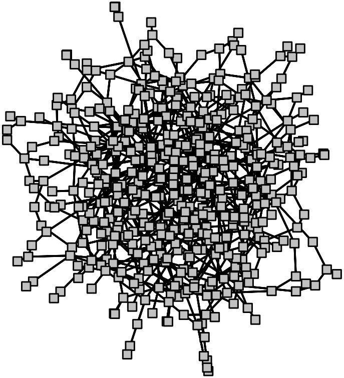
The densest subgraph of sparse random graphs



Justin Salez (Université Paris 7)
with Venkat Anantharam (UC Berkeley)

The objective method (Aldous-Steele, 2004)

The objective method (Aldous-Steele, 2004)

- **Context:** given a large interacting system (graph), one is interested in a macroscopic quantity which depends on the microscopic contribution of each particle (vertices).

The objective method (Aldous-Steele, 2004)

- **Context:** given a large interacting system (graph), one is interested in a macroscopic quantity which depends on the microscopic contribution of each particle (vertices).
- **Key assumption:** no long-range interactions, i.e. the microscopic contribution of each particle is essentially insensitive to remote perturbations of the system.

The objective method (Aldous-Steele, 2004)

- **Context:** given a large interacting system (graph), one is interested in a macroscopic quantity which depends on the microscopic contribution of each particle (vertices).
- **Key assumption:** no long-range interactions, i.e. the microscopic contribution of each particle is essentially insensitive to remote perturbations of the system.
- **Expected consequences:**

The objective method (Aldous-Steele, 2004)

- **Context:** given a large interacting system (graph), one is interested in a macroscopic quantity which depends on the microscopic contribution of each particle (vertices).
- **Key assumption:** no long-range interactions, i.e. the microscopic contribution of each particle is essentially insensitive to remote perturbations of the system.
- **Expected consequences:**
    1. efficient approximability by local distributed algorithms;

The objective method (Aldous-Steele, 2004)

- **Context:** given a large interacting system (graph), one is interested in a macroscopic quantity which depends on the microscopic contribution of each particle (vertices).
- **Key assumption:** no long-range interactions, i.e. the microscopic contribution of each particle is essentially insensitive to remote perturbations of the system.
- **Expected consequences:**
  1. efficient approximability by local distributed algorithms;
  2. existence of an infinite-volume limit.

The objective method (Aldous-Steele, 2004)

- **Context:** given a large interacting system (graph), one is interested in a macroscopic quantity which depends on the microscopic contribution of each particle (vertices).
- **Key assumption:** no long-range interactions, i.e. the microscopic contribution of each particle is essentially insensitive to remote perturbations of the system.
- **Expected consequences:**
  1. efficient approximability by local distributed algorithms;
  2. existence of an infinite-volume limit.
- **Idea:** formalize that via *local weak convergence*, and use this framework to replace the asymptotic study of large graphs by the direct analysis of their infinite-volume limits.

Local convergence around a fixed root

Local convergence around a fixed root

(G, o) : countable, locally finite, connected rooted graph

Local convergence around a fixed root

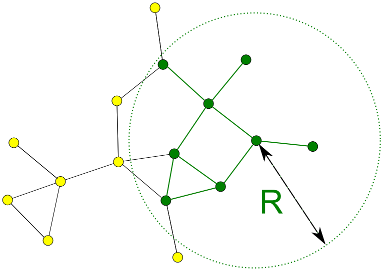
(G, o) : countable, locally finite, connected rooted graph

$[G, o]_R$: ball of radius R around o in G

Local convergence around a fixed root

(G, o) : countable, locally finite, connected rooted graph

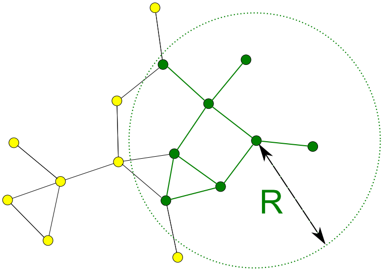
$[G, o]_R$: ball of radius R around o in G



Local convergence around a fixed root

(G, o) : countable, locally finite, connected rooted graph

$[G, o]_R$: ball of radius R around o in G

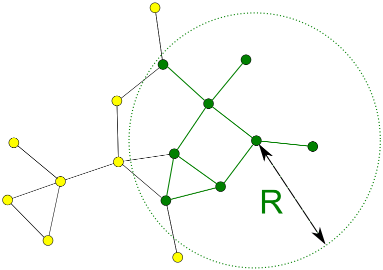


$$(G_n, o_n) \xrightarrow[n\to\infty]{} (G, o)$$

Local convergence around a fixed root

(G, o) : countable, locally finite, connected rooted graph

$[G, o]_R$: ball of radius R around o in G



$(G_n, o_n) \xrightarrow[n\to\infty]{} (G, o)$ if for each **fixed** R, there is $n_R \in \mathbb{N}$ such that

$$n \geq n_R \implies [G_n, o_n]_R \equiv [G, o]_R$$

Local weak convergence (Benjamini-Schramm, 2001)

$\mathcal{G}_\star = \{\text{locally finite connected rooted graphs}\}$.

Local weak convergence (Benjamini-Schramm, 2001)

$\mathcal{G}_\star = \{\text{locally finite connected rooted graphs}\}$.

$G = (V, E)$: finite unrooted, possibly disconnected graph.

Local weak convergence (Benjamini-Schramm, 2001)

$\mathcal{G}_\star = \{\text{locally finite connected rooted graphs}\}$.

$G = (V, E)$: finite unrooted, possibly disconnected graph.

Consider the law on $\mathcal{G}_\star$ obtained by choosing a root $o \in V$ uniformly at random, and restricting to its connected component:

Local weak convergence (Benjamini-Schramm, 2001)

$\mathcal{G}_\star = \{\text{locally finite connected rooted graphs}\}$.

$G = (V, E)$: finite unrooted, possibly disconnected graph.

Consider the law on $\mathcal{G}_\star$ obtained by choosing a root $o \in V$ uniformly at random, and restricting to its connected component:

$$\mathcal{L}_G := \frac{1}{|V|} \sum_{o \in V} \delta_{[G,o]}.$$

Local weak convergence (Benjamini-Schramm, 2001)

$\mathcal{G}_\star = \{\text{locally finite connected rooted graphs}\}$.

$G = (V, E)$: finite unrooted, possibly disconnected graph.

Consider the law on $\mathcal{G}_\star$ obtained by choosing a root $o \in V$ uniformly at random, and restricting to its connected component:

$$\mathcal{L}_G := \frac{1}{|V|} \sum_{o \in V} \delta_{[G,o]}.$$

$\mathcal{L}_G$ is an element of $\mathcal{P}(\mathcal{G}_\star) := \{\text{probability measures on } \mathcal{G}_\star\}$.

Local weak convergence (Benjamini-Schramm, 2001)

$\mathcal{G}_\star = \{\text{locally finite connected rooted graphs}\}$.

$G = (V, E)$: finite unrooted, possibly disconnected graph.

Consider the law on $\mathcal{G}_\star$ obtained by choosing a root $o \in V$ uniformly at random, and restricting to its connected component:

$$\mathcal{L}_G := \frac{1}{|V|} \sum_{o \in V} \delta_{[G,o]}.$$

$\mathcal{L}_G$ is an element of $\mathcal{P}(\mathcal{G}_\star) := \{\text{probability measures on } \mathcal{G}_\star\}$.

$\{G_n\}_{n \geq 1}$: sequence of finite graphs.

Local weak convergence (Benjamini-Schramm, 2001)

$\mathcal{G}_\star = \{\text{locally finite connected rooted graphs}\}$.

$G = (V, E)$: finite unrooted, possibly disconnected graph.

Consider the law on $\mathcal{G}_\star$ obtained by choosing a root $o \in V$ uniformly at random, and restricting to its connected component:

$$\mathcal{L}_G := \frac{1}{|V|} \sum_{o \in V} \delta_{[G,o]}.$$

$\mathcal{L}_G$ is an element of $\mathcal{P}(\mathcal{G}_\star) := \{\text{probability measures on } \mathcal{G}_\star\}$.

$\{G_n\}_{n \geq 1}$: sequence of finite graphs. If $\{\mathcal{L}_{G_n}\}_{n \geq 1}$ admits a weak limit $\mathcal{L} \in \mathcal{P}(\mathcal{G}_\star)$, then call $\mathcal{L}$ the **local weak limit** of $\{G_n\}_{n \geq 1}$.

Local weak convergence (Benjamini-Schramm, 2001)

$\mathcal{G}_\star = \{\text{locally finite connected rooted graphs}\}$.

$G = (V, E)$: finite unrooted, possibly disconnected graph.

Consider the law on $\mathcal{G}_\star$ obtained by choosing a root $o \in V$ uniformly at random, and restricting to its connected component:

$$\mathcal{L}_G := \frac{1}{|V|} \sum_{o \in V} \delta_{[G,o]}.$$

$\mathcal{L}_G$ is an element of $\mathcal{P}(\mathcal{G}_\star) := \{\text{probability measures on } \mathcal{G}_\star\}$.

$\{G_n\}_{n\geq 1}$: sequence of finite graphs. If $\{\mathcal{L}_{G_n}\}_{n\geq 1}$ admits a weak limit $\mathcal{L} \in \mathcal{P}(\mathcal{G}_\star)$, then call $\mathcal{L}$ the **local weak limit** of $\{G_n\}_{n\geq 1}$.

- $\mathcal{L}$ describes the local geometry of G_n around a random node

Examples of local weak limits

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ = dirac at $(\mathbb{Z}^d, 0)$

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ = dirac at $(\mathbb{Z}^d, 0)$
- G_n = random $d-$regular graph on n nodes

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ = dirac at $(\mathbb{Z}^d, 0)$
- G_n = random $d-$regular graph on n nodes
  $\mathcal{L}$ = dirac at the $d-$regular infinite rooted tree

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n $=$ box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ $=$ dirac at $(\mathbb{Z}^d, 0)$
- G_n $=$ random $d-$regular graph on n nodes
  $\mathcal{L}$ $=$ dirac at the $d-$regular infinite rooted tree
- G_n $=$ Erdős-Rényi graph with $p_n = \frac{c}{n}$ on n nodes

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ = dirac at $(\mathbb{Z}^d, 0)$
- G_n = random $d-$regular graph on n nodes
  $\mathcal{L}$ = dirac at the $d-$regular infinite rooted tree
- G_n = Erdős-Rényi graph with $p_n = \frac{c}{n}$ on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree Poisson(c)

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ = dirac at $(\mathbb{Z}^d, 0)$
- G_n = random $d-$regular graph on n nodes
  $\mathcal{L}$ = dirac at the $d-$regular infinite rooted tree
- G_n = Erdős-Rényi graph with $p_n = \frac{c}{n}$ on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree Poisson(c)
- G_n = random graph with degree distribution π on n nodes

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ = dirac at $(\mathbb{Z}^d, 0)$
- G_n = random $d-$regular graph on n nodes
  $\mathcal{L}$ = dirac at the $d-$regular infinite rooted tree
- G_n = Erdős-Rényi graph with $p_n = \frac{c}{n}$ on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree Poisson(c)
- G_n = random graph with degree distribution π on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree distribution π

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ = dirac at $(\mathbb{Z}^d, 0)$
- G_n = random $d-$regular graph on n nodes
  $\mathcal{L}$ = dirac at the $d-$regular infinite rooted tree
- G_n = Erdős-Rényi graph with $p_n = \frac{c}{n}$ on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree Poisson(c)
- G_n = random graph with degree distribution π on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree distribution π
- G_n = uniform random tree on n nodes

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ = dirac at $(\mathbb{Z}^d, 0)$
- G_n = random $d-$regular graph on n nodes
  $\mathcal{L}$ = dirac at the $d-$regular infinite rooted tree
- G_n = Erdős-Rényi graph with $p_n = \frac{c}{n}$ on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree Poisson(c)
- G_n = random graph with degree distribution π on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree distribution π
- G_n = uniform random tree on n nodes
  $\mathcal{L}$ = *Infinite Skeleton Tree* (Grimmett, 1980)

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ = dirac at $(\mathbb{Z}^d, 0)$
- G_n = random $d-$regular graph on n nodes
  $\mathcal{L}$ = dirac at the $d-$regular infinite rooted tree
- G_n = Erdős-Rényi graph with $p_n = \frac{c}{n}$ on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree Poisson(c)
- G_n = random graph with degree distribution π on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree distribution π
- G_n = uniform random tree on n nodes
  $\mathcal{L}$ = *Infinite Skeleton Tree* (Grimmett, 1980)
- G_n = preferential attachment graph on n nodes

Examples of local weak limits

Note: graphs must be **sparse**, i.e. $|E| \asymp |V|$

- G_n = box of size $n \times \ldots \times n$ in $\mathbb{Z}^d$
  $\mathcal{L}$ = dirac at $(\mathbb{Z}^d, 0)$
- G_n = random $d-$regular graph on n nodes
  $\mathcal{L}$ = dirac at the $d-$regular infinite rooted tree
- G_n = Erdős-Rényi graph with $p_n = \frac{c}{n}$ on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree Poisson(c)
- G_n = random graph with degree distribution π on n nodes
  $\mathcal{L}$ = law of a Galton-Watson tree with degree distribution π
- G_n = uniform random tree on n nodes
  $\mathcal{L}$ = *Infinite Skeleton Tree* (Grimmett, 1980)
- G_n = preferential attachment graph on n nodes
  $\mathcal{L}$ = *Polya-point graph* (Berger-Borgs-Chayes-Sabery, 2009)

An illustration: the nullity of large graphs

An illustration: the nullity of large graphs

$$\mu_G(\{0\}) = \frac{\dim \ker(\mathrm{A_G})}{|V|}.$$

An illustration: the nullity of large graphs

$\mu_G(\{0\}) = \frac{\dim\ker(\mathrm{A_G})}{|V|}$. Asymptotics when G is large ?

An illustration: the nullity of large graphs

$\mu_G(\{0\}) = \frac{\dim \ker(\mathrm{A_G})}{|V|}$. Asymptotics when G is large ?

Conjecture (Bauer-Golinelli 2001). For G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$,

$$\mu_{G_n}(\{0\}) \xrightarrow[n\to\infty]{} \lambda^* + e^{-c\lambda^*} + c\lambda^* e^{-c\lambda^*} - 1,$$

where $\lambda^* \in [0,1]$ is the smallest root of $\lambda = e^{-ce^{-c\lambda}}$.

An illustration: the nullity of large graphs

$\mu_G(\{0\}) = \frac{\dim \ker(A_G)}{|V|}$. Asymptotics when G is large ?

Conjecture (Bauer-Golinelli 2001). For G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$,

$$\mu_{G_n}(\{0\}) \xrightarrow[n\to\infty]{} \lambda^* + e^{-c\lambda^*} + c\lambda^* e^{-c\lambda^*} - 1,$$

where $\lambda^* \in [0,1]$ is the smallest root of $\lambda = e^{-ce^{-c\lambda}}$.

Theorem (Bordenave-Lelarge-S., 2011)

An illustration: the nullity of large graphs

$\mu_G(\{0\}) = \frac{\dim \ker(A_G)}{|V|}$. Asymptotics when G is large ?

Conjecture (Bauer-Golinelli 2001). For G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$,

$$\mu_{G_n}(\{0\}) \xrightarrow[n\to\infty]{} \lambda^* + e^{-c\lambda^*} + c\lambda^* e^{-c\lambda^*} - 1,$$

where $\lambda^* \in [0,1]$ is the smallest root of $\lambda = e^{-ce^{-c\lambda}}$.

Theorem (Bordenave-Lelarge-S., 2011)

1. $G_n \to \mathcal{L} \implies \mu_{G_n}(\{0\}) \to \mu_{\mathcal{L}}(\{0\})$.

An illustration: the nullity of large graphs

$\mu_G(\{0\}) = \frac{\dim \ker(\mathrm{A_G})}{|V|}$. Asymptotics when G is large ?

Conjecture (Bauer-Golinelli 2001). For G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$,

$$\mu_{G_n}(\{0\}) \xrightarrow[n\to\infty]{} \lambda^* + e^{-c\lambda^*} + c\lambda^* e^{-c\lambda^*} - 1,$$

where $\lambda^* \in [0,1]$ is the smallest root of $\lambda = e^{-ce^{-c\lambda}}$.

Theorem (Bordenave-Lelarge-S., 2011)

1. $G_n \to \mathcal{L} \implies \mu_{G_n}(\{0\}) \to \mu_{\mathcal{L}}(\{0\})$.
2. When $\mathcal{L} = \textsc{Galton-Watson}(\pi)$,

$$\mu_{\mathcal{L}}(\{0\}) = \min_{\lambda=\lambda^{**}} \left\{ f'(1)\lambda\lambda^* + f(1-\lambda) + f(1-\lambda^*) - 1 \right\},$$

where $f(z) = \sum_n \pi_n z^n$ and $\lambda^* = f'(1-\lambda)/f'(1)$.

Continuity with respect to local weak convergence

Continuity with respect to local weak convergence

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** only.

Continuity with respect to local weak convergence

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** only.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

Continuity with respect to local weak convergence

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** only.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

- **Algorithmic implication:** Φ is efficiently approximable via local distributed algorithms, independently of network size.

Continuity with respect to local weak convergence

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** only.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

- **Algorithmic implication:** Φ is efficiently approximable via local distributed algorithms, independently of network size.
- **Analytic implication:** Φ admits a limit along most sparse graph sequences. The distributional self-similarity of $\mathcal{L}$ may even allow for an explicit determination of $\Phi(\mathcal{L})$.

Continuity with respect to local weak convergence

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** only.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

- **Algorithmic implication:** Φ is efficiently approximable via local distributed algorithms, independently of network size.
- **Analytic implication:** Φ admits a limit along most sparse graph sequences. The distributional self-similarity of $\mathcal{L}$ may even allow for an explicit determination of $\Phi(\mathcal{L})$.
- **Examples:**

Continuity with respect to local weak convergence

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** only.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

- **Algorithmic implication:** Φ is efficiently approximable via local distributed algorithms, independently of network size.
- **Analytic implication:** Φ admits a limit along most sparse graph sequences. The distributional self-similarity of $\mathcal{L}$ may even allow for an explicit determination of $\Phi(\mathcal{L})$.
- **Examples:** number of spanning trees (Lyons, 2005),

Continuity with respect to local weak convergence

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** only.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

- **Algorithmic implication:** Φ is efficiently approximable via local distributed algorithms, independently of network size.
- **Analytic implication:** Φ admits a limit along most sparse graph sequences. The distributional self-similarity of $\mathcal{L}$ may even allow for an explicit determination of $\Phi(\mathcal{L})$.
- **Examples:** number of spanning trees (Lyons, 2005), spectrum and rank (Bordenave-Lelarge-S, 2011),

Continuity with respect to local weak convergence

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** only.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

- **Algorithmic implication:** Φ is efficiently approximable via local distributed algorithms, independently of network size.
- **Analytic implication:** Φ admits a limit along most sparse graph sequences. The distributional self-similarity of $\mathcal{L}$ may even allow for an explicit determination of $\Phi(\mathcal{L})$.
- **Examples:** number of spanning trees (Lyons, 2005), spectrum and rank (Bordenave-Lelarge-S, 2011), matching polynomial (idem, 2013),

Continuity with respect to local weak convergence

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** only.
- This can be rigorously formalized by a continuity theorem:

