconditional log-concave
  (Adamczak, Chafaï '13)

- rrd matrix (conjectured)

- Dependent entries?
The circular law universality class

- iid, heavier tails (Bordenave, Caputo, Chafaï '10)
- iid, finite 2nd moment (Tao-Vu '08)
- Bernoulli
- Ginibre
- Exchangeable array + moment hypothesis (Adamczak, Chafaï, Wolff '14)
- Circular law
- rrd matrix $d \to \infty$
- Uniform doubly-stoch. (Nguyen '12)
- Unconditional log-concave (Adamczak, Chafaï '13)

Dependent entries?
The circular law universality class

- iid, heavier tails (Bordenave, Caputo, Chafaï '10)
- iid, finite 2nd moment (Tao-Vu '08)
- Bernoulli
- Ginibre
- Exchangeable array + moment hypothesis (Adamczak, Chafaï, Wolff '14)
- rrd matrix (conjectured)
  - \( d \to \infty \)
  - \( d \) fixed
- Uniform doubly-stoch. (Nguyen '12)
- Unconditional log-concave (Adamczak, Chafaï '13)

Dependent entries?
The circular law universality class

- iid, heavier tails (Bordenave, Caputo, Chafaï '10)
- iid, finite 2nd moment (Tao-Vu '08)
  - Bernoulli
  - Ginibre
- Exchangeable array + moment hypothesis (Adamczak, Chafaï, Wolff '14)

**Circular law**
- rrd matrix
  - $d \rightarrow \infty$
- Uniform doubly-stoch. (Nguyen '12)
- Unconditional log-concave (Adamczak, Chafaï '13)

**Oriented Kesten-McKay law**
- (conjectured)
  - $d$ fixed

Dependent entries?
The circular law universality class

iid, heavier tails (Bordenave, Caputo, Chafaï '10)

iid, finite 2nd moment (Tao-Vu '08)

- Bernoulli
- Ginibre
- Exchangeable array + moment hypothesis (Adamczak, Chafaï, Wolff '14)

Circular law

*Uniform doubly-stoch.* (Nguyen '12)

*Unconditional log-concave* (Adamczak, Chafaï '13)

*Dependent entries?*

Oriented Kesten-McKay law

*rrd matrix* $d \to \infty$

*Sum of $d$ iid Haar unitaries* (Basak, Dembo '12)

*d fixed*
Figure: Empirical eigenvalue distributions for simulated $8000 \times 8000$ rescaled r.r.d. matrices $\frac{1}{\sqrt{d}} A$ for $d = 3$ (left), 10 (middle), and 100 (right). Predictions from the oriented Kesten–McKay law are plotted in red.
We consider signed r.r.d. matrices $A \circ X = (a_{ij}x_{ij})$, where

- $A \in \mathcal{M}_{n,d}$ is an r.r.d. matrix,
- $X$ is an i.i.d. matrix with $\pm 1$ Bernoulli entries, independent of $A$.

**Theorem (C. '15)**

Fix $p \in (0, 1)$ and put $d = \lfloor pn \rfloor$. Then as $n \to \infty$, the empirical spectral distribution of $\frac{1}{\sqrt{d}}A \circ X$ converges weakly in probability to the uniform measure on $B_{\mathbb{C}}(0, 1)$.

- Stated for i.i.d. signs, but the proof only needs the entries of $X$ to have $4 + \varepsilon$ finite moments.
- **Work in progress:** remove $X$, extend to sparse case $d = o(n)$ (more on this later).
Girko’s Hermitization approach

For a Borel probability measure $\mu$, define the \textit{log potential}:

$$U_\mu(z) := \int_C \log |\lambda - z| \, d\mu(\lambda).$$

Two sides to why this is useful:

1) Borel measures on $\mathbb{C}$ are characterized by their log-potentials:

$$\mu = \frac{1}{2\pi} \Delta U_\mu.$$

2) Determinant identity:

$$\prod_{i=1}^n |\lambda_i(M)| = |\text{det}(M)| = \prod_{i=1}^n s_i(M)$$

where $s_1(M) \geq \cdots \geq s_n(M)$ are the singular values.
How does one prove circular laws?

Putting these together:

- For a sequence of $n \times n$ matrices $(M_n)_{n \geq 1}$, to show
  \[ \mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(M_n)} \] converges, suffices to show pointwise convergence of

  \[ U_{\mu_n}(z) = \int_{\mathbb{C}} \log |\lambda - z| d\mu_n(\lambda) = \frac{1}{n} \sum_{i=1}^{n} \log |\lambda_i(M_n - zI_n)| \]

  \[ = \frac{1}{n} \sum_{i=1}^{n} \log s_i(M_n - zI_n) = \int_{\mathbb{R}_+} \log(s) d\nu_{M_n - zI_n}(s). \]

- **Gain:** $\nu_{M_n - zI_n}$ are ESDs of *Hermitian* random matrices, which are (for our purposes) well understood.

