Large deviations of subgraph counts for sparse random graphs

Banff, Canada, August 7, 2019

Amir Dembo, Stanford University

Joint works with Nicholas Cook and with Sohom Bhattacharya

Celebrating HT Yau's 60-th birthday!

Universality for typical behavior: Examples

- **CLT:** for $X_1, X_2, \ldots$ iid, $\mathbb{E}\, X_1 = 0$, $\mathbb{E}\, X_1^2 = 1$,

$$\forall\ a < b, \quad \mathbb{P}\left\{ \tfrac{X_1 + \cdots + X_N}{\sqrt{N}} \in [a, b] \right\} \longrightarrow \gamma([a, b]) \qquad (\textit{universal}).$$

Universality for typical behavior: Examples

- **CLT:** for $X_1, X_2, \dots$ iid, $\mathbb{E}\, X_1 = 0$, $\mathbb{E}\, X_1^2 = 1$,

$$\forall\ a < b, \quad \mathbb{P}\left\{ \frac{X_1 + \cdots + X_N}{\sqrt{N}} \in [a,b] \right\} \longrightarrow \gamma([a,b]) \qquad (\textit{universal}).$$

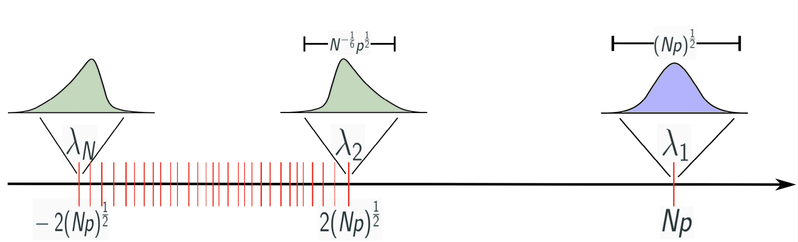
- Let $\boldsymbol{A} = (a_{ij})_{i,j=1}^N$ adjacency matrix for the Erdős–Rényi graph $G(N,p)$ with $0 < p \ll 1$, eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$.

Universality for typical behavior: Examples

- **CLT:** for $X_1, X_2, \dots$ iid, $\mathbb{E}\, X_1 = 0$, $\mathbb{E}\, X_1^2 = 1$,

$$\forall\ a < b, \quad \mathbb{P}\left\{ \frac{X_1 + \cdots + X_N}{\sqrt{N}} \in [a,b] \right\} \longrightarrow \gamma([a,b]) \qquad (\textit{universal}).$$

- Let $\boldsymbol{A} = (a_{ij})_{i,j=1}^N$ adjacency matrix for the Erdős–Rényi graph $G(N,p)$ with $0 < p \ll 1$, eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$.



λ_1 is asymptotically Gaussian.

For $p \gg N^{-2/3}$: $\lambda_2, -\lambda_N$ follow the Tracy–Widom law [Lee–Schnelli '16].

Large deviations: Beyond universality

CLT: for $X_1, X_2, \ldots$ iid, $\mathbb{E}\, X_1 = 0$, $\mathbb{E}\, X_1^2 = 1$,

$$\forall\, a < b, \quad \mathbb{P}\left\{ \frac{X_1 + \cdots + X_N}{\sqrt{N}} \in [a, b] \right\} \longrightarrow \gamma([a, b]) \qquad \textit{(universal)}.$$

Large deviations: Beyond universality

CLT: for $X_1, X_2, \dots$ iid, $\mathbb{E}\, X_1 = 0$, $\mathbb{E}\, X_1^2 = 1$,

$$\forall\ a < b, \quad \mathbb{P}\left\{\frac{X_1 + \cdots + X_N}{\sqrt{N}} \in [a, b]\right\} \longrightarrow \gamma([a, b]) \qquad (\textit{universal}).$$

Compare Cramér's **Large deviations principle (LDP)**:

$$\forall\ a < b, \quad \frac{1}{N} \log \mathbb{P}\left\{\frac{X_1 + \cdots + X_N}{N} \in [a, b]\right\} \longrightarrow -\inf_{x \in [a,b]} J(x),$$

where J is the non-universal *rate function* depending strongly on the law of X_1 (particularly its tail behavior).

Large deviations: Beyond universality

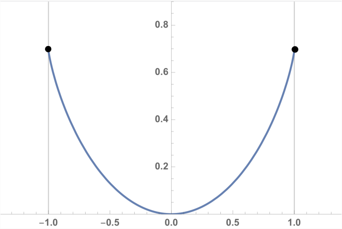
CLT: for $X_1, X_2, \ldots$ iid, $\mathbb{E}\, X_1 = 0$, $\mathbb{E}\, X_1^2 = 1$,

$$\forall\ a < b, \quad \mathbb{P}\left\{ \frac{X_1 + \cdots + X_N}{\sqrt{N}} \in [a,b] \right\} \longrightarrow \gamma([a,b]) \qquad (\textit{universal}).$$

Compare Cramér's **Large deviations principle (LDP)**:

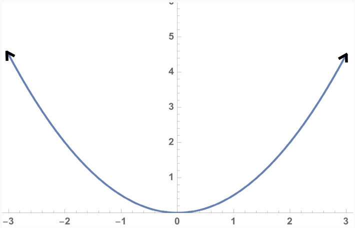
$$\forall\ a < b, \quad \frac{1}{N} \log \mathbb{P}\left\{ \frac{X_1 + \cdots + X_N}{N} \in [a,b] \right\} \longrightarrow -\inf_{x \in [a,b]} J(x),$$

where J is the non-universal *rate function* depending strongly on the law of X_1 (particularly its tail behavior).



Rademacher

$\frac{1+x}{2} \log(1+x) + \frac{1-x}{2} \log(1-x)$



Gaussian

$\frac{x^2}{2}$

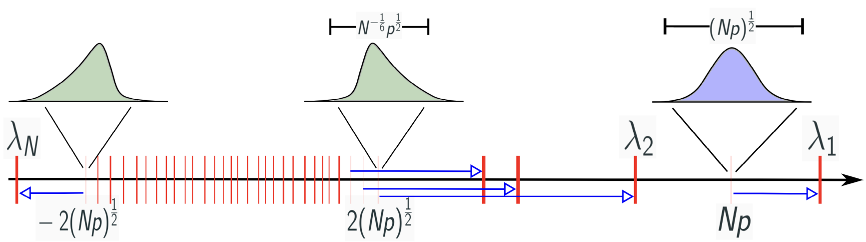
Nonlinear large deviations (Chatterjee–D. '14)

How about nonlinear functionals?

Nonlinear large deviations (Chatterjee–D. '14)

How about nonlinear functionals?

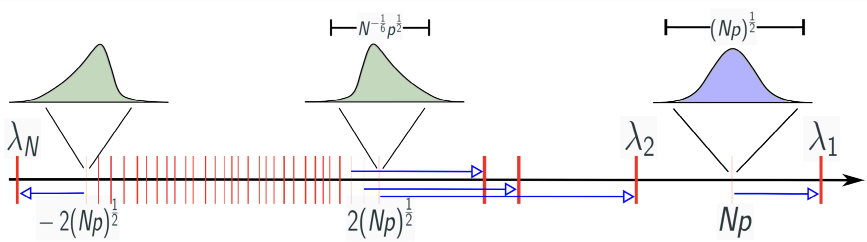
Example: Extreme eigenvalues of random matrices / random graphs.



Nonlinear large deviations (Chatterjee–D. '14)

How about nonlinear functionals?

Example: Extreme eigenvalues of random matrices / random graphs.

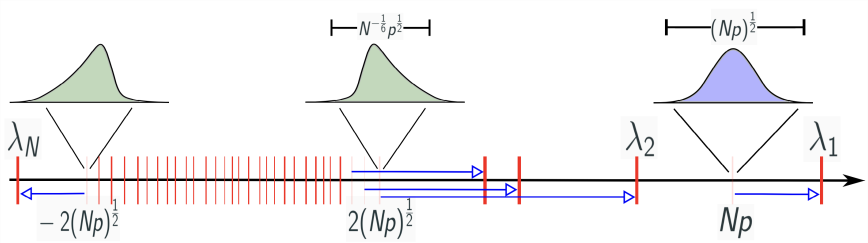


In this talk we focus on low-degree polynomials of Bernoulli variables.
(Tails for eigenvalues will be under the hood.)

Nonlinear large deviations (Chatterjee–D. '14)

How about nonlinear functionals?

Example: Extreme eigenvalues of random matrices / random graphs.



In this talk we focus on low-degree polynomials of Bernoulli variables. (Tails for eigenvalues will be under the hood.)

Note we consider outliers at scale Np (for LDP at scale of the bulk cf. Guionnet–Husson 17' for $p = 1/2$).

Subgraph counts in $G(N, p)$

- Let $\boldsymbol{G} \sim G(N, p)$ be an Erdős–Rényi graph on vertices $[N] = \{1, \ldots, N\}$

Subgraph counts in $G(N, p)$

- Let $\boldsymbol{G} \sim G(N, p)$ be an Erdős–Rényi graph on vertices $[N] = \{1, \ldots, N\}$
- Number of triangles in $\boldsymbol{G}$: $\quad \mathcal{N}_\Delta(\boldsymbol{G}) = \sum_{\{i,j,k\} \subset [N]} a_{ij} a_{jk} a_{ik}$

  (recall the adjacency matrix $\boldsymbol{A} = (a_{ij})_{i,j=1}^N$ with $a_{ij} = \mathbf{1}_{\{i,j\} \text{ is an edge}}$).

Subgraph counts in $G(N, p)$

- Let $\boldsymbol{G} \sim G(N, p)$ be an Erdős–Rényi graph on vertices $[N] = \{1, \ldots, N\}$
- Number of triangles in $\boldsymbol{G}$: $\quad \mathcal{N}_\Delta(\boldsymbol{G}) = \sum_{\{i,j,k\} \subset [N]} a_{ij} a_{jk} a_{ik}$

  (recall the adjacency matrix $\boldsymbol{A} = (a_{ij})_{i,j=1}^N$ with $a_{ij} = \mathbf{1}_{\{i,j\} \text{ is an edge}}$).

  $\mathbb{E}\,\mathcal{N}_\Delta(\boldsymbol{G}) = \binom{N}{3} p^3$.

Subgraph counts in $G(N,p)$

- Let $\boldsymbol{G} \sim G(N,p)$ be an Erdős–Rényi graph on vertices $[N] = \{1,\dots,N\}$
- Number of triangles in $\boldsymbol{G}$: $\mathcal{N}_\Delta(\boldsymbol{G}) = \sum_{\{i,j,k\}\subset[N]} a_{ij}a_{jk}a_{ik}$

  (recall the adjacency matrix $\boldsymbol{A} = (a_{ij})_{i,j=1}^N$ with $a_{ij} = \mathbf{1}_{\{i,j\} \text{ is an edge}}$).

  $\mathbb{E}\,\mathcal{N}_\Delta(\boldsymbol{G}) = \binom{N}{3}p^3$.
- **Question:** Conditional on $\boldsymbol{G}$ having extra triangles, i.e. $\left\{\mathcal{N}_\Delta(\boldsymbol{G}) \geq \binom{N}{3}q^3\right\}$ for some $q > p$, how are the edges distributed?

Subgraph counts in $G(N, p)$

- Let $\boldsymbol{G} \sim G(N, p)$ be an Erdős–Rényi graph on vertices $[N] = \{1, \ldots, N\}$
- Number of triangles in $\boldsymbol{G}$: $\mathcal{N}_\Delta(\boldsymbol{G}) = \sum_{\{i,j,k\} \subset [N]} a_{ij} a_{jk} a_{ik}$

  (recall the adjacency matrix $\boldsymbol{A} = (a_{ij})_{i,j=1}^N$ with $a_{ij} = \mathbf{1}_{\{i,j\} \text{ is an edge}}$).

  $\mathbb{E}\,\mathcal{N}_\Delta(\boldsymbol{G}) = \binom{N}{3} p^3$.
- **Question:** Conditional on $\boldsymbol{G}$ having extra triangles, i.e. $\left\{\mathcal{N}_\Delta(\boldsymbol{G}) \geq \binom{N}{3} q^3\right\}$ for some $q > p$, how are the edges distributed?
  A few possibilities:

Subgraph counts in $G(N, p)$

- Let $\boldsymbol{G} \sim G(N, p)$ be an Erdős–Rényi graph on vertices $[N] = \{1, \ldots, N\}$
- Number of triangles in $\boldsymbol{G}$: $\quad \mathcal{N}_\Delta(\boldsymbol{G}) = \sum_{\{i,j,k\} \subset [N]} a_{ij} a_{jk} a_{ik}$

  (recall the adjacency matrix $\boldsymbol{A} = (a_{ij})_{i,j=1}^N$ with $a_{ij} = \mathbf{1}_{\{i,j\} \text{ is an edge}}$).

  $\mathbb{E} \mathcal{N}_\Delta(\boldsymbol{G}) = \binom{N}{3} p^3$.
- **Question:** Conditional on $\boldsymbol{G}$ having extra triangles, i.e. $\left\{ \mathcal{N}_\Delta(\boldsymbol{G}) \geq \binom{N}{3} q^3 \right\}$ for some $q > p$, how are the edges distributed?
  A few possibilities:

  (A) As in $G(N, q)$?

Subgraph counts in $G(N,p)$

- Let $\boldsymbol{G} \sim G(N,p)$ be an Erdős–Rényi graph on vertices $[N] = \{1, \dots, N\}$
- Number of triangles in $\boldsymbol{G}$: $\mathcal{N}_\Delta(\boldsymbol{G}) = \sum_{\{i,j,k\} \subset [N]} a_{ij} a_{jk} a_{ik}$

  (recall the adjacency matrix $\boldsymbol{A} = (a_{ij})_{i,j=1}^N$ with $a_{ij} = \mathbf{1}_{\{i,j\} \text{ is an edge}}$).

  $\mathbb{E}\mathcal{N}_\Delta(\boldsymbol{G}) = \binom{N}{3} p^3$.

- **Question:** Conditional on $\boldsymbol{G}$ having extra triangles, i.e. $\left\{\mathcal{N}_\Delta(\boldsymbol{G}) \geq \binom{N}{3} q^3\right\}$ for some $q > p$, how are the edges distributed?
  A few possibilities:

  (A) As in $G(N,q)$?

  (B) As in $G(N,p)$ with a small planted clique?

Subgraph counts in $G(N,p)$

- Let $\boldsymbol{G} \sim G(N,p)$ be an Erdős–Rényi graph on vertices $[N] = \{1, \ldots, N\}$
- Number of triangles in $\boldsymbol{G}$: $\mathcal{N}_\Delta(\boldsymbol{G}) = \sum_{\{i,j,k\} \subset [N]} a_{ij} a_{jk} a_{ik}$

  (recall the adjacency matrix $\boldsymbol{A} = (a_{ij})_{i,j=1}^N$ with $a_{ij} = \mathbf{1}_{\{i,j\} \text{ is an edge}}$).

  $\mathbb{E} \mathcal{N}_\Delta(\boldsymbol{G}) = \binom{N}{3} p^3$.
- **Question:** Conditional on $\boldsymbol{G}$ having extra triangles, i.e. $\left\{ \mathcal{N}_\Delta(\boldsymbol{G}) \geq \binom{N}{3} q^3 \right\}$ for some $q > p$, how are the edges distributed?
  A few possibilities:

  (A) As in $G(N,q)$?

  (B) As in $G(N,p)$ with a small planted clique?

  (C) As in $G(N,p)$ with a small planted hub?

Subgraph counts in $G(N,p)$

- Let $\boldsymbol{G} \sim G(N,p)$ be an Erdős–Rényi graph on vertices $[N] = \{1,\dots,N\}$
- Number of triangles in $\boldsymbol{G}$: $\mathcal{N}_\Delta(\boldsymbol{G}) = \sum_{\{i,j,k\}\subset[N]} a_{ij}a_{jk}a_{ik}$

  (recall the adjacency matrix $\boldsymbol{A} = (a_{ij})_{i,j=1}^N$ with $a_{ij} = \mathbf{1}_{\{i,j\} \text{ is an edge}}$).

  $\mathbb{E}\mathcal{N}_\Delta(\boldsymbol{G}) = \binom{N}{3}p^3$.

- **Question:** Conditional on $\boldsymbol{G}$ having extra triangles, i.e. $\left\{\mathcal{N}_\Delta(\boldsymbol{G}) \ge \binom{N}{3}q^3\right\}$ for some $q > p$, how are the edges distributed? A few possibilities:

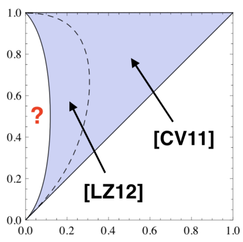
  (A) As in $G(N,q)$?

  (B) As in $G(N,p)$ with a small planted clique?

  (C) As in $G(N,p)$ with a small planted hub?

  Answer is (A) for much (not all!) of $0 < p < q < 1$ fixed.

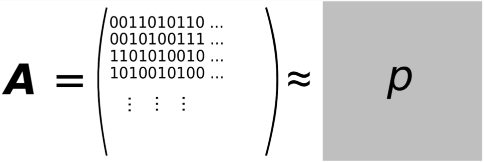
  [Chatterjee–Varadhan '11]+[Lubetzky–Zhao '12].



Subgraph counts in $G(N, p)$

Conjecture: Let H have max degree D. For $N^{-1/D} \ll p \ll 1$, depending on the size of δ,

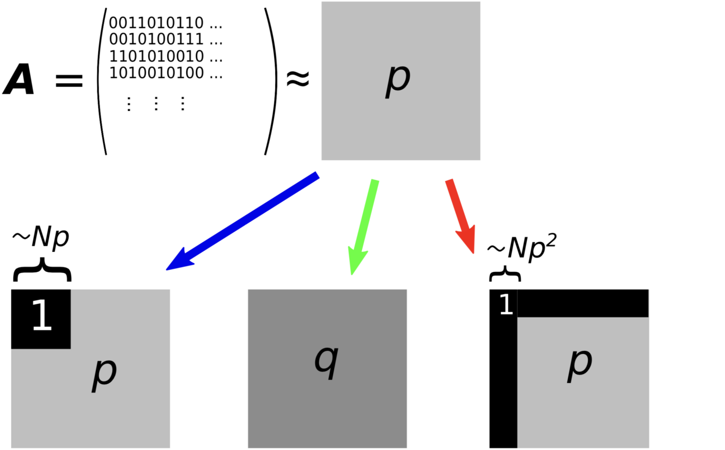
$$G \,\Big|\, \Big\{ \mathcal{N}_H(G) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_H(G) \Big\} \quad \approx \quad G(N,p) + \text{planted clique or hub.}$$



Subgraph counts in $G(N, p)$

Conjecture: Let H have max degree D. For $N^{-1/D} \ll p \ll 1$, depending on the size of δ,

$$\boldsymbol{G} \,\Big|\, \Big\{ \mathcal{N}_H(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_H(\boldsymbol{G}) \Big\} \quad \approx \quad G(N, p) + \text{planted clique or hub.}$$



The "infamous" upper tail for triangle counts [Janson–Ruciński '02]

- Upper tail up to constant factors in the exponent:

$$\mathbb{P}\{\mathcal{N}_\Delta(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_\Delta(\boldsymbol{G})\} = p^{\Theta_\delta(N^2 p^2)}, \quad p \geq (\log N)/N.$$

  [Chatterjee '12], [DeMarco–Kahn '12]

The "infamous" upper tail for triangle counts [Janson–Ruciński '02]

- Upper tail up to constant factors in the exponent:

$$\mathbb{P}\{\mathcal{N}_\Delta(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_\Delta(\boldsymbol{G})\} = p^{\Theta_\delta(N^2p^2)}, \quad p \geq (\log N)/N.$$

[Chatterjee '12], [DeMarco–Kahn '12]

- Recent works find the leading exponential order:

$$\mathbb{P}\{\mathcal{N}_\Delta(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_\Delta(\boldsymbol{G})\} = p^{(1+o(1))\min\{\frac{\delta^{2/3}}{2},\frac{\delta}{3}\}N^2p^2}$$

(matching probabilities for planted clique or hub of appropriate size) for $N^{-\kappa} \ll p \ll 1$, with

The “infamous” upper tail for triangle counts [Janson–Ruciński ’02]

- Upper tail up to constant factors in the exponent:

$$\mathbb{P}\{\mathcal{N}_\Delta(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_\Delta(\boldsymbol{G})\} = p^{\Theta_\delta(N^2p^2)}, \quad p \geq (\log N)/N.$$

  [Chatterjee ’12], [DeMarco–Kahn ’12]

- Recent works find the leading exponential order:

$$\mathbb{P}\{\mathcal{N}_\Delta(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_\Delta(\boldsymbol{G})\} = p^{(1+o(1))\min\{\frac{\delta^{2/3}}{2},\frac{\delta}{3}\}N^2p^2}$$

  (matching probabilities for planted clique or hub of appropriate size) for $N^{-\kappa} \ll p \ll 1$, with

  * $\kappa = \frac{1}{41} - \epsilon$ [Chatterjee–D. ’14] + [Lubetzky–Zhao ’14]
  * $\kappa = \frac{1}{18} - \epsilon$ [Eldan ’16]
  * $\kappa = \frac{1}{3}$ [Cook–D. ’18] (or $\frac{1}{2} - \epsilon$ for cycles of $\ell \geq 4$).
  * $\kappa = \frac{1}{2} - \epsilon$ [Augeri ’18] for cycles of length $\ell \geq 3$.
  * $\kappa = 1 - \epsilon$ [Harel–Mousset–Samotij ’19].

The "infamous" upper tail for triangle counts [Janson–Ruciński '02]

- Upper tail up to constant factors in the exponent:

$$\mathbb{P}\{\mathcal{N}_\Delta(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_\Delta(\boldsymbol{G})\} = p^{\Theta_\delta(N^2p^2)}, \quad p \geq (\log N)/N.$$

  [Chatterjee '12], [DeMarco–Kahn '12]

- Recent works find the leading exponential order:

$$\mathbb{P}\{\mathcal{N}_\Delta(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_\Delta(\boldsymbol{G})\} = p^{(1+o(1))\min\{\frac{\delta^{2/3}}{2},\frac{\delta}{3}\}N^2p^2}$$

  (matching probabilities for planted clique or hub of appropriate size) for $N^{-\kappa} \ll p \ll 1$, with

  * $\kappa = \frac{1}{41} - \epsilon$ [Chatterjee–D. '14] + [Lubetzky–Zhao '14]
  * $\kappa = \frac{1}{18} - \epsilon$ [Eldan '16]
  * $\kappa = \frac{1}{3}$ [Cook–D. '18] (or $\frac{1}{2} - \epsilon$ for cycles of $\ell \geq 4$).
  * $\kappa = \frac{1}{2} - \epsilon$ [Augeri '18] for cycles of length $\ell \geq 3$.
  * $\kappa = 1 - \epsilon$ [Harel–Mousset–Samotij '19].

- How about general subgraphs?

Upper tail for general subgraph counts

- Let $H = (V, E)$ connected of max degree D, and assume $N^{-\kappa(H)} \ll p \ll 1$ for some $\kappa(H) \in (0, 1)$.

Upper tail for general subgraph counts

- Let $H = (V, E)$ connected of max degree D, and assume $N^{-\kappa(H)} \ll p \ll 1$ for some $\kappa(H) \in (0, 1)$.
- [Chatterjee–D. '14] + [Bhattacharya–Ganguly–Lubetzky–Zhao '16]:

$$\mathbb{P}\left\{\mathcal{N}_H(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_H(\boldsymbol{G})\right\} = p^{(1+o(1))c_H(\delta)N^2p^D}$$

matching the probability of a planted clique or hub up to sub-exponential factors, assuming $\kappa(H) = \frac{c}{D|E|}$.

Upper tail for general subgraph counts

- Let $H = (V, E)$ connected of max degree D, and assume $N^{-\kappa(H)} \ll p \ll 1$ for some $\kappa(H) \in (0, 1)$.
- [Chatterjee–D. '14] + [Bhattacharya–Ganguly–Lubetzky–Zhao '16]:

$$\mathbb{P}\left\{\mathcal{N}_H(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_H(\boldsymbol{G})\right\} = p^{(1+o(1))c_H(\delta)N^2 p^D}$$

matching the probability of a planted clique or hub up to sub-exponential factors, assuming $\kappa(H) = \frac{c}{D|E|}$.

(formula for $c_H(\delta)$ obtained by [BGLZ '16] as solution to the (LDP) variational problem, valid down to $\kappa(H) = 1/D$.)

Upper tail for general subgraph counts

- Let $H = (V, E)$ connected of max degree D, and assume $N^{-\kappa(H)} \ll p \ll 1$ for some $\kappa(H) \in (0, 1)$.
- [Chatterjee–D. '14] + [Bhattacharya–Ganguly–Lubetzky–Zhao '16]:

$$\mathbb{P}\left\{\mathcal{N}_H(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_H(\boldsymbol{G})\right\} = p^{(1+o(1))c_H(\delta)N^2p^D}$$

  matching the probability of a planted clique or hub up to sub-exponential factors, assuming $\kappa(H) = \frac{c}{D|E|}$.

  (formula for $c_H(\delta)$ obtained by [BGLZ '16] as solution to the (LDP) variational problem, valid down to $\kappa(H) = 1/D$.)
- [Eldan '16] + [BGLZ '16]: can take $\kappa(H) = \frac{1}{6|E|} - \epsilon$.

Upper tail for general subgraph counts

- Let $H = (V, E)$ connected of max degree D, and assume $N^{-\kappa(H)} \ll p \ll 1$ for some $\kappa(H) \in (0, 1)$.
- [Chatterjee–D. '14] + [Bhattacharya–Ganguly–Lubetzky–Zhao '16]:

$$\mathbb{P}\left\{\mathcal{N}_H(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_H(\boldsymbol{G})\right\} = p^{(1+o(1))c_H(\delta)N^2 p^D}$$

  matching the probability of a planted clique or hub up to sub-exponential factors, assuming $\kappa(H) = \frac{c}{D|E|}$.

  (formula for $c_H(\delta)$ obtained by [BGLZ '16] as solution to the (LDP) variational problem, valid down to $\kappa(H) = 1/D$.)
- [Eldan '16] + [BGLZ '16]: can take $\kappa(H) = \frac{1}{6|E|} - \epsilon$.
- [Cook–D. '18]: $\kappa(H) = \frac{1}{3D-2} - \epsilon$.

Upper tail for general subgraph counts

Theorem (Cook–D. '18)

Fix $H = (V, E)$ connected of max degree $D \geq 2$. If $N^{-\frac{1}{3D-2}+\epsilon} \leq p \ll 1$ then

$$\mathbb{P}\left\{\mathcal{N}_H(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_H(\boldsymbol{G})\right\} = p^{(1+o(1))c_H(\delta)N^2 p^D}.$$

Upper tail for general subgraph counts

Theorem (Cook–D. '18)

Fix $H = (V, E)$ connected of max degree $D \geq 2$. If $N^{-\frac{1}{3D-2}+\epsilon} \leq p \ll 1$ then

$$\mathbb{P}\left\{\mathcal{N}_H(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_H(\boldsymbol{G})\right\} = p^{(1+o(1))c_H(\delta)N^2p^D}.$$

- This is currently the best result for general H, but see
  - [Cook–D. '18], [Augeri '18] for sharpening in case of cycles (exploiting relationship to the spectrum of $\boldsymbol{A}$);
  - improvement to $\kappa(H) = \frac{2}{D} - \epsilon$ for H non-bipartite D-regular in [Harel–Mousset–Samotij '19].

Upper tail for general subgraph counts

Theorem (Cook–D. '18)

Fix $H = (V, E)$ connected of max degree $D \geq 2$. If $N^{-\frac{1}{3D-2}+\epsilon} \leq p \ll 1$ then

$$\mathbb{P}\left\{\mathcal{N}_H(\boldsymbol{G}) \geq (1+\delta)\,\mathbb{E}\,\mathcal{N}_H(\boldsymbol{G})\right\} = p^{(1+o(1))c_H(\delta)N^2 p^D}.$$

- This is currently the best result for general H, but see
  - ∗ [Cook–D. '18], [Augeri '18] for sharpening in case of cycles (exploiting relationship to the spectrum of $\boldsymbol{A}$);
  - ∗ improvement to $\kappa(H) = \frac{2}{D} - \epsilon$ for H non-bipartite D-regular in [Harel–Mousset–Samotij '19].
- We also get:
  - ∗ lower tails (reduction to variational problem – can solve only for Sidorenko graphs);
  - ∗ upper tails for $\lambda_1, \lambda_2, -\lambda_N$ (together with subsequent work by [Bhattacharya–Ganguly '18] solving the LDP variational problem).

Previous approaches to upper tails

- [Chatterjee–D. '14]: large deviations for nonlinear functions $f : \{0,1\}^d \to \mathbb{R}$ through the study of Gibbs measures μ with density $\mu(\{x\}) \propto e^{h(x)}$ for some Hamiltonian $h : \{0,1\}^d \to \mathbb{R}$.

Previous approaches to upper tails

- [Chatterjee–D. '14]: large deviations for nonlinear functions $f : \{0,1\}^d \to \mathbb{R}$ through the study of Gibbs measures μ with density $\mu(\{x\}) \propto e^{h(x)}$ for some Hamiltonian $h : \{0,1\}^d \to \mathbb{R}$.
- Taking $e^{h(x)}$ as a "smooth" approximation to the indicator function $1_{f(x)\geq t}$, recover estimates on $\mathbb{P}(f(X) \geq t)$ from estimates on the partition function $Z = \sum_{x\in\{0,1\}^d} e^{h(x)}$.

Previous approaches to upper tails

- [Chatterjee–D. '14]: large deviations for nonlinear functions $f : \{0,1\}^d \to \mathbb{R}$ through the study of Gibbs measures μ with density $\mu(\{x\}) \propto e^{h(x)}$ for some Hamiltonian $h : \{0,1\}^d \to \mathbb{R}$.
- Taking $e^{h(x)}$ as a "smooth" approximation to the indicator function $1_{f(x)\geq t}$, recover estimates on $\mathbb{P}(f(X) \geq t)$ from estimates on the partition function $Z = \sum_{x\in\{0,1\}^d} e^{h(x)}$.
- Obtain conditions for validity of the naïve mean field approximation:

$$\log Z = \sup_{\nu\in M_1(\{0,1\}^d)} \int h d\nu - H(\nu\|\mu) \approx \sup_{\substack{\nu\in M_1(\{0,1\}^d) \\ \text{product measures}}} \int h d\nu - H(\nu\|\mu)$$

where $H(\nu\|\mu)$ is the relative entropy.

Previous approaches to upper tails

- [Chatterjee–D. '14]: large deviations for nonlinear functions $f : \{0,1\}^d \to \mathbb{R}$ through the study of Gibbs measures μ with density $\mu(\{x\}) \propto e^{h(x)}$ for some Hamiltonian $h : \{0,1\}^d \to \mathbb{R}$.
- Taking $e^{h(x)}$ as a "smooth" approximation to the indicator function $1_{f(x)\geq t}$, recover estimates on $\mathbb{P}(f(X) \geq t)$ from estimates on the partition function $Z = \sum_{x\in\{0,1\}^d} e^{h(x)}$.
- Obtain conditions for validity of the naïve mean field approximation:

$$\log Z = \sup_{\nu \in M_1(\{0,1\}^d)} \int h d\nu - H(\nu\|\mu) \approx \sup_{\substack{\nu \in M_1(\{0,1\}^d) \\ \text{product measures}}} \int h d\nu - H(\nu\|\mu)$$

  where $H(\nu\|\mu)$ is the relative entropy.
- Extended and refined by [Yan '15], [Eldan '16], [Augeri '18], [Austin '18].

Previous approaches to upper tails

- [Chatterjee–D. '14]: large deviations for nonlinear functions $f : \{0,1\}^d \to \mathbb{R}$ through the study of Gibbs measures μ with density $\mu(\{x\}) \propto e^{h(x)}$ for some Hamiltonian $h : \{0,1\}^d \to \mathbb{R}$.
- Taking $e^{h(x)}$ as a "smooth" approximation to the indicator function $1_{f(x)\geq t}$, recover estimates on $\mathbb{P}(f(X) \geq t)$ from estimates on the partition function $Z = \sum_{x\in\{0,1\}^d} e^{h(x)}$.
- Obtain conditions for validity of the naïve mean field approximation:

$$\log Z = \sup_{\nu \in M_1(\{0,1\}^d)} \int h d\nu - H(\nu\|\mu) \approx \sup_{\substack{\nu \in M_1(\{0,1\}^d) \\ \text{product measures}}} \int h d\nu - H(\nu\|\mu)$$

  where $H(\nu\|\mu)$ is the relative entropy.
- Extended and refined by [Yan '15], [Eldan '16], [Augeri '18], [Austin '18].
- Disadvantage: Errors in the passage from indicator functions to smooth approximations cause a sub-optimal range of sparsity.

Dense case (Chatterjee–Varadhan '11)

- For a sequence of probability measures μ_N on a common topological space $\mathcal{X}$, large deviations principle (LDP) yields asymptotics of form

$$\mu_N(\mathcal{E}) \approx \exp\big(-v_N \inf_{x\in\mathcal{E}} J(x)\big), \qquad \mathcal{E} \subseteq \mathcal{X},$$

  for a rate function J and speed v_N.

Dense case (Chatterjee–Varadhan '11)

- For a sequence of probability measures μ_N on a common topological space $\mathcal{X}$, large deviations principle (LDP) yields asymptotics of form

$$\mu_N(\mathcal{E}) \approx \exp\big(-v_N \inf_{x\in\mathcal{E}} J(x)\big), \qquad \mathcal{E} \subseteq \mathcal{X},$$

for a rate function J and speed v_N.

- In dense case (p fixed), C–V get an LDP for $\mu_N(\cdot) = \mathbb{P}(\boldsymbol{G} \in \cdot)$. What does it mean? μ_N live on separate spaces $\mathcal{G}_N \cong \{0,1\}^{\binom{N}{2}}$...

Dense case (Chatterjee–Varadhan '11)

- For a sequence of probability measures μ_N on a common topological space $\mathcal{X}$, large deviations principle (LDP) yields asymptotics of form

$$\mu_N(\mathcal{E}) \approx \exp\big(- v_N \inf_{x\in\mathcal{E}} J(x)\big), \qquad \mathcal{E} \subseteq \mathcal{X},$$

  for a rate function J and speed v_N.

- In dense case (p fixed), C–V get an LDP for $\mu_N(\cdot) = \mathbb{P}(\boldsymbol{G} \in \cdot)$. What does it mean? μ_N live on separate spaces $\mathcal{G}_N \cong \{0,1\}^{\binom{N}{2}}$...

- The space of **graphons** provides a "completion" of $\bigcup_{N\geq 1} \mathcal{G}_N$:

$$\mathcal{W} := \{g : [0,1]^2 \to [0,1] \text{ symmetric, Lebesgue measurable}\},$$

  equipped with a topology coming from the cut-norm:

$$\|f\|_\square := \sup_{S,T\subseteq[0,1]} \left| \int_{S\times T} f(x,y)dxdy \right|.$$

Dense case (Chatterjee–Varadhan '11)

- For a sequence of probability measures μ_N on a common topological space $\mathcal{X}$, large deviations principle (LDP) yields asymptotics of form

$$\mu_N(\mathcal{E}) \approx \exp\big(-v_N \inf_{x\in\mathcal{E}} J(x)\big), \qquad \mathcal{E} \subseteq \mathcal{X},$$

  for a rate function J and speed v_N.

- In dense case (p fixed), C–V get an LDP for $\mu_N(\cdot) = \mathbb{P}(\boldsymbol{G} \in \cdot)$. What does it mean? μ_N live on separate spaces $\mathcal{G}_N \cong \{0,1\}^{\binom{N}{2}}$...

- The space of **graphons** provides a "completion" of $\bigcup_{N\geq 1} \mathcal{G}_N$:

$$\mathcal{W} := \{g : [0,1]^2 \to [0,1] \text{ symmetric, Lebesgue measurable}\},$$

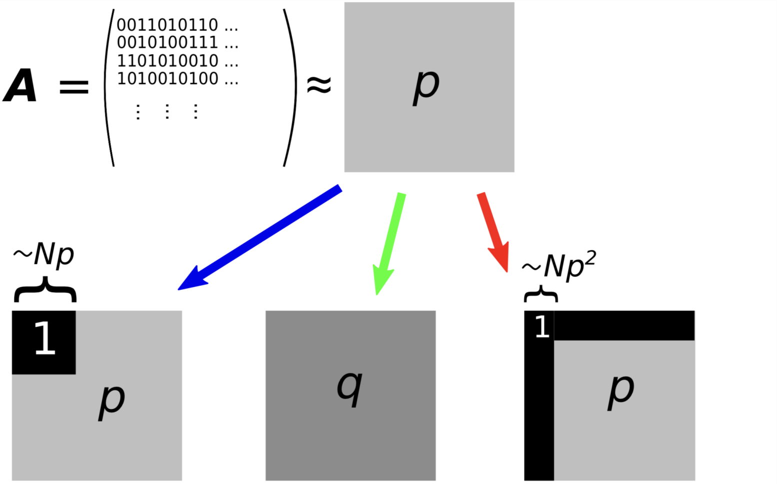
  equipped with a topology coming from the cut-norm:

$$\|f\|_\square := \sup_{S,T\subseteq[0,1]} \left| \int_{S\times T} f(x,y)dxdy \right|.$$

- Graphons are limits of rescaled adjacency matrices, and $\|\cdot\|_\square$ extends the matrix cut-norm $\|M\|_\square = \max_{U,V\subseteq[N]} \left|\sum_{(i,j)\in U\times V} M_{ij}\right|$.

Dense case (Chatterjee–Varadhan '11)

Identify a finite graph $G \in \mathcal{G}_N$ with $g \in \mathcal{W}$ via its adjacency matrix A, putting $g(x, y) := A_{\lfloor Nx \rfloor, \lfloor Ny \rfloor}$. General $g \in \mathcal{W}$ is like a "continuum adjacency matrix".



Dense case (Chatterjee–Varadhan ’11)

Graphon space provides a topological reformulation of the classic regularity method from extremal graph theory.

Dense case (Chatterjee–Varadhan ’11)

Graphon space provides a topological reformulation of the classic regularity method from extremal graph theory.

Key fact 1: The space of graphons with cut-norm topology is compact ($\approx$ Szemerédi’s regularity lemma).

Dense case (Chatterjee–Varadhan ’11)

Graphon space provides a topological reformulation of the classic regularity method from extremal graph theory.

Key fact 1: The space of graphons with cut-norm topology is compact ($\approx$ Szemerédi’s regularity lemma).

Theorem (Chatterjee–Varadhan)

Fix $p \in (0,1)$ and for $N \geq 1$ let $\boldsymbol{G}_N \sim G(N,p)$. The sequence of probability measures $\mu_N(\cdot) = \mathbb{P}(\boldsymbol{G}_N \in \cdot)$ on the topological space of graphons satisfies an LDP (of speed N^2, with explicit rate function).

Dense case (Chatterjee–Varadhan '11)

Graphon space provides a topological reformulation of the classic regularity method from extremal graph theory.

Key fact 1: The space of graphons with cut-norm topology is compact ($\approx$ Szemerédi's regularity lemma).

Theorem (Chatterjee–Varadhan)

Fix $p \in (0,1)$ and for $N \geq 1$ let $\boldsymbol{G}_N \sim G(N,p)$. The sequence of probability measures $\mu_N(\cdot) = \mathbb{P}(\boldsymbol{G}_N \in \cdot)$ on the topological space of graphons satisfies an LDP (of speed N^2, with explicit rate function).

Key fact 2: the subgraph counting functions $\mathcal{N}_H(G)$, suitably extended to graphons, are continuous in the cut-norm topology.
($\approx$ the counting lemma).

Dense case (Chatterjee–Varadhan '11)

Graphon space provides a topological reformulation of the classic regularity method from extremal graph theory.

Key fact 1: The space of graphons with cut-norm topology is compact ($\approx$ Szemerédi's regularity lemma).

Theorem (Chatterjee–Varadhan)

Fix $p \in (0,1)$ and for $N \geq 1$ let $\boldsymbol{G}_N \sim G(N,p)$. The sequence of probability measures $\mu_N(\cdot) = \mathbb{P}(\boldsymbol{G}_N \in \cdot)$ on the topological space of graphons satisfies an LDP (of speed N^2, with explicit rate function).

Key fact 2: the subgraph counting functions $\mathcal{N}_H(G)$, suitably extended to graphons, are continuous in the cut-norm topology.
($\approx$ the counting lemma).

Corollary: upper tails for subgraph counts $\mathcal{N}_H(\boldsymbol{G})$
(just apply the LDP to super-level sets).

Dense case (Chatterjee–Varadhan '11)

Graphon space provides a topological reformulation of the classic regularity method from extremal graph theory.

Key fact 1: The space of graphons with cut-norm topology is compact ($\approx$ Szemerédi's regularity lemma).

Theorem (Chatterjee–Varadhan)

Fix $p \in (0,1)$ and for $N \geq 1$ let $\boldsymbol{G}_N \sim G(N,p)$. The sequence of probability measures $\mu_N(\cdot) = \mathbb{P}(\boldsymbol{G}_N \in \cdot)$ on the topological space of graphons satisfies an LDP (of speed N^2, with explicit rate function).

Key fact 2: the subgraph counting functions $\mathcal{N}_H(G)$, suitably extended to graphons, are continuous in the cut-norm topology.
($\approx$ the counting lemma).

Corollary: upper tails for subgraph counts $\mathcal{N}_H(\boldsymbol{G})$
(just apply the LDP to super-level sets).

Moral: the cut-norm topology is the right topology if you're interested in subgraph counts.

Sparse case: Sharpening the regularity method

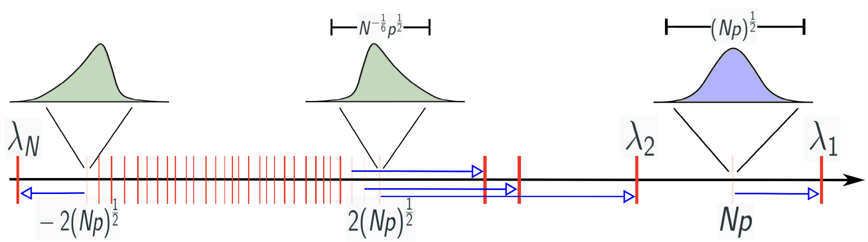
- Regularity and counting lemmas aren't accurate enough to analyze sparse graphs (and unfortunately they're sharp).

Sparse case: Sharpening the regularity method

- Regularity and counting lemmas aren't accurate enough to analyze sparse graphs (and unfortunately they're sharp).
- Existing sparse graph limit theories, such as L^p-graphons [Borgs–Chayes–Cohn–Zhao '14], lack a strong enough counting lemma.

Sparse case: Sharpening the regularity method

- Regularity and counting lemmas aren't accurate enough to analyze sparse graphs (and unfortunately they're sharp).
- Existing sparse graph limit theories, such as L^p-graphons [Borgs–Chayes–Cohn–Zhao '14], lack a strong enough counting lemma.
- We get much improved regularity and counting lemmas after cutting out appropriate small "bad" events (involving outlier eigenvalues).



Spectral regularity lemma for random graphs

Write $\mathcal{A}_N = \{0,1\}^{\binom{N}{2}}$ for the space of adjacency matrices and $\mathcal{X}_N = [0,1]^{\binom{N}{2}}$ for its convex hull (weighted adjacency matrices).

Proposition (Quantitative compactness for $\mathcal{A}_N$)

Let $N \in \mathbb{N}, K \geq 1$, $p \in (0,1)$ with $Np \geq \log N$, and $1 \leq R \leq Np$. There exists a partition $\mathcal{A}_N = \bigsqcup_{j=0}^{J} \mathcal{E}_j$ with the following properties:

(a) $\log J \lesssim RN \log(3 + \frac{R}{Kp})$;

(b) $\mathbb{P}\{\boldsymbol{A}_{N,p} \in \mathcal{E}_0\} \lesssim \exp(-cK^2N^2p^2)$;

(c) For each $1 \leq j \leq J$, there exists $Y_j \in \mathcal{X}_N$ of rank at most R such that $\|A - Y_j\|_{\mathrm{op}} \lesssim \frac{KNp}{\sqrt{R}}$ for all $A \in \mathcal{E}_j$.

Spectral counting lemma for random graphs

Proposition (Lipschitz continuity for homomorphism counts)

Let $H = (V, E)$ of max degree D.

Let $N \in \mathbb{N}$ and $p \in (0, 1)$. For $K \geq 1$ set

$$\mathcal{E}_H(K) = \left\{ X \in \mathcal{X}_N : \exists F \leq H \text{ with } \hom(F, X) > KN^{|V_F|} p^{|E_F|} \right\}.$$

Spectral counting lemma for random graphs

Proposition (Lipschitz continuity for homomorphism counts)

Let $H = (V, E)$ of max degree D.

Let $N \in \mathbb{N}$ and $p \in (0, 1)$. For $K \geq 1$ set

$$\mathcal{E}_H(K) = \Big\{ X \in \mathcal{X}_N : \exists F \leq H \text{ with } \hom(F, X) > K N^{|V_F|} p^{|E_F|} \Big\}.$$

(a) If $N^{-1/D} < p < 1$, then for any $K \geq 2$,

$$\mathbb{P}\left\{ \boldsymbol{A}_{N,p} \in \mathcal{E}_H(K) \right\} \lesssim_H \exp\Big(- c(H) K^{1/|V|} N^2 p^D \Big).$$

Spectral counting lemma for random graphs

Proposition (Lipschitz continuity for homomorphism counts)

Let $H = (V, E)$ of max degree D.
Let $N \in \mathbb{N}$ and $p \in (0,1)$. For $K \geq 1$ set

$$\mathcal{E}_H(K) = \Big\{ X \in \mathcal{X}_N : \exists F \leq H \text{ with } \hom(F, X) > KN^{|V_F|} p^{|E_F|} \Big\}.$$

(a) If $N^{-1/D} < p < 1$, then for any $K \geq 2$,

$$\mathbb{P}\left\{ \boldsymbol{A}_{N,p} \in \mathcal{E}_H(K) \right\} \lesssim_H \exp\Big(- c(H) K^{1/|V|} N^2 p^D \Big).$$

(b) For any $X, Y \in \mathcal{X}_N$ with $X \notin \mathcal{E}_H(K)$, for all $F \leq H$,

$$|\hom(F, X) - \hom(F, Y)| \lesssim_H KN^{|V_F|} p^{|E_F|} \frac{\|X - Y\|_{\mathrm{op}}}{Np^{D}}.$$

Beyond G(N,p)

Special properties of $G(N,p)$ and event $\{\mathcal{N}_H(\boldsymbol{G}) \geq t\}$:

- Independence (of edges)
- Homogeneity (exchangability, same p)
- One dimensional (one H)

Theorem (D.-Bhattacharya '19)

[Cook-D. '18] conclusions extend to:

- Uniform random graph $G^{(m)}(N)$, number of edges $m = \binom{N}{2}p$.
- Random d-regular graph $G^d(N)$, degree $d = Np$ (if H regular).
- $\mathbb{P}\left\{\mathcal{N}_{H_i}(\boldsymbol{G}) \geq (1+\delta_i)\,\mathbb{E}\mathcal{N}_{H_i}(\boldsymbol{G}), i \leq k\right\}$ joint upper tail.
- Inhomogeneous $G(N,\mathbf{p})$ as in Stochastic block model.

Semi-universal: [Cook-D. '18] reduction to variational problem is robust.
But [BGLZ '16] solution - special for $G(N,p)$; Re-done (change $c_H(\delta)$).

Thank you and

Many happy birthdays – HT!