The maximum of the characteristic polynomial for a random permutation matrix

Random Matrices and Free Probability Workshop

IPAM, 2018/05/16

Nick Cook, UCLA

Based on joint work with Ofer Zeitouni

Model and previous work

- Let P_N be a uniform random $N \times N$ permutation matrix and let $\chi_N(z) = \det(zI_N - P_N)$ denote its characteristic polynomial.
- Consider the random field $X_N : \mathbb{R}/\mathbb{Z} \to [-\infty, \infty)$

$$X_N(t) = \log|\chi_N(e(-t))| = \log|\det(I_N - e(t)P_N)|$$

  where $e(t) := \exp(2\pi\mathrm{i}t)$.
- Hambly–Keevash–O'Connell–Stark '99 obtained a CLT:
  For fixed $t \in \mathbb{R}/\mathbb{Z}$ of finite type,

$$\frac{X_N(t)}{\sqrt{\frac{\pi^2}{12}\log N}} \to N(0,1)$$

  (and similarly for the imaginary part of $\log \chi_N$).
- Note X_N is badly behaved at rational points (atom at $-\infty$).
- Multidimensional CLT obtained by Dang–Zeindler '13.

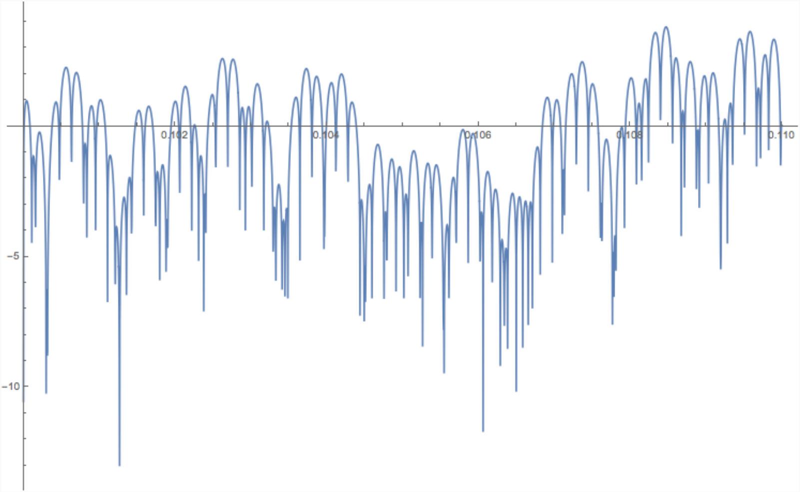


Figure 1: Simulated $X_N(t)$ with $N = 10^4$ for $t \in [0.1, 0.11]$.
Cycle structure of P_N: 6310, 1914, 909, 668, 79, 47, 33, 19, 12, 5, 3, 1.
(Generated using the Chinese restaurant process.)

Numerical simulations

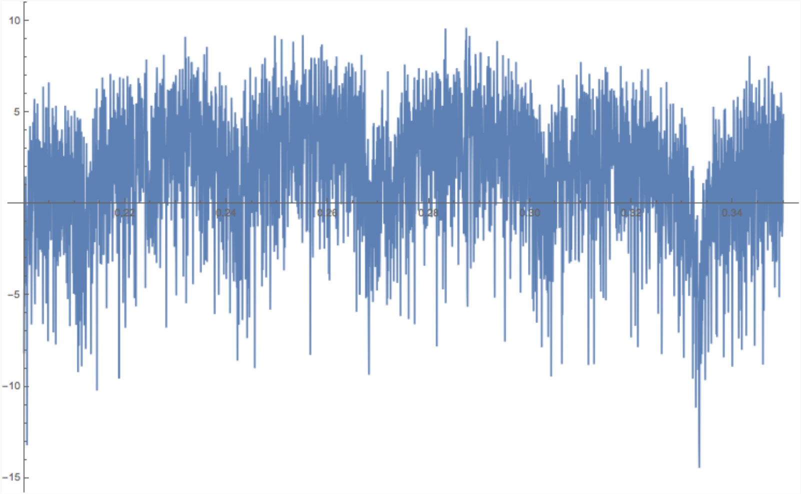


Figure 2: Simulated $X_N(t)$ with $N = 10^9$ for $t \in [0.2, 0.35]$.
Cycle structure of P_N: 892,060,223, 78,087,020, 19,479,718, 9,152,317, 630,684, 352,623, 114,502, 104,059, 8,973, 8,193, 1,641, 33, 5, 3, 2, 2, 1, 1.