- **Loss:** $s \mapsto \log(s) \notin BC(\mathbb{R}_+)$, has singularities at 0 and $\infty$. 
For a signed r.r.d. matrix $A_n \circ X_n$, write $\nu_{n,z} = \nu \frac{1}{\sqrt{d}} A_n \circ X_n - zI_n$.

Step 1: Show $\nu_{n,z}$ converges weakly in probability to a deterministic limit $\nu_z$ for all $z \in \mathbb{C}$.

i.e. $\forall f \in BC(\mathbb{R}_+), \forall \varepsilon > 0$,

$$\mathbb{P} \left( \left| \int_{\mathbb{R}_+} f \, d\nu_{n,z} - \int_{\mathbb{R}_+} f \, d\nu_z \right| > \varepsilon \right) = o(1)$$

Step 2: Prove bounds on extreme singular values.

2a) Show $s_1(\frac{1}{\sqrt{d}} A_n \circ X_n - zI_n) = O(1)$ with high probability (w.h.p.)

2b) Show $s_n(\frac{1}{\sqrt{d}} A_n \circ X_n - zI_n) \geq n^{-c}$ w.h.p.
Proof outline

For a signed r.r.d. matrix $A_n \circ X_n$, write $\nu_{n,z} = \nu \frac{1}{\sqrt{d}} A_n \circ X_n - zI_n$.

Step 1: Show $\nu_{n,z}$ converges weakly in probability to a deterministic limit $\nu_z$ for all $z \in \mathbb{C}$.

i.e. $\forall f \in BC(\mathbb{R}^+), \forall \varepsilon > 0,$

$$\mathbb{P} \left( \left| \int_{\mathbb{R}^+} f \, d\nu_{n,z} - \int_{\mathbb{R}^+} f \, d\nu_z \right| > \varepsilon \right) = o(1)$$

Step 2: Prove bounds on extreme singular values.

2a) Show $s_1(\frac{1}{\sqrt{d}} A_n \circ X_n - zI_n) = O(1)$ with high probability (w.h.p.)

2b) Show $s_n(\frac{1}{\sqrt{d}} A_n \circ X_n - zI_n) \geq n^{-c}$ w.h.p.
Step 1: weak convergence of singular value distributions

**Step 1:** prove weak convergence of empirical singular value distributions

\[ \nu_{n,z} = \nu \frac{1}{\sqrt{d}} A \circ X - z I = \frac{1}{n} \sum_{i=1}^{n} \delta_{s_i} \left( \frac{1}{\sqrt{d}} A \circ X - z I \right). \]

Idea (following Tran–Vu–Wang '10):
- Replace \( A \) with a 0/1 matrix
  \[ B = (b_{ij})_{1 \leq i, j \leq n}, \quad b_{ij} \text{ i.i.d. Bernoulli}(d/n) \]
  independent of \( X \). \( B \circ X \) has i.i.d. entries.
- Note \( A \overset{d}{=} B \mid \{ B \in \mathcal{M}_{n,d} \} \).
  For a “bad event” \( B \) we can bound

\[ \mathbb{P}(A \in B) = \mathbb{P}(B \in B \mid B \in \mathcal{M}_{n,d}) \leq \frac{\mathbb{P}(B \in B)}{\mathbb{P}(B \in \mathcal{M}_{n,d})}. \]
Step 1: a comparison trick

For a “bad event” $\mathcal{B}$ we can bound

$$\mathbb{P}(A \in \mathcal{B}) = \mathbb{P}(B \in \mathcal{B} | \mathcal{E}_{n,d}) \leq \frac{\mathbb{P}(B \in \mathcal{B})}{\mathbb{P}(B \in \mathcal{M}_{n,d})}.$$ 

Lemma (Tran)

$$\mathbb{P}(B \in \mathcal{M}_{n,d}) = \exp(-O(n\sqrt{d})).$$

Want to show: for any $f \in BC(\mathbb{R}_+)$, $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\int_{\mathbb{R}_+} f \, d\nu_{n,z} - \int_{\mathbb{R}_+} f \, d\nu_z\right| > \varepsilon\right) = o(1)$$

Denoting $\tilde{\nu}_{n,z} = \nu \frac{1}{\sqrt{d}} B \circ X_{-zI}$, it suffices to show

$$\mathbb{P}\left(\left|\int_{\mathbb{R}_+} f \, d\tilde{\nu}_{n,z} - \int_{\mathbb{R}_+} f \, d\nu_z\right| > \varepsilon\right) \ll e^{-Cn\sqrt{d}}.$$
Step 1: a comparison trick

Want to show: \( \mathbb{P} \left( \left| \int_{\mathbb{R}^+} f \, d\tilde{\nu}_{n,z} - \int_{\mathbb{R}^+} f \, d\nu_z \right| > \varepsilon \right) \ll e^{-Cn\sqrt{d}}. \)

- Desired bound is too small to apply work of Bourgade–Yau–Yin ’12 on the local law.
- Instead we go back to an argument of Guionnet–Zeitouni ’00:
  - Lemma: if \( f : \mathbb{R}_+ \rightarrow \mathbb{R} \) is convex and 1-Lipschitz, then
    \[
    F = B \mapsto \int_{\mathbb{R}^+} f d\nu \frac{1}{\sqrt{d}} B \circ X - z I
    \]
    is convex and 1-Lipschitz on \( \mathcal{M}_n(\mathbb{C}) \) (in Frobenius norm).
  - Applying Talagrand’s isoperimetric inequality:
    \[
    \mathbb{P} \left( |F(B) - \mathbb{E} F(B)| \geq \varepsilon \right) = O(e^{-c\varepsilon^2 nd}).
    \]

  Extend to general \( f \) by an approximation argument.

- This argument applies for \( A \) drawn uniformly from any set \( S \subset \mathcal{M}_n(\{0,1\}) \) satisfying \( \mathbb{P}(B \in S) \geq \exp(-o(nd)) \).
Step 2: smallest singular value

- Consider a random $n \times n$ matrix of the form

$$M = A \circ X + B$$

with: $X$ i.i.d., $A$ fixed 0/1 matrix, $B$ fixed.

- We control the lower tail of $s_n(M)$ under a quasirandomness hypothesis on $A$
  ("super-regularity", c.f. Szemerédi’s regularity lemma).

**Theorem (C. ’15)**

Assume $A$ satisfies [quasirandomness hypothesis], $\|B\| = O(\sqrt{n})$, and $|x_{ij}| = O(1)$ for all $i, j \in [n]$. Then for all $t > 0$,

$$\mathbb{P} \left( s_n(M) \leq tn^{-1/2} \right) \lesssim t + n^{-1/2}.$$ 

- Similar result by Rudelson–Zeitouni for the case that $x_{ij}$ are Gaussian, under a weaker expansion-type assumption on $A$.

- From (C. ’14): the r.r.d. matrix $A$ is super-regular w.h.p.
We can extend the argument for Step 1 (convergence of singular value distributions) to the r.r.d. matrix $A$ with $d = n^\epsilon$.

The main difficulty is to obtain control of the least singular value. In this direction we have the following:

**Theorem (C. '14)**

There are absolute constants $C, c > 0$ such that the following holds. If $C \log^2 n \leq d \leq \frac{n}{2}$, then

$$\mathbb{P}(s_n(A) = 0) = O(d^{-c}).$$

(We can take $c = .05$.)

**Conjecture**

There are constants $C, c > 0$ such that for any $d \in [3, n - 3]$,

$$\mathbb{P}(s_n(A) = 0) \leq Cn^{-c}.$$
Spectral concentration from classical concentration

- Proofs of upper bounds on $s_1(M) = \|M\|_{op}$ reduce to an application of *concentration of measure*.
- Proofs of lower bounds on $s_n(M) = \|M^{-1}\|_{op}^{-1}$ reduce to the application of *anti-concentration* or “small ball” estimates.

**Theorem (Anti-concentration for random walks, Erdős ’40s)**

Let $\xi_1, \ldots, \xi_n$ be i.i.d. uniform Bernoulli signs, and $x \in \mathbb{R}^n$. Then for any $a \in \mathbb{R}$,

$$\mathbb{P} \left( \sum_{j=1}^{n} \xi_j x_j = a \right) \lesssim |\{j : x_j \neq 0\}|^{-1/2}.$$

- More sophisticated bounds have been developed by Tao–Vu and Rudelson–Vershynin using *Inverse Littlewood-Offord theory*.
- This is our hammer – where is the nail?
Local symmetries: switchings (after McKay)

- In a regular digraph, we can change between vertices $i_1, i_2, j_1, j_2$ and preserve $d$-regularity.
- In the adjacency matrix, this corresponds to switching between

$$
I_2 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad J_2 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
$$

at the $(i_1, i_2) \times (j_1, j_2)$ minor.
- Idea: apply several independent switchings, encode outcomes with i.i.d. signs $\xi_j$. 

Idea: apply several independent switchings, encode outcomes with i.i.d. signs $\xi_j$. 

Idea: apply several independent switchings, encode outcomes with i.i.d. signs $\xi_j$. 

Idea: apply several independent switchings, encode outcomes with i.i.d. signs $\xi_j$. 

Idea: apply several independent switchings, encode outcomes with i.i.d. signs $\xi_j$. 

Idea: apply several independent switchings, encode outcomes with i.i.d. signs $\xi_j$. 

Idea: apply several independent switchings, encode outcomes with i.i.d. signs $\xi_j$. 

Idea: apply several independent switchings, encode outcomes with i.i.d. signs $\xi_j$. 

Idea: apply several independent switchings, encode outcomes with i.i.d. signs $\xi_j$. 

Idea: apply several independent switchings, encode outcomes with i.i.d. signs $\xi_j$.
Where is the nail?

\[
\begin{align*}
\begin{bmatrix}
1 & 1 & \cdots & 1 \\
1 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix}
\end{align*}
\]

Conditional on \( R_3, \ldots, R_n \), the only randomness is in the choice of sets \( \text{Ex}(1, 2), \text{Ex}(2, 1) \).

Let \( \pi : \text{Ex}(1, 2) \rightarrow \text{Ex}(2, 1) \) uniform random bijection.

Conditional on \( \pi \), independently resample the 2 \times 2 minors \( M_{(1,2) \times (j, \pi(j))} \).
Conditional on $R_3, \ldots, R_n$, the only randomness is in the choice of sets $Ex(1, 2)$, $Ex(2, 1)$.

Let $\pi : Ex(1, 2) \to Ex(2, 1)$ uniform random bijection.

Conditional on $\pi$, independently resample the $2 \times 2$ minors $M_{(1,2) \times (j, \pi(j))}$. 
Where is the nail?

In the randomness of the resampling, $R_1 \cdot u$ is a random walk with steps $u_j - u_{\pi(j)}$. (Found the nail!)

Key technical proposition: normal vectors $u$ have small level sets.

Combining this with the randomness of $\pi$ guarantees most steps are non-zero.

What if $\text{Ex}(1, 2)$ is small?

\[
\begin{bmatrix}
1 & \cdots & 1 & 1 & 0 & \cdots & 1 & 0 & 1 & \cdots & 0 & 0 & \cdots \\
2 & \cdots & 1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 1 & 0 & \cdots \\
\vdots & & & & & & & & & & & & \\
\end{bmatrix}
\]
Problem: what if vertices 1, 2 have large codegree?

Solution: use the method of exchangeable pairs for concentration of measure (Chatterjee '06) with a “reflection” coupling to show codegrees concentrate around $d^2/n$.

Also obtain control on edge densities:

For $S, T \subset [n]$ and $\varepsilon \geq 0$,

$$
\mathbb{P} \left( \left| \frac{e(S, T)}{\frac{d}{n} |S||T|} - 1 \right| \geq \varepsilon \right) \leq 2 \exp \left( - \frac{c\varepsilon^2 d}{1 + \varepsilon \frac{d}{n} |S||T|} \right).
$$

In recent work with Larry Goldstein and Toby Johnson, we obtain exponential tail bounds for more general statistics using size biased couplings.

Allowed us to extend a bound $\lambda_2(A) = O(\sqrt{d})$ on the second eigenvalue of a random regular (undirected) graph to allow $d = O(n^{2/3})$ (previous results were limited to $d = o(n^{1/2})$).
Summary of toy problem

- To show
  \[ \mathbb{P} \left( R_1 \in \text{span}(R_3, \ldots, R_n) \right) = o(1) \]
  
  we defined a coupling \((M, \tilde{M}, \pi, \xi)\) on an enlarged probability space, with \(M \overset{d}{=} \tilde{M}\), and sought to show
  \[ \mathbb{P} \left( \tilde{R}_1 \in \text{span}(R_3, \ldots, R_n) \mid M \right) = o(1). \]

1. **The randomness of** \(M\): \(\text{Ex}(1, 2)\) is large with high probability.
2. **The randomness of** \(\pi\): the random walk \(\tilde{R}_1 \cdot u\) takes many steps with high probability.
3. **The randomness of** \(\xi = (\xi_1, \ldots, \xi_n)\) (encoding the resampling of \(2 \times 2\) minors): used with Erdős’ anti-concentration bound to finish the proof.