Sparse random graphs with unusually large subgraph counts

Random excursions with Jean Bertoin, July 7, 2021

Amir Dembo, Stanford University

Joint works with Nicholas Cook, Huy Tuan Pham and Sohom Bhattacharya

Overview of the talk

I. Motivation.

II. Conditional structure on LD event: dense graphs.

III. Naïve Mean Field (NMF) and Exponential Random Graph Models (ERGMs).

IV. Typical structure of sparse ERGMs and for sparse graphs given LD event.

V. Some key ideas within the proofs.

I. Extremal vs typical behavior vs large deviations (triangle counts)

Many problems in combinatorics are of the form:

For $\mathcal{G}_n$ discrete, understand the behavior of some function $t : \mathcal{G}_n \to \mathbb{R}$ under a constraint on another function $e : \mathcal{G}_n \to \mathbb{R}$.

I. Extremal vs typical behavior vs large deviations (triangle counts)

Many problems in combinatorics are of the form:

For $\mathcal{G}_n$ discrete, understand the behavior of some function $t : \mathcal{G}_n \to \mathbb{R}$ under a constraint on another function $e : \mathcal{G}_n \to \mathbb{R}$.

Example, triangle counts: $\mathcal{G}_n \cong \{0,1\}^{\binom{n}{2}}$, all simple graphs over $[n]$, $t(G)$, $e(G)$, count numbers of triangles and edges in G

$$t(G) = \mathcal{N}_{K_3}(G) = \sum_{\{i,j,k\} \subset [n]} a_{ij} a_{jk} a_{ik} \,, \qquad e(G) = \mathcal{N}_{K_2}(G) = |E(G)|$$

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

I. Extremal vs typical behavior vs large deviations (triangle counts)

Many problems in combinatorics are of the form:

For $\mathcal{G}_n$ discrete, understand the behavior of some function $t : \mathcal{G}_n \to \mathbb{R}$ under a constraint on another function $e : \mathcal{G}_n \to \mathbb{R}$.

Example, triangle counts: $\mathcal{G}_n \cong \{0,1\}^{\binom{n}{2}}$, all simple graphs over $[n]$, $t(G)$, $e(G)$, count numbers of triangles and edges in G

$$t(G) = \mathcal{N}_{K_3}(G) = \sum_{\{i,j,k\} \subset [n]} a_{ij} a_{jk} a_{ik} \ , \qquad e(G) = \mathcal{N}_{K_2}(G) = |E(G)|$$

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

Extremal problem: Maximize $t(G)$ over $G \in \mathcal{G}_n$ such that $e(G) \leq m$.

I. Extremal vs typical behavior vs large deviations (triangle counts)

Many problems in combinatorics are of the form:

For $\mathcal{G}_n$ discrete, understand the behavior of some function $t : \mathcal{G}_n \to \mathbb{R}$ under a constraint on another function $e : \mathcal{G}_n \to \mathbb{R}$.

Example, triangle counts: $\mathcal{G}_n \cong \{0,1\}^{\binom{n}{2}}$, all simple graphs over $[n]$, $t(G)$, $e(G)$, count numbers of triangles and edges in G

$$t(G) = \mathcal{N}_{K_3}(G) = \sum_{\{i,j,k\} \subset [n]} a_{ij} a_{jk} a_{ik} \, , \qquad e(G) = \mathcal{N}_{K_2}(G) = |E(G)|$$

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

Extremal problem: Maximize $t(G)$ over $G \in \mathcal{G}_n$ such that $e(G) \le m$.

Typical behavior: $\boldsymbol{G} = \boldsymbol{G}(n,p)$ an Erdős–Rényi graph on set of vertices $[n]$,

$$e(\boldsymbol{G}) \approx \mathbb{E}\mathcal{N}_{K_2}(\boldsymbol{G}) = \binom{n}{2} p = m, \qquad t(\boldsymbol{G}) \approx \mathbb{E}\mathcal{N}_{K_3}(\boldsymbol{G}) = \binom{n}{3} p^3 = g_n(m).$$

I. Extremal vs typical behavior vs large deviations (triangle counts)

Many problems in combinatorics are of the form:

For $\mathcal{G}_n$ discrete, understand the behavior of some function $t : \mathcal{G}_n \to \mathbb{R}$ under a constraint on another function $e : \mathcal{G}_n \to \mathbb{R}$.

Example, triangle counts: $\mathcal{G}_n \cong \{0,1\}^{\binom{n}{2}}$, all simple graphs over $[n]$, $t(G)$, $e(G)$, count numbers of triangles and edges in G

$$t(G) = \mathcal{N}_{K_3}(G) = \sum_{\{i,j,k\}\subset[n]} a_{ij}a_{jk}a_{ik} \ , \qquad e(G) = \mathcal{N}_{K_2}(G) = |E(G)|$$

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

Extremal problem: Maximize $t(G)$ over $G \in \mathcal{G}_n$ such that $e(G) \le m$.

Typical behavior: $\boldsymbol{G} = \boldsymbol{G}(n,p)$ an Erdős–Rényi graph on set of vertices $[n]$,

$$e(\boldsymbol{G}) \approx \mathbb{E}\mathcal{N}_{K_2}(\boldsymbol{G}) = \binom{n}{2}p = m, \qquad t(\boldsymbol{G}) \approx \mathbb{E}\mathcal{N}_{K_3}(\boldsymbol{G}) = \binom{n}{3}p^3 = g_n(m).$$

Large deviations: In between (soft-extremizing).

Q.: Law of $\boldsymbol{G}$ given $\mathcal{T}_{K_3}(\delta) = \left\{\mathcal{N}_{K_3}(\boldsymbol{G}) \ge \binom{n}{3}q^3\right\}$, for $q = (1+\delta)^{1/3}p$?

A.: Usually relies on the (LD) asymptotic of $U_{n,p}(K_3,\delta) := -\log \mathbb{P}(\mathcal{T}_{K_3}(\delta))$.

II. Typical structure of $G(n, p)$ on LD upper tail event

- $\boldsymbol{G} = \boldsymbol{G}(n, p)$ an Erdős–Rényi graph on vertices $[n]$, law denoted $\mu_{n,p}(G)$.

II. Typical structure of $G(n,p)$ on LD upper tail event

- $\boldsymbol{G} = \boldsymbol{G}(n,p)$ an Erdős–Rényi graph on vertices $[n]$, law denoted $\mu_{n,p}(G)$.
- Triangles count: $\mathcal{N}_{K_3}(\boldsymbol{G}) = \sum_{\{i,j,k\}\subset[n]} a_{ij} a_{jk} a_{ik}, \qquad \mathbb{E}\mathcal{N}_{K_3}(\boldsymbol{G}) = \binom{n}{3} p^3.$

II. Typical structure of $\boldsymbol{G}(n,p)$ on LD upper tail event

- $\boldsymbol{G} = \boldsymbol{G}(n,p)$ an Erdős–Rényi graph on vertices $[n]$, law denoted $\mu_{n,p}(G)$.
- Triangles count: $\mathcal{N}_{K_3}(\boldsymbol{G}) = \sum_{\{i,j,k\}\subset[n]} a_{ij}a_{jk}a_{ik}, \qquad \mathbb{E}\mathcal{N}_{K_3}(\boldsymbol{G}) = \binom{n}{3}p^3.$
- **Question:** Conditional on extra triangles, i.e. $\mathcal{T}_{K_3}(\delta) = \left\{ \mathcal{N}_{K_3}(\boldsymbol{G}) \geq \binom{n}{3}q^3 \right\}$ for $q = (1+\delta)^{1/3}p > p$, how are the edges distributed?

II. Typical structure of $G(n,p)$ on LD upper tail event

- $\boldsymbol{G} = \boldsymbol{G}(n,p)$ an Erdős–Rényi graph on vertices $[n]$, law denoted $\mu_{n,p}(G)$.
- Triangles count: $\mathcal{N}_{K_3}(\boldsymbol{G}) = \sum_{\{i,j,k\}\subset[n]} a_{ij}a_{jk}a_{ik}$, $\quad \mathbb{E}\mathcal{N}_{K_3}(\boldsymbol{G}) = \binom{n}{3}p^3$.
- **Question:** Conditional on extra triangles, i.e. $\mathcal{T}_{K_3}(\delta) = \left\{\mathcal{N}_{K_3}(\boldsymbol{G}) \geq \binom{n}{3}q^3\right\}$ for $q = (1+\delta)^{1/3}p > p$, how are the edges distributed?

  A few possibilities:

II. Typical structure of $\boldsymbol{G}(n,p)$ on LD upper tail event

- $\boldsymbol{G} = \boldsymbol{G}(n,p)$ an Erdős–Rényi graph on vertices $[n]$, law denoted $\mu_{n,p}(G)$.
- Triangles count: $\mathcal{N}_{K_3}(\boldsymbol{G}) = \sum_{\{i,j,k\}\subset[n]} a_{ij}a_{jk}a_{ik}, \qquad \mathbb{E}\mathcal{N}_{K_3}(\boldsymbol{G}) = \binom{n}{3}p^3.$
- **Question:** Conditional on extra triangles, i.e. $\mathcal{T}_{K_3}(\delta) = \left\{\mathcal{N}_{K_3}(\boldsymbol{G}) \geq \binom{n}{3}q^3\right\}$ for $q = (1+\delta)^{1/3}p > p$, how are the edges distributed?

  A few possibilities:

  (A) As in $\boldsymbol{G}(n,q)$?

II. Typical structure of $\boldsymbol{G}(n,p)$ on LD upper tail event

- $\boldsymbol{G} = \boldsymbol{G}(n,p)$ an Erdős–Rényi graph on vertices $[n]$, law denoted $\mu_{n,p}(G)$.
- Triangles count: $\mathcal{N}_{K_3}(\boldsymbol{G}) = \sum_{\{i,j,k\}\subset[n]} a_{ij}a_{jk}a_{ik}, \qquad \mathbb{E}\mathcal{N}_{K_3}(\boldsymbol{G}) = \binom{n}{3}p^3.$
- **Question:** Conditional on extra triangles, i.e. $\mathcal{T}_{K_3}(\delta) = \left\{ \mathcal{N}_{K_3}(\boldsymbol{G}) \geq \binom{n}{3}q^3 \right\}$ for $q = (1+\delta)^{1/3}p > p$, how are the edges distributed?

  A few possibilities:

  (A) As in $\boldsymbol{G}(n,q)$?

  (B) As in $\boldsymbol{G}(n,p)$ with a small planted clique?

II. Typical structure of $G(n,p)$ on LD upper tail event

- $\boldsymbol{G} = \boldsymbol{G}(n,p)$ an Erdős–Rényi graph on vertices $[n]$, law denoted $\mu_{n,p}(G)$.
- Triangles count: $\mathcal{N}_{K_3}(\boldsymbol{G}) = \sum_{\{i,j,k\}\subset[n]} a_{ij}a_{jk}a_{ik}, \qquad \mathbb{E}\mathcal{N}_{K_3}(\boldsymbol{G}) = \binom{n}{3}p^3.$
- **Question:** Conditional on extra triangles, i.e. $\mathcal{T}_{K_3}(\delta) = \left\{\mathcal{N}_{K_3}(\boldsymbol{G}) \geq \binom{n}{3}q^3\right\}$ for $q = (1+\delta)^{1/3}p > p$, how are the edges distributed?

  A few possibilities:

  (A) As in $\boldsymbol{G}(n,q)$?

  (B) As in $\boldsymbol{G}(n,p)$ with a small planted clique?

  (C) As in $\boldsymbol{G}(n,p)$ with a small planted hub?

II. Typical structure of $G(n, p)$ on LD upper tail event

- $\boldsymbol{G} = \boldsymbol{G}(n, p)$ an Erdős–Rényi graph on vertices $[n]$, law denoted $\mu_{n,p}(G)$.
- Triangles count: $\mathcal{N}_{K_3}(\boldsymbol{G}) = \sum_{\{i,j,k\}\subset[n]} a_{ij}a_{jk}a_{ik}$, $\quad \mathbb{E}\mathcal{N}_{K_3}(\boldsymbol{G}) = \binom{n}{3}p^3$.
- **Question:** Conditional on extra triangles, i.e. $\mathcal{T}_{K_3}(\delta) = \{\mathcal{N}_{K_3}(\boldsymbol{G}) \geq \binom{n}{3}q^3\}$ for $q = (1+\delta)^{1/3}p > p$, how are the edges distributed?

  A few possibilities:

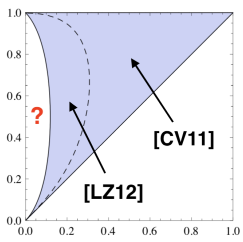
  (A) As in $\boldsymbol{G}(n, q)$?

  (B) As in $\boldsymbol{G}(n, p)$ with a small planted clique?

  (C) As in $\boldsymbol{G}(n, p)$ with a small planted hub?

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

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



III. Gibbs measures and Naïve Mean Field (NMF) approximation

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

III. Gibbs measures and Naïve Mean Field (NMF) approximation

- [Chatterjee–D. '14]: LD for nonlinear functions $f : \{0,1\}^d \to \mathbb{R}$ through the study of Gibbs measures μ_h with density $\mu_h(\{x\}) \propto e^{h(x)}$ for some Hamiltonian $h : \{0,1\}^d \to \mathbb{R}$.
- Taking $e^{h(x)}$ as a "smooth" approximation to the indicator function $\mathbf{1}_{\{f(x) \geq L\}}$, recovers estimates on $\mu_\star(\{x : f(x) \geq L\})$ via the partition function $Z = \sum_{x \in \{0,1\}^d} e^{h(x)} \mu_\star(x)$.

III. Gibbs measures and Naïve Mean Field (NMF) approximation

- [Chatterjee–D. '14]: LD for nonlinear functions $f : \{0,1\}^d \to \mathbb{R}$ through the study of Gibbs measures μ_h with density $\mu_h(\{x\}) \propto e^{h(x)}$ for some Hamiltonian $h : \{0,1\}^d \to \mathbb{R}$.
- Taking $e^{h(x)}$ as a "smooth" approximation to the indicator function $\mathbf{1}_{\{f(x)\geq L\}}$, recovers estimates on $\mu_\star(\{x : f(x) \geq L\})$ via the partition function $Z = \sum_{x\in\{0,1\}^d} e^{h(x)} \mu_\star(x)$.
- Obtain conditions for validity of the NMF approximation:

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

for reference (uniform) measure $\mu_\star$ and the relative entropy $D(\nu\|\mu_\star)$.

III. Gibbs measures and Naïve Mean Field (NMF) approximation

- [Chatterjee–D. '14]: LD for nonlinear functions $f : \{0,1\}^d \to \mathbb{R}$ through the study of Gibbs measures μ_h with density $\mu_h(\{x\}) \propto e^{h(x)}$ for some Hamiltonian $h : \{0,1\}^d \to \mathbb{R}$.
- Taking $e^{h(x)}$ as a "smooth" approximation to the indicator function $\mathbf{1}_{\{f(x)\geq L\}}$, recovers estimates on $\mu_\star(\{x : f(x) \geq L\})$ via the partition function $Z = \sum_{x\in\{0,1\}^d} e^{h(x)} \mu_\star(x)$.
- Obtain conditions for validity of the NMF approximation:

$$\log Z = \sup_{\nu \in M_1(\{0,1\}^d)} \left\{ \int h d\nu - D(\nu \| \mu_\star) \right\} \approx \sup_{\substack{\nu \in M_1(\{0,1\}^d) \\ \text{product measures}}} \left\{ \int h d\nu - D(\nu \| \mu_\star) \right\}$$

  for reference (uniform) measure $\mu_\star$ and the relative entropy $D(\nu \| \mu_\star)$.
- Extended and refined by [Yan '15], [Eldan '16], [Augeri '18], [Austin '18].

III. Gibbs measures and Naïve Mean Field (NMF) approximation

- [Chatterjee–D. '14]: LD for nonlinear functions $f : \{0,1\}^d \to \mathbb{R}$ through the study of Gibbs measures μ_h with density $\mu_h(\{x\}) \propto e^{h(x)}$ for some Hamiltonian $h : \{0,1\}^d \to \mathbb{R}$.
- Taking $e^{h(x)}$ as a "smooth" approximation to the indicator function $\mathbf{1}_{\{f(x)\geq L\}}$, recovers estimates on $\mu_\star(\{x : f(x) \geq L\})$ via the partition function $Z = \sum_{x\in\{0,1\}^d} e^{h(x)}\mu_\star(x)$.
- Obtain conditions for validity of the NMF approximation:

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

  for reference (uniform) measure $\mu_\star$ and the relative entropy $D(\nu\|\mu_\star)$.
- 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 (for $\mu_\star = \mu_{n,p}$, $p \ll 1$).

III. Exponential Random Graph Models (ERGMs)

- $p \in (0,1)$, $\underline{F} = (F_1, \dots, F_\ell)$, connected graphs of max degree Δ.
- Gibbs measures for (Erdős–Rényi law) $\mu_{n,p}$ on $\mathcal{G}_n$ called ERGMs:

$$\mu_{n,p,\beta}(G) = \frac{1}{Z_{n,p}(\beta)} e^{\beta R_{n,p} \cdot h_{\underline{F}}(G)} \mu_{n,p}(G)\,, \qquad G \in \mathcal{G}_n\,,$$

$$h_{\underline{F}}(G) = \sum_{k=1}^{\ell} f_k(t_{F_k}(G))\,, \qquad t_F(G) = \frac{\mathcal{N}_F(G)}{\mathbb{E}\mathcal{N}_F(\boldsymbol{G})}\,.$$

$\beta > 0$, $f_k(\cdot) \uparrow$ promote the appearance of F_k-s in typical sample $\boldsymbol{G}$ of law $\mu_{n,p,\beta}$

The collection $\{\mathcal{N}_{F_k}(\boldsymbol{G}), k \le \ell\}$ forms a *sufficient statistic* for β.

III. Exponential Random Graph Models (ERGMs)

- $p \in (0,1)$, $\underline{F} = (F_1, \ldots, F_\ell)$, connected graphs of max degree Δ.
- Gibbs measures for (Erdős–Rényi law) $\mu_{n,p}$ on $\mathcal{G}_n$ called ERGMs:

$$\mu_{n,p,\beta}(G) = \frac{1}{Z_{n,p}(\beta)} e^{\beta R_{n,p} \cdot h_{\underline{F}}(G)} \mu_{n,p}(G)\,, \qquad G \in \mathcal{G}_n\,,$$

$$h_{\underline{F}}(G) = \sum_{k=1}^{\ell} f_k(t_{F_k}(G))\,, \qquad t_F(G) = \frac{\mathcal{N}_F(G)}{\mathbb{E}\mathcal{N}_F(\boldsymbol{G})}\,.$$

$\beta > 0$, $f_k(\cdot) \uparrow$ promote the appearance of F_k-s in typical sample $\boldsymbol{G}$ of law $\mu_{n,p,\beta}$

The collection $\{\mathcal{N}_{F_k}(\boldsymbol{G}), k \le \ell\}$ forms a *sufficient statistic* for β.

- LD for (joint) upper tails $\mathcal{T}_{\underline{F}}(\underline{\delta}) = \{\, t_{F_k}(\boldsymbol{G}) \ge 1 + \delta_k,\ k \in [\ell]\}$:

If $-R_{n,p}^{-1} \log \mu_{n,p}(\mathcal{T}_{\underline{F}}(\underline{\delta})) \to J_{\underline{F}}(\underline{\delta})$, then (by Varadhan's LD Lemma),

$$R_{n,p}^{-1} \log Z_{n,p}(\beta) \to \Psi_{\underline{F}}(\beta) = \sup_{\underline{\delta} \in \mathbb{R}_+^{\ell}} \Big\{ \beta \sum_{k=1}^{\ell} f_k(1+\delta_k) - J_{\underline{F}}(\underline{\delta}) \Big\}\,.$$

III. Exponential Random Graph Models (ERGMs)

- $p \in (0,1)$, $\underline{F} = (F_1, \dots, F_\ell)$, connected graphs of max degree Δ.
- Gibbs measures for (Erdős–Rényi law) $\mu_{n,p}$ on $\mathcal{G}_n$ called ERGMs:

$$\mu_{n,p,\beta}(G) = \frac{1}{Z_{n,p}(\beta)} e^{\beta R_{n,p} \cdot h_{\underline{F}}(G)} \mu_{n,p}(G)\,, \qquad G \in \mathcal{G}_n\,,$$

$$h_{\underline{F}}(G) = \sum_{k=1}^{\ell} f_k(t_{F_k}(G))\,, \qquad t_F(G) = \frac{\mathcal{N}_F(G)}{\mathbb{E}\mathcal{N}_F(\boldsymbol{G})}\,.$$

$\beta > 0$, $f_k(\cdot) \uparrow$ promote the appearance of F_k-s in typical sample $\boldsymbol{G}$ of law $\mu_{n,p,\beta}$

The collection $\{\mathcal{N}_{F_k}(\boldsymbol{G}), k \le \ell\}$ forms a *sufficient statistic* for β.

- LD for (joint) upper tails $\mathcal{T}_{\underline{F}}(\underline{\delta}) = \{\, t_{F_k}(\boldsymbol{G}) \ge 1 + \delta_k,\ k \in [\ell]\}$:

If $-R_{n,p}^{-1} \log \mu_{n,p}(\mathcal{T}_{\underline{F}}(\underline{\delta})) \to J_{\underline{F}}(\underline{\delta})$, then (by Varadhan's LD Lemma),

$$R_{n,p}^{-1} \log Z_{n,p}(\beta) \to \Psi_{\underline{F}}(\beta) = \sup_{\underline{\delta} \in \mathbb{R}_+^\ell} \Big\{ \beta \sum_{k=1}^{\ell} f_k(1+\delta_k) - J_{\underline{F}}(\underline{\delta}) \Big\}\,.$$

- Structure of a typical sample from $\mu_{n,p,\beta}$ determined by the structure of Erdős–Rényi graph $\boldsymbol{G}$, conditional upon the rare event $\mathcal{T}_{\underline{F}}(\underline{\delta})$.

III. Edge–F model

- Edge–F model (ERGM with $\ell = 1$; sociology literature):

$$\mu_{n,p,\beta}(G) = \frac{1}{Z_{n,p}(\beta)} e^{\beta R_{n,p} f(t_F(G))} \mu_{n,p}(G)\,, \qquad t_F(G) = \frac{\mathcal{N}_F(G)}{\mathbb{E}\mathcal{N}_F(\boldsymbol{G})}\,.$$

Dense case (p and $\beta \in \mathbb{R}$ fixed, $R_{n,p} = n^2$, $f(x) = x$):
[Bhamidi–Bressler–Sly '08], [Chatterjee–Diaconis '11], [Lubetzky–Zhao '12].

III. Edge–F model

- Edge–F model (ERGM with $\ell = 1$; sociology literature):

$$\mu_{n,p,\beta}(G) = \frac{1}{Z_{n,p}(\beta)} e^{\beta R_{n,p} f(t_F(G))} \mu_{n,p}(G)\,, \qquad t_F(G) = \frac{\mathcal{N}_F(G)}{\mathbb{E}\mathcal{N}_F(\boldsymbol{G})}\,.$$

  Dense case (p and $\beta \in \mathbb{R}$ fixed, $R_{n,p} = n^2$, $f(x) = x$):
  [Bhamidi–Bressler–Sly '08], [Chatterjee–Diaconis '11], [Lubetzky–Zhao '12].

- Problems with ERGMs:
  - For p and $\beta > 0$ fixed, $\boldsymbol{G}$ looks like an (A) $\boldsymbol{G}(n, q)$ graph!
  - Degeneracy: in some range of p, β, typical $\boldsymbol{G}$ is close to empty or full.

III. Edge–F model

- Edge–F model (ERGM with $\ell = 1$; sociology literature):

$$\mu_{n,p,\beta}(G) = \frac{1}{Z_{n,p}(\beta)} e^{\beta R_{n,p} f(t_F(G))} \mu_{n,p}(G)\,, \qquad t_F(G) = \frac{\mathcal{N}_F(G)}{\mathbb{E}\mathcal{N}_F(\boldsymbol{G})}\,.$$

  Dense case (p and $\beta \in \mathbb{R}$ fixed, $R_{n,p} = n^2$, $f(x) = x$):
  [Bhamidi–Bressler–Sly '08], [Chatterjee–Diaconis '11], [Lubetzky–Zhao '12].

- Problems with ERGMs:
  - For p and $\beta > 0$ fixed, $\boldsymbol{G}$ looks like an (A) $\boldsymbol{G}(n, q)$ graph!
  - Degeneracy: in some range of p, β, typical $\boldsymbol{G}$ is close to empty or full.

- [Cook–D. '21$^+$]: Quantitative control on $\log Z_{n,p}(\beta)$, $p \ll 1$, allows sparse ERGMs (richer structure, see [Eldan–Gross '17]).

  Resolve degeneracy via $f(x) = x^\gamma$ (with $\gamma < \gamma_F < 1$; for dense case [LZ12]).

IV. Upper tail for subgraph counts

- For F of max degree Δ and rate $R_{n,p} = n^2 p^\Delta \log(1/p)$, we aim at

$$U_{n,p}(F,\delta) = -\log \mu_{n,p}\big(\mathcal{T}_F(\delta)\big) \approx R_{n,p} J_F(\delta)\,,$$

  matching $-\log$ probability of a planted clique or hub of appropriate sizes.

- [Chatterjee–D. '14] for $n^{-a_F} \ll p \ll 1$, $a_F = \frac{c}{\Delta e(F)}$, get NMF for upper tail:

$$U_{n,p}(F,\delta) \approx \inf_{X \in \mathcal{X}_n = [0,1]^{\binom{n}{2}}} \big\{ D(\mu_X \| \mu_{n,p}) : t_F(X) \geq 1+\delta \big\} := \Phi_{n,p}(F,\delta)\,.$$

  [B. Bhattacharya–Ganguly–Lubetzky–Zhao '16]: $R_{n,p}^{-1}\Phi_{n,p}(F,\delta) \to J_F(\delta)$ when $n^{-1/\Delta} \ll p \ll 1$. Also, [Eldan '16] improves to $a_F = \frac{1}{6e(F)} - \varepsilon$.

- For F of max degree Δ and rate $R_{n,p} = n^2 p^\Delta \log(1/p)$, we aim at

$$U_{n,p}(F,\delta) = -\log \mu_{n,p}\big(\mathcal{T}_F(\delta)\big) \approx R_{n,p} J_F(\delta)\,,$$

  matching $-\log$ probability of a planted clique or hub of appropriate sizes.
- [Chatterjee–D. '14] for $n^{-a_F} \ll p \ll 1$, $a_F = \frac{c}{\Delta e(F)}$, get NMF for upper tail:

$$U_{n,p}(F,\delta) \approx \inf_{X \in \mathcal{X}_n = [0,1]^{\binom{n}{2}}} \big\{ D(\mu_X \| \mu_{n,p}) : t_F(X) \geq 1+\delta \big\} := \Phi_{n,p}(F,\delta)\,.$$

  [B. Bhattacharya–Ganguly–Lubetzky–Zhao '16]: $R_{n,p}^{-1} \Phi_{n,p}(F,\delta) \to J_F(\delta)$ when $n^{-1/\Delta} \ll p \ll 1$. Also, [Eldan '16] improves to $a_F = \frac{1}{6e(F)} - \varepsilon$.

- [Cook–D. '18]: $a_F = \frac{1}{2\Delta}$; [Cook–D.–Pham '21]: $a_F = \frac{1}{\Delta+1}$.
- Special H: [Augeri '18]: $a_{C_\ell} = \frac{1}{2}$, $\ell \geq 3$; [Cook–D. '18]: $\frac{2}{\Delta+1}$, for F a star; [Basak–Basu '19], after [Harel–Mousset–Samotij '19]: $\frac{2}{\Delta} - \varepsilon$, for Δ-regular F.

IV. Beyond $\boldsymbol{G}(n, p)$ and beyond a single event $\mathcal{T}_F(\delta)$

$\boldsymbol{G}(n, p)$: Independence (of edges), Homogeneity (exchangeability, same p).

IV. Beyond $G(n,p)$ and beyond a single event $\mathcal{T}_F(\delta)$

$\boldsymbol{G}(n,p)$: Independence (of edges), Homogeneity (exchangeability, same p).

Theorem [D.-S. Bhattacharya '19]

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

- Uniform random graph $\boldsymbol{G}^{(m)}(n)$, number of edges $m = \binom{n}{2} p$.
- Random d-regular graph, degree $d = np$ (for F "regular plus trees").
- Inhomogeneous $\boldsymbol{G}(n, \mathbf{p})$ (as in the Stochastic Block Model).

IV. Beyond $G(n, p)$ and beyond a single event $\mathcal{T}_F(\delta)$

$\boldsymbol{G}(n, p)$: Independence (of edges), Homogeneity (exchangeability, same p).

Theorem [D.-S. Bhattacharya '19]

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

- Uniform random graph $\boldsymbol{G}^{(m)}(n)$, number of edges $m = \binom{n}{2} p$.
- Random d-regular graph, degree $d = np$ (for F "regular plus trees").
- Inhomogeneous $\boldsymbol{G}(n, \mathbf{p})$ (as in the Stochastic Block Model).

NMF semi-universal. Re-do [BGLZ16] part, changes $J_F(\delta)$ (see [Gunby '20]).

IV. Beyond $G(n,p)$ and beyond a single event $\mathcal{T}_F(\delta)$

$\boldsymbol{G}(n,p)$: Independence (of edges), Homogeneity (exchangeability, same p).

Theorem [D.-S. Bhattacharya '19]

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

- Uniform random graph $\boldsymbol{G}^{(m)}(n)$, number of edges $m = \binom{n}{2}p$.
- Random d-regular graph, degree $d = np$ (for F "regular plus trees").
- Inhomogeneous $\boldsymbol{G}(n, \mathbf{p})$ (as in the Stochastic Block Model).

NMF semi-universal. Re-do [BGLZ16] part, changes $J_F(\delta)$ (see [Gunby '20]).

Theorem [CDP '21]

$\{F_k\}$ of max degree Δ, $n^{-1/(\Delta+1)} \ll p$. NMF for joint upper tail ($\underline{\delta} \in \mathbb{R}_+^\ell$):

$$U_{n,p}(\underline{F}, \underline{\delta}) \approx \Phi_{n,p}(\underline{F}, \underline{\delta}) = \inf_{X \in \mathcal{X}_n} \left\{ D(\mu_X \| \mu_{n,p}) : t_{F_k}(X) \geq 1 + \delta_k \,, \quad \forall k \in [\ell] \right\}.$$

[D.–S. Bhattacharya '19]: $\qquad R_{n,p}^{-1} \Phi_{n,p}(\underline{F}, \underline{\delta}) \to J_{\underline{F}}(\underline{\delta})$.

IV. Typical structure on upper tail event: sparse random graphs

F of max degree Δ, $\mathcal{T}_F(\delta) = \{t_F(\boldsymbol{G}) \geq 1 + \delta\}$, $t_F(\boldsymbol{G}) = \mathcal{N}_F(\boldsymbol{G}) / \mathbb{E}\mathcal{N}_F(\boldsymbol{G})$.

Conj: $\left\{\boldsymbol{G} \,\middle|\, \mathcal{T}_F(\delta)\right\} \quad \approx \quad \boldsymbol{G}(n,p)$ + planted clique or hub.

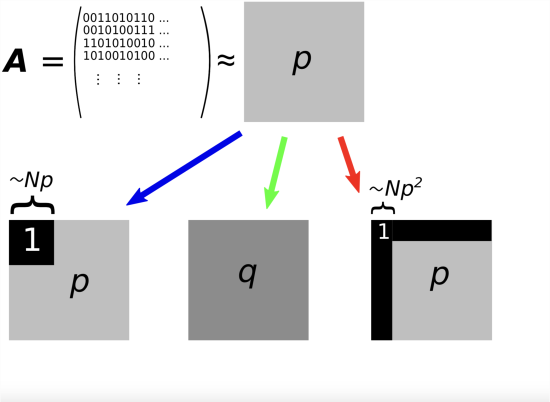
($n^{-1/\Delta} \ll p \ll 1$, with clique vs hub determined by F, δ).

F of max degree Δ, $\mathcal{T}_F(\delta) = \{t_F(\boldsymbol{G}) \geq 1 + \delta\}$, $t_F(\boldsymbol{G}) = \mathcal{N}_F(\boldsymbol{G}) / \mathbb{E}\mathcal{N}_F(\boldsymbol{G})$.

Conj: $\left\{\boldsymbol{G} \,\middle|\, \mathcal{T}_F(\delta)\right\} \quad \approx \quad \boldsymbol{G}(n,p)$ + planted clique or hub.

($n^{-1/\Delta} \ll p \ll 1$, with clique vs hub determined by F, δ).

$$[N \mapsto n,\ p \mapsto p^{\Delta/2}] \qquad R = n^2 p^{\Delta} \log(1/p), \qquad e^{-JR} = p^{\mathbf{J\,n^2 p^{\Delta}}}$$

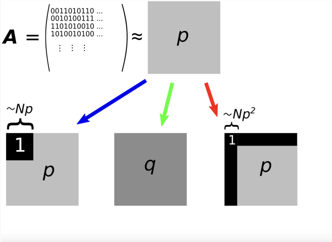


IV. Typical structure on the upper tail event: sparse random graphs

F of max degree Δ, $\mathcal{T}_F(\delta) = \{t_F(\mathbf{G}) \geq 1 + \delta\}$, $t_F(\mathbf{G}) = \mathcal{N}_F(\mathbf{G}) / \mathbb{E}\mathcal{N}_F(\mathbf{G})$.

Conj: $\mathbf{G} \,\big|\, \mathcal{T}_F(\delta) \quad \approx \quad \mathbf{G}(n,p)$ + planted clique or hub.

($n^{-1/\Delta} \ll p \ll 1$, with clique vs hub determined by F, δ).

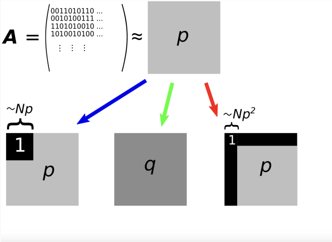


- [HMS19]: $n^{-2/\Delta+\varepsilon} \ll p \ll 1$, conditional on $\mathcal{T}_{K_{\Delta+1}}(\delta)$, WHP $\mathbf{G}$ contains almost-clique [or almost-hub] of the correct area ($= Jn^2p^\Delta$).

F of max degree Δ, $\mathcal{T}_F(\delta) = \{t_F(\boldsymbol{G}) \geq 1+\delta\}$, $t_F(\boldsymbol{G}) = \mathcal{N}_F(\boldsymbol{G}) / \mathbb{E}\mathcal{N}_F(\boldsymbol{G})$.

Conj: $\boldsymbol{G} \,\Big|\, \mathcal{T}_F(\delta) \quad \approx \quad \boldsymbol{G}(n,p)$ + planted clique or hub.

$(n^{-1/\Delta} \ll p \ll 1$, with clique vs hub determined by F, $\delta)$.



- [HMS19]: $n^{-2/\Delta+\varepsilon} \ll p \ll 1$, conditional on $\mathcal{T}_{K_{\Delta+1}}(\delta)$, WHP $\boldsymbol{G}$ contains almost-clique [or almost-hub] of the correct area $(= Jn^2p^\Delta)$.
- [Cook–D. '21$^+$]: Same for any F, when $n^{-1/(\Delta+1)} \ll p \ll 1$
  & stronger statement of the above **Conj.** when $n^{-1/(2\Delta_\star)} \ll p \ll 1$.
  (here $2\Delta_\star := \max_{(u,v)\in E(F)}\{\deg(u)+\deg(v)\} \in [\Delta+1, 2\Delta]$).

IV. Typical sparse ERGM sample

Fix $\underline{F}$, $\underline{\delta} \in \mathbb{R}_+^\ell$, $\beta > 0$. $\exists \mathcal{Q}(\underline{\delta})$ finite set of optimizers in variation problem $J_{\underline{F}}(\underline{\delta})$ and a finite(?) set $\mathcal{R}(\beta)$ of optimizers for $\Psi_{\underline{F}}(\beta)$ (e.g. in Edge-K_3 model).

Fix $\underline{F}$, $\underline{\delta} \in \mathbb{R}_+^\ell$, $\beta > 0$. $\exists \mathcal{Q}(\underline{\delta})$ finite set of optimizers in variation problem $J_{\underline{F}}(\underline{\delta})$ and a finite(?) set $\mathcal{R}(\beta)$ of optimizers for $\Psi_{\underline{F}}(\beta)$ (e.g. in Edge-K_3 model).

Theorem [CD21$^+$]

Suppose $n^{-1/(\Delta+1)} \ll p \ll 1$ and $\varepsilon > 0$ fixed. Then, WHP on $\mu_{n,p}(\,\cdot\,|\mathcal{T}_{\underline{F}}(\underline{\delta}))$, for some $(a,b) \in \mathcal{Q}(\underline{\delta})$ and I, I' disjoint with $|I| = [\sqrt{ap^\Delta} n]$, $|I'| = [bp^\Delta n]$:

- min-degree$(\boldsymbol{G}[I]) \geq (1-\epsilon)|I|$,
- degrees of at least $(1-\varepsilon)|I'|$ vertices of $\boldsymbol{G}$ exceed $(1-\varepsilon)n$.

If $n^{-1/(2\Delta_\star)} \ll p \ll 1$, WHP also spectral norm $\|A_{\boldsymbol{G}} - X^{I,I'}\|_{\ell_2 \to \ell_2} = o(p^{\Delta/2} n)$ for some such I, I' and the clique-hub matrix

$$X_{ij}^{I,I'} = p + (1-p)\left[\mathbb{1}_{i,j\in I} + \mathbb{1}_{(i,j)\in I' \times [n] \setminus I'}\right].$$

Taking $(a,b) \in \mathcal{R}(\beta)$, the above applies WHP for typical $\boldsymbol{G}$ from the ERGM $\mu_{n,p,\beta}$.

Fix $\underline{F}$, $\underline{\delta} \in \mathbb{R}_+^{\ell}$, $\beta > 0$. $\exists \mathcal{Q}(\underline{\delta})$ finite set of optimizers in variation problem $J_{\underline{F}}(\underline{\delta})$ and a finite(?) set $\mathcal{R}(\beta)$ of optimizers for $\Psi_{\underline{F}}(\beta)$ (e.g. in Edge-K_3 model).

Theorem [CD21$^+$]

Suppose $n^{-1/(\Delta+1)} \ll p \ll 1$ and $\varepsilon > 0$ fixed. Then, WHP on $\mu_{n,p}(\,\cdot\,|\mathcal{T}_{\underline{F}}(\underline{\delta}))$, for some $(a, b) \in \mathcal{Q}(\underline{\delta})$ and I, I' disjoint with $|I| = [\sqrt{ap^{\Delta}}n]$, $|I'| = [bp^{\Delta}n]$:

- min-degree$(\boldsymbol{G}[I]) \geq (1 - \epsilon)|I|$,
- degrees of at least $(1 - \varepsilon)|I'|$ vertices of $\boldsymbol{G}$ exceed $(1 - \varepsilon)n$.

If $n^{-1/(2\Delta_\star)} \ll p \ll 1$, WHP also spectral norm $\|A_{\boldsymbol{G}} - X^{I,I'}\|_{\ell_2 \to \ell_2} = o(p^{\Delta/2}n)$ for some such I, I' and the clique-hub matrix

$$X_{ij}^{I,I'} = p + (1-p)\left[\mathbb{1}_{i,j \in I} + \mathbb{1}_{(i,j) \in I' \times [n] \setminus I'}\right].$$

Taking $(a, b) \in \mathcal{R}(\beta)$, the above applies WHP for typical $\boldsymbol{G}$ from the ERGM $\mu_{n,p,\beta}$.

Proof crucially uses stability results for the relevant variational problems.

V. Dense graphs: LDP for graphons ([CV11])

- The LDP with rate R_n and rate function $J(\cdot)$, for μ_n on common space $\mathcal{X}$ is of form

$$-\log \mu_n(\mathcal{E}) \approx R_n \inf_{x \in \mathcal{E}} \{J(x)\}, \qquad \mathcal{E} \subseteq \mathcal{X}.$$

V. Dense graphs: LDP for graphons ([CV11])

- The LDP with rate R_n and rate function $J(\cdot)$, for μ_n on common space $\mathcal{X}$ is of form

$$-\log \mu_n(\mathcal{E}) \approx R_n \inf_{x \in \mathcal{E}} \{J(x)\}, \qquad \mathcal{E} \subseteq \mathcal{X}.$$

- Identifying $G \in \mathcal{G}_n$ via $\mathbf{A} = \mathbf{A}_G$, with $g_G(x, y) := \mathbf{A}_{\lfloor nx \rfloor, \lfloor ny \rfloor}$ embeds finite graphs in

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

(where graphon $g \in \mathcal{W}$ is like a "continuum adjacency matrix").

V. Dense graphs: LDP for graphons ([CV11])

- The LDP with rate R_n and rate function $J(\cdot)$, for μ_n on common space $\mathcal{X}$ is of form

$$-\log \mu_n(\mathcal{E}) \approx R_n \inf_{x \in \mathcal{E}} \{J(x)\}, \qquad \mathcal{E} \subseteq \mathcal{X}.$$

- Identifying $G \in \mathcal{G}_n$ via $\mathbf{A} = \mathbf{A}_G$, with $g_G(x,y) := \mathbf{A}_{\lfloor nx \rfloor, \lfloor ny \rfloor}$ embeds finite graphs in

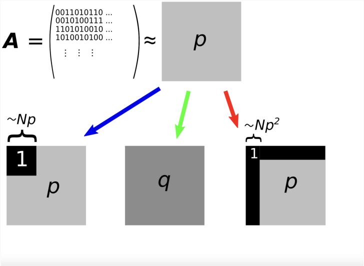
$$\mathcal{W} := \{g : [0,1]^2 \to [0,1] \text{ symmetric, Lebesgue measurable}\}$$

(where graphon $g \in \mathcal{W}$ is like a "continuum adjacency matrix").

- For p fixed, [CV11] get LDP for $\mu_{n,p}(\cdot) = \mathbb{P}(\boldsymbol{G} \in \cdot)$, in $\mathcal{W}$ equipped with the cut-norm:

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

(extension of matrix cut-norm $\|M\|_\square = \max_{U,V \subseteq [n]} \{ |\langle \mathbb{1}_U, M \mathbb{1}_V \rangle| \}$).



V. Sparse graphs: quantitative regularity via covering by convex sets [CD18]

$(\mathcal{W}, \square)$ is a topological reformulation of regularity method from extremal graph theory.

$\approx$ Szemerédi's regularity lemma: quotient (vertex relabeling) makes $(\mathcal{W}, \square)$ compact.

$\Longrightarrow$ **Key** to [CV11] proof of the LDP of rate n^2 for $\mu_{n,p}$ in $\mathcal{W}$ (p fixed).

$\approx$ Counting lemma: $G \mapsto \mathcal{N}_F(G)$ extends continuously to $(\mathcal{W}, \square)$.

$\Longrightarrow$ Applying the LDP to super-level sets, yields upper tails for $\mathcal{N}_F(\boldsymbol{G})$.

V. Sparse graphs: quantitative regularity via covering by convex sets [CD18]

$(\mathcal{W}, \square)$ is a topological reformulation of regularity method from extremal graph theory.

$\approx$ Szemerédi's regularity lemma: quotient (vertex relabeling) makes $(\mathcal{W}, \square)$ compact.

$\Longrightarrow$ **Key** to [CV11] proof of the LDP of rate n^2 for $\mu_{n,p}$ in $\mathcal{W}$ (p fixed).

$\approx$ Counting lemma: $G \mapsto \mathcal{N}_F(G)$ extends continuously to $(\mathcal{W}, \square)$.

$\Longrightarrow$ Applying the LDP to super-level sets, yields upper tails for $\mathcal{N}_F(\boldsymbol{G})$.

Moral: Cut-norm topology perfect for subgraph counts in dense graphs.

V. Sparse graphs: quantitative regularity via covering by convex sets [CD18]

$(\mathcal{W}, \square)$ is a topological reformulation of regularity method from extremal graph theory.

$\approx$ Szemerédi's regularity lemma: quotient (vertex relabeling) makes $(\mathcal{W}, \square)$ compact.

$\Longrightarrow$ **Key** to [CV11] proof of the LDP of rate n^2 for $\mu_{n,p}$ in $\mathcal{W}$ (p fixed).

$\approx$ Counting lemma: $G \mapsto \mathcal{N}_F(G)$ extends continuously to $(\mathcal{W}, \square)$.

$\Longrightarrow$ Applying the LDP to super-level sets, yields upper tails for $\mathcal{N}_F(\boldsymbol{G})$.

Moral: Cut-norm topology perfect for subgraph counts in dense graphs.

Sparse case: The regularity and counting lemmas are not accurate enough to analyze sparse graphs (and unfortunately both are sharp).

V. Sparse graphs: quantitative regularity via covering by convex sets [CD18]

$(\mathcal{W}, \Box)$ is a topological reformulation of regularity method from extremal graph theory.

$\approx$ Szemerédi's regularity lemma: quotient (vertex relabeling) makes $(\mathcal{W}, \Box)$ compact.

$\Longrightarrow$ **Key** to [CV11] proof of the LDP of rate n^2 for $\mu_{n,p}$ in $\mathcal{W}$ (p fixed).

$\approx$ Counting lemma: $G \mapsto \mathcal{N}_F(G)$ extends continuously to $(\mathcal{W}, \Box)$.

$\Longrightarrow$ Applying the LDP to super-level sets, yields upper tails for $\mathcal{N}_F(\boldsymbol{G})$.

Moral: Cut-norm topology perfect for subgraph counts in dense graphs.

Sparse case: The regularity and counting lemmas are not accurate enough to analyze sparse graphs (and unfortunately both are sharp).

Solution:

$\log \mu_{n,p}(\mathcal{K}) \leq -R_{n,p} \inf_{x \in \mathcal{K}} \{J(x)\}$ for any $\mathcal{K} \subset \mathcal{X}_n$ convex

(non-asymptotic LD bound due to the min-max theorem).

V. Sparse graphs: quantitative regularity via covering by convex sets [CD18]

$(\mathcal{W}, \Box)$ is a topological reformulation of regularity method from extremal graph theory.

$\approx$ Szemerédi's regularity lemma: quotient (vertex relabeling) makes $(\mathcal{W}, \Box)$ compact.

$\Longrightarrow$ **Key** to [CV11] proof of the LDP of rate n^2 for $\mu_{n,p}$ in $\mathcal{W}$ (p fixed).

$\approx$ Counting lemma: $G \mapsto \mathcal{N}_F(G)$ extends continuously to $(\mathcal{W}, \Box)$.

$\Longrightarrow$ Applying the LDP to super-level sets, yields upper tails for $\mathcal{N}_F(\boldsymbol{G})$.

Moral: Cut-norm topology perfect for subgraph counts in dense graphs.

Sparse case: The regularity and counting lemmas are not accurate enough to analyze sparse graphs (and unfortunately both are sharp).

Solution:

$\log \mu_{n,p}(\mathcal{K}) \leq -R_{n,p} \inf_{x \in \mathcal{K}}\{J(x)\}$ for any $\mathcal{K} \subset \mathcal{X}_n$ convex

(non-asymptotic LD bound due to the min-max theorem).

$\Longrightarrow$ Suffices to cover $\mathcal{G}_n \subset \mathcal{X}_n$ by $\mathcal{E}_K$ with $\log \mu_{n,p}(\mathcal{E}_K) \leq -R_{n,p}\, K$ (K-negligible),
together with $\exp(o(R_{n,p}))$ convex sets (= quantitative compactness),
on each of which the variation of $\mathcal{N}_F(\cdot)$ is controlled (= quantitative counting lemma).

V. Spectral quantitative regularity [CD18]

Recall $R_{n,p} = n^2 p^\Delta \log(1/p)$.

Regard $G \in \mathcal{G}_n \subset \mathcal{X}_n = [0,1]^{\binom{n}{2}}$, adjacency matrices for edge-weighted graphs.

Matrix spectral norm: $\|M\|_{\ell_2 \to \ell_2} = \sup_{\|u\|_2, \|v\|_2 \le 1} |\langle u, Mv \rangle|$.

V. Spectral quantitative regularity [CD18]

Recall $R_{n,p} = n^2 p^\Delta \log(1/p)$.

Regard $G \in \mathcal{G}_n \subset \mathcal{X}_n = [0,1]^{\binom{n}{2}}$, adjacency matrices for edge-weighted graphs.

Matrix spectral norm: $\|M\|_{\ell_2 \to \ell_2} = \sup_{\|u\|_2, \|v\|_2 \le 1} |\langle u, Mv \rangle|$.

Spectral regularity & counting lemma [CD18]

Fix Δ', $\varepsilon > 0$ (small), $L, K \geq 1$ (large).

(R) For $np^{\Delta'} \gg \log n$ can cover $\mathcal{X}_n$ by K-negligible $\mathcal{E}_K$ and $\exp(o(R_{n,p}))$ spectral-norm balls of diameter $\varepsilon \boldsymbol{np^{\Delta'/2}}$.

V. Spectral quantitative regularity [CD18]

Recall $R_{n,p} = n^2 p^\Delta \log(1/p)$.

Regard $G \in \mathcal{G}_n \subset \mathcal{X}_n = [0,1]^{\binom{n}{2}}$, adjacency matrices for edge-weighted graphs.

Matrix spectral norm: $\|M\|_{\ell_2 \to \ell_2} = \sup_{\|u\|_2, \|v\|_2 \leq 1} |\langle u, Mv \rangle|$.

Spectral regularity & counting lemma [CD18]

Fix Δ', $\varepsilon > 0$ (small), $L, K \geq 1$ (large).

(R) For $np^{\Delta'} \gg \log n$ can cover $\mathcal{X}_n$ by K-negligible $\mathcal{E}_K$ and $\exp(o(R_{n,p}))$ spectral-norm balls of diameter $\varepsilon np^{\Delta'/2}$.

(C) If $\mathcal{K} \subseteq \mathcal{X}_n$ convex of spectral diameter at most $\varepsilon np^{\Delta_\star}$ is such that $\inf_{X \in \mathcal{K}} \{t_{F'}(X)\} \leq L$ for strictly induced sub-graphs $F' \prec F$, then

$$\sup_{X,Y \in \mathcal{K}} |t_F(X) - t_F(Y)| \leq c_F L \varepsilon \,.$$

V. Quantitative regularity: choice of norm ([CDP21] vs. [CD18])

Matrix cut-norm with cutoff scale $n_0 = np^{\Delta-1}$:

$\|M\|_{\mathrm{B}^*} = \sup_{U,V\subset[n]} \left\{ \frac{|\langle \mathbb{1}_U, M\,\mathbb{1}_V\rangle|}{\|\mathbb{1}_U\otimes\mathbb{1}_V\|_{\mathrm{B}}} \right\}$, $\qquad \|\mathbb{1}_U\otimes\mathbb{1}_V\|_{\mathrm{B}} = (|U|\vee n_0)\cdot(|V|\vee n_0)$.

$\mathbb{1}_I\otimes\mathbb{1}_I$ for clique on $|I| = O(np^{\Delta/2}) \geq n_0$, vs. hub on $I'\times[n]$, $|I'| = O(np^{\Delta}) \ll n_0$.

V. Quantitative regularity: choice of norm ([CDP21] vs. [CD18])

Matrix cut-norm with cutoff scale $n_0 = np^{\Delta-1}$:

$\|M\|_{B^*} = \sup_{U,V\subset[n]} \left\{\frac{|\langle \mathbb{1}_U, M\,\mathbb{1}_V\rangle|}{\|\,\mathbb{1}_U \otimes \mathbb{1}_V\,\|_B}\right\}, \qquad \|\,\mathbb{1}_U \otimes \mathbb{1}_V\,\|_B = (|U| \vee n_0)\cdot(|V| \vee n_0).$

$\mathbb{1}_I \otimes \mathbb{1}_I$ for clique on $|I| = O(np^{\Delta/2}) \geq n_0$, vs. hub on $I' \times [n]$, $|I'| = O(np^{\Delta}) \ll n_0$.

B^*-regularity & counting lemma [CDP21]

Fix $\Delta' = \Delta + 1$, $\varepsilon > 0$ (small), $L, K \geq 1$ (large).

(R) For $np^{\Delta'} \gg \log n$, can cover $\mathcal{G}_n$ by K-negligible set $\mathcal{E}_K$ and $\exp(o(R_{n,p}))$ balls in $\mathcal{X}_n$ of B^*-diameter εp.

(C) If $\mathcal{K} \subseteq \mathcal{G}_n$ of B^*-diameter at most εp is such that $\inf_{X\in\mathcal{K}}\{t_{F'}(X)\} \leq L$ for proper sub-graphs $F' < F$, then

$$\sup_{X,Y\in \mathrm{conv}(\mathcal{K})} |t_F(X) - t_F(Y)| \leq c_F L\varepsilon\,.$$

Happy birthday Jean