Main result: Law of large numbers for the maximum of X_N

$X_N(t) = \log|\det(I_N - e(t)P_N)|.$

Theorem (C., Zeitouni '18)

$$\frac{1}{\log N} \sup_{t\in\mathbb{R}/\mathbb{Z}} X_N(t) \to x_c \approx 0.677 \quad \text{in probability.}$$

(Informally:

$$\sup_{|z|=1} |\chi_N(z)| = N^{x_c + o(1)} \qquad \text{with high probability.)}$$

Related work: Maximum of the CUE field

Replacing P_N with a Haar unitary U_N, we obtain the CUE field:

$$X_N^{\text{cue}}(t) = \log|\det(I_N - e(t)U_N)|.$$

Conjecture (Fyodorov–Hiary–Keating '12):

$$M_N := \sup_{t\in\mathbb{R}/\mathbb{Z}} X_N^{\text{cue}}(t) - \left(\log N - \frac{3}{4}\log\log N\right) \quad \text{converges in distribution.}$$

$$\sup_{t\in\mathbb{R}/\mathbb{Z}} X_N^{\text{cue}}(t) = \log N + o_P(\log N)$$ Arguin–Belius–Bourgade '15

Related work of Arguin–Belius–Bourgade–Soundararajan–Radziwiłł'16 on the Riemann ζ function.

All proofs proceed by exposing an underlying branching structure.

Note that, unlike X_N, distribution of X_N^{cue} is invariant under rotations.

Related work: Maximum of the CUE field

Replacing P_N with a Haar unitary U_N, we obtain the CUE field:

$$X_N^{\text{cue}}(t) = \log|\det(I_N - e(t)U_N)|.$$

Conjecture (Fyodorov–Hiary–Keating '12):

$$M_N := \sup_{t\in\mathbb{R}/\mathbb{Z}} X_N^{\text{cue}}(t) - \left(\log N - \frac{3}{4}\log\log N\right) \quad \text{converges in distribution.}$$

$$\sup_{t\in\mathbb{R}/\mathbb{Z}} X_N^{\text{cue}}(t) = \log N + o_P(\log N) \qquad \text{Arguin–Belius–Bourgade '15}$$

$$= \log N - \frac{3}{4}\log\log N + o_P(\log\log N) \qquad \text{Paquette–Zeitouni '15}$$

Related work of Arguin–Belius–Bourgade–Soundararajan–Radziwiłł'16 on the Riemann ζ function.

All proofs proceed by exposing an underlying branching structure.

Note that, unlike X_N, distribution of X_N^{cue} is invariant under rotations.

Related work: Maximum of the CUE field

Replacing P_N with a Haar unitary U_N, we obtain the CUE field:

$$X_N^{\mathrm{cue}}(t) = \log|\det(I_N - e(t)U_N)|.$$

Conjecture (Fyodorov–Hiary–Keating '12):

$$M_N := \sup_{t\in\mathbb{R}/\mathbb{Z}} X_N^{\mathrm{cue}}(t) - \left(\log N - \frac{3}{4}\log\log N\right) \quad \text{converges in distribution.}$$

$$\begin{aligned}
\sup_{t\in\mathbb{R}/\mathbb{Z}} X_N^{\mathrm{cue}}(t) &= \log N + o_P(\log N) && \text{Arguin–Belius–Bourgade '15} \\
&= \log N - \frac{3}{4}\log\log N + o_P(\log\log N) && \text{Paquette–Zeitouni '15} \\
&= \log N - \frac{3}{4}\log\log N + O_P(1) && \text{Chhaibi–Madaule–Najnudel '16}
\end{aligned}$$

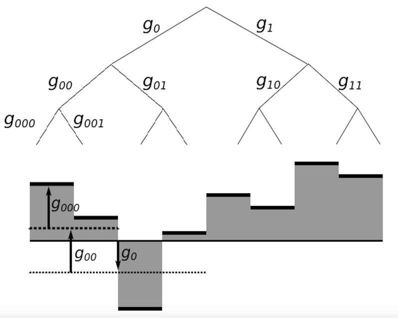
Related work of Arguin–Belius–Bourgade–Soundararajan–Radziwiłł'16 on the Riemann ζ function.

All proofs proceed by exposing an underlying branching structure.

Note that, unlike X_N, distribution of X_N^{cue} is invariant under rotations.

Maximal displacement for branching random walk (BRW)

- X_N and X_N^{cue} are *logarithmically correlated fields*
- Archetypical log-correlated field is (binary, Gaussian) BRW, which we view as a random field $X_n^{\text{brw}}(t)$, $t \in [0,1]$.
- For each $t \in [0,1]$, $X_n^{\text{brw}}(t) \sim N(0,n)$.



Decorrelation of increments: Denoting by $\text{anc}(s,t) \in [1,n]$ the generation where lineages of s,t split, we have

$$\text{Cov}\big(X_n(s) - X_m(s), X_n(t) - X_m(t)\big) = 0 \quad \text{for any } \text{anc}(s,t) < m < n.$$

Maximal displacement for branching random walk (BRW)

Theorem (Hammersley '74, Kingman '75, Biggins '77)

$$\frac{1}{n}\sup_{t\in[0,1)} X_n^{\mathrm{brw}}(t) \to b_c = \sqrt{2\log 2} \quad \text{in probability.}$$

- We recall some key proof ideas going back to Bramson '78.
- Let $T_n = \{k2^{-n} \in [0,1)\}$ and put $\mathcal{S}_n(b) = \{t \in T_n : X_n^{\mathrm{brw}}(t) \geq bn\}$.
- Upper bound. First moment method (union bound):

$$\mathbb{E}\,|\mathcal{S}_n(b_c+\epsilon)| = 2^n\,\mathbb{P}(X_n^{\mathrm{brw}}(1) \geq (b_c+\epsilon)n) \leq e^{-c(\epsilon)n}.$$

  Then apply Markov's inequality.
- Lower bound. Same computation shows $\mathbb{E}\,|\mathcal{S}_n(b_c-\epsilon)| \geq e^{c'(\epsilon)n}$, so we'd be done if we can establish concentration, i.e.

$$\frac{\mathrm{Var}\,|\mathcal{S}_n(b_c-\epsilon)|}{(\mathbb{E}\,|\mathcal{S}_n(b_c-\epsilon)|)^2} = o(1).$$

  But this is false. (See the board...)

Maximal displacement for branching random walk (BRW)

- Modified second moment: Rather than counting high points $\mathcal{S}_n(b)$, we should only count points that make steady progress toward $b_c n$.
- Toy computation: Let $X_n^{(1)} = X_{n/2}^{\text{brw}}$, $X_n^{(2)} = X_n^{\text{brw}} - X_{n/2}^{\text{brw}}$, and put

$$\mathcal{E}_2(t) = \left\{ X_n^{(1)}(t),\, X_n^{(2)}(t) \geq \frac{bn}{2} \right\}.$$

For pairs (s,t) with $\text{anc}(s,t) < n/2$, we can bound

$$\begin{aligned} \mathbb{P}(\mathcal{E}_2(s) \cap \mathcal{E}_2(t)) &= \mathbb{P}\left(X_{n/2}^{(1)}(s), X_{n/2}^{(1)}(t) \geq \frac{bn}{2} \right) \mathbb{P}\left(X_n^{(2)}(s), X_n^{(2)}(t) \geq \frac{bn}{2} \right) \\ &\leq \mathbb{P}\left(X_{n/2}^{(1)}(s) \geq \frac{bn}{2} \right) \mathbb{P}\left(X_n^{(2)}(s) \geq \frac{bn}{2} \right)^2 \end{aligned}$$

where we used the decorrelation of increments.

- Proceeding in a similar manner, we can show the greatest contribution to the second moment of the number of "steadily advancing" particles comes from pairs (s,t) whose lineages split early on.

First steps: Discretize and pass to Poisson model

$X_N(t) = \log|\det(I_N - e(t)P_N)|$, $t \in \mathbb{R}/\mathbb{Z}$.

- It is enough to show

$$\max_{t \in T_N} X_N(t) = (1 + o(1)) \log N \quad \text{w.h.p.}$$

  where T_N is a mesh for $\mathbb{R}/\mathbb{Z}$ of size $|T_N| = O(N)$.

- X_N only depends on the cycle structure of P_N:

$$X_N(t) = \sum_{1 \le \ell \le N} C_\ell(P_N) \log|1 - e(\ell t)|,$$

  where $C_\ell(P_N)$ is the number of cycles of length ℓ in P_N.

- **Arratia–Tavare '92**: Let Z_ℓ be independent $\mathrm{Poi}(1/\ell)$ variables. Then

$$d_{\mathrm{TV}}\big((C_\ell(P_N))_{\ell \le L}, (Z_\ell)_{\ell \le L}\big) \to 0 \quad \text{if } L = o(N).$$

First steps: Discretize and pass to Poisson model

- Let $\omega(1) \leq W \leq N^{o(1)}$ be a slowly growing function of N and split

$$X_N(t) = \sum_{\ell \leq N/W} C_\ell(P_N) \log|1 - e(\ell t)| + \sum_{N/W < \ell \leq N} C_\ell(P_N) \log|1 - e(\ell t)|$$
$$=: X_N^{\leq}(t) + X_N^{>}(t).$$

  Letting

$$Y_N(t) = \sum_{\ell \leq N} Z_\ell \log|1 - e(\ell t)|,$$

  we have $X_N^{\leq}(t) \approx Y_{N/W}(t)$ in distribution.

- Second moment computation shows

$$X_N^{>}(t) \leq \sum_{N/W < \ell \leq N} C_\ell(P_N) \leq O(\log W) = o(\log N) \quad \text{w.h.p.}$$

  so for the upper bound it suffices to consider the Poisson field Y_N.

Upper bound: first attempt

- Let $\mathcal{S}(Y_N, x) = \{t \in T_N : Y_N(t) \geq x \log N\}$. We want to show that for any fixed $\epsilon > 0$,

$$|\mathcal{S}(Y_N, x_c + \epsilon)| = 0 \quad \text{w.h.p.}$$

- First moment:

$$\mathbb{E}\,|\mathcal{S}(Y_N, x)| = \sum_{t \in T_N} \mathbb{P}(Y_N(t) \geq x \log N).$$

- To estimate tail events, estimate MGFs:

$$\mathbb{E}\, e^{\beta Y_N(t)} = \prod_{\ell \leq N} \mathbb{E} \exp(\beta Z_\ell \log |1 - e(\ell t)|) = \exp\left(\sum_{\ell \leq N} \frac{1}{\ell} \Big(|1 - e(\ell t)|^\beta - 1 \Big) \right).$$

- For t irrational, the sum is $\sim (\log N)\big(\int_0^1 |1 - e(u)|^\beta du - 1 \big)$.
- This eventually shows $\mathbb{E}\,|\mathcal{S}(Y_N, 1 + \epsilon)| = o(1)$.

Fix: Condition on a "typical" distribution of cycle lengths

- Problem: First moment is thrown off by "clumping" events, where there is an atypically large number of cycles of comparable length (the corresponding waves $t \mapsto \log|1 - e(\ell t)|$ add constructively).
- Solution: Consider a sequence of windows $I_k = [e^{\delta k}, e^{\delta(k+1)}) \subset [N]$ with δ small (actually $o(1)$) and put

$$\mathcal{N}(I_k) = \sum_{\ell \in I_k} Z_\ell, \qquad Y_{I_k}(t) = \sum_{\ell \in I_k} Z_\ell \log|1 - e(\ell t)|.$$

  Expect $\mathbb{E}\mathcal{N}(I_k) \approx \delta$ cycles in each window. We condition on a "typical" sequence $(\mathcal{N}(I_k))_{k \geq 1}$ (with $\mathcal{N}(I_k) \in \{0, 1\}$ for most k).
- Let $\mathcal{K} = \{k : \mathcal{N}(I_k) = 1\}$. Then $|\mathcal{K}| \approx \log N$, and for $k \in \mathcal{K}$, the increments

$$Y_{I_k}(t) = \log|1 - e(\boldsymbol{\ell}_k t)|, \qquad \mathbb{P}(\boldsymbol{\ell}_k = \ell) \propto \frac{1_{\ell \in I_k}}{\ell}$$

  are still independent and have comparable contribution to the fluctuations of Y_N.

Upper bound

- Under the conditioning on $(\mathcal{N}(I_k))_k$ we have

$$\mathbb{E}\, e^{\beta Y_N(t)} \approx \prod_{k\in\mathcal{K}} \mathbb{E}\exp\left(\beta Y_{I_k}(t)\right) = \prod_{k\in\mathcal{K}}\left(\frac{1}{\sum_{\ell\in I_k}\frac{1}{\ell}}\sum_{\ell\in I_k}\frac{|1-e(\ell t)|^\beta}{\ell}\right).$$

- If t is irrational then

$$\frac{1}{\sum_{\ell\in I_k}\frac{1}{\ell}}\sum_{\ell\in I_k}\frac{|1-e(\ell t)|^\beta}{\ell} = \int_0^1 |1-e(u)|^\beta du + o_{\beta,t}(1).$$

- As we'll be taking a union bound over $t\in T_N$, we need to quantify dependence of the error on t. We get satisfactory errors outside low-frequency Bohr sets

$$B_\xi(\eta) = \{t\in\mathbb{R}/\mathbb{Z} : \|\xi t\|_{\mathbb{R}/\mathbb{Z}} \le \eta\}.$$

We call $t\in \mathrm{Maj}(\xi_0,\eta) := \bigcup_{\xi\le\xi_0} B_\xi(\eta)$ major arc points.

Upper bound: Major and minor arcs

- While $Y_N(t)$ is badly behaved on major arcs, we can show

$$\sup_{t \in \mathrm{Maj}(\xi_0,\eta)} Y_N(t) \le -c(\xi_0,\eta)\log^2 N \quad \text{w.h.p.}$$

- On the complementary set of minor arcs, Y_N behaves like a sequence of iid variables:

$$Y_{I_k}(t) = \log|1-e(\ell_k t)| \approx \log|1-e(\boldsymbol{u}_k)|, \qquad \boldsymbol{u}_k \sim \mathit{Unif}(\mathbb{R}/\mathbb{Z}).$$

- Letting $\lambda(\beta) = \log \mathbb{E}\exp\left(\beta \log|1-e(\boldsymbol{u})|\right) = \log \int_0^1 |1-e(u)|^\beta du$, for major arcs we have a first moment estimate for $|\mathcal{S}(Y_N, x) \setminus \mathrm{Maj}\,|$:

$$\sum_{t \in T_N \setminus \mathrm{Maj}(\xi_0,\eta)} \mathbb{P}\left(Y_N(t) \ge x \log N\right) \approx |T_N| \exp\left(-\lambda^*(x)|\mathcal{K}|\right) \approx N^{1-\lambda^*(x)}.$$

  where λ^* is the Legendre transform of λ.
- $x_c \approx 0.677$ is the unique solution to $\lambda^*(x) = 1$.

Lower bound: Early, middle, and late generations

For $M \le N$, denote the truncated field

$$X_{\le M}(t) = \sum_{\ell \le M} C_\ell(P_N) \log |1 - e(\ell t)|,$$

and the set of "survivors"

$$\mathcal{S}_{\le M}(x) = \{t \in T_N : X_{\le M}(t) \ge x \log M\}.$$

We show $\mathcal{S}_{\le N}(x_c - \epsilon)$ is nonempty by tracking the population of survivors across three epochs:

1. **Early generations:** $|\mathcal{S}_{\le N^c}(-C)| \gtrsim N$ w.h.p.
2. **Middle generations:** $|\mathcal{S}_{\le N/W}(x_c - \epsilon)| \ge N^{c(\epsilon)}$ w.h.p.
3. **Late generations:** $|\mathcal{S}_{\le N}(x_c - 2\epsilon)| \ge 1$ w.h.p.

For early and middle generations we can replace $X_{\le M}$ with Y_M.

Lower bound: Middle generations

- We follow the second moment argument for BRW. Requires approximate decorrelation of tail events for increments $\Delta_{m,n}(t) := \sum_{k\in\mathcal{K}\cap[m,n)} Y_{I_k}(t)$.
- For $\xi_0 \in \mathbb{N}$ put

$$d_{\xi_0}(s,t) = \min_{\xi,\xi'\in\{-\xi_0,\dots,\xi_0\}\setminus\{0\}} \|\xi s + \xi' t\|_{\mathbb{R}/\mathbb{Z}}.$$

  Lower bound on $d_{\xi_0}(s,t)$ gives quantitative linear indepedence of $1, s, t$ over $\mathbb{Z}$.
- If $m = \omega(1)$, $s, t \notin \mathrm{Maj}(\xi_0, \eta)$ and $d_{\xi_0}(s,t) > \eta$ for $\xi_0 = \omega(n^2)$ and $\eta \gg e^{-\delta m}$, then we have an estimate of the form

$$\mathbb{P}\left(\Delta_{m,n}(s), \Delta_{m,n}(t) \geq x\right) \leq (1+o(1))\,\mathbb{P}(\Delta_{m,n}(s) \geq x)^2.$$

Lower bound: Late generations ($|\mathcal{S}_{\le N}(x_c - 2\epsilon)| \ge 1$)

- Poisson model is no longer available. We condition on the $N^{c(\epsilon)}$ survivors of Middle Ages $\mathcal{S}_{\le N/W} := \mathcal{S}_{\le N/W}(x_c - \epsilon)$ and want to show they aren't all wiped out by the high frequency tail $X_N^{>} = X_N - X_{\le N/W}$.
- Note that $\log|1 - e(\ell t)| \le -\epsilon \log N$ for t in the Bohr set

$$B_\ell(N^{-\epsilon}) = \{t \in \mathbb{R}/\mathbb{Z} : \|\ell t\|_{\mathbb{R}/\mathbb{Z}} \le N^{-\epsilon}\}.$$

  So if $\mathcal{S}_{\le N/W} \subset B_\ell(N^{-\varepsilon})$ for some $\ell \in (N/W, N]$, there is a chance the whole population gets wiped out.
- We employ a structural dichotomy for $\mathcal{S}_{\le N/W}$: Either
  - (Structured case) $\mathcal{S}_{\le N/W}$ has large overlap with $B_\ell(N^{-\epsilon})$ for $\ge N/W^3$ different frequencies $\ell \in (N/W, N]$, or
  - (Unstructured case) it doesn't.

  In the unstructured case, we can show it is unlikely any of the exceptional frequencies are selected (using the switchings method).

Lower bound: Late generations

(Structured case)

$\mathcal{S}_{\leq N/W}$ has large overlap with $B_\ell(N^{-\epsilon})$ for $\geq N/W^3$ different frequencies $\ell \in (N/W, N]$.

In this case we can use pigeonholing and a Vinogradov-type lemma to show the $\mathcal{S}_{\leq N/W}$ must actually contain an element that is major arc, specifically an element of $B_\xi(\eta)$ for some

$$\xi = W^{O(1)}, \qquad \eta = W^{O(1)} N^{-\epsilon-1}.$$

But such major arc points are unlikely to be in $\mathcal{S}_{\leq N/W}$ in the first place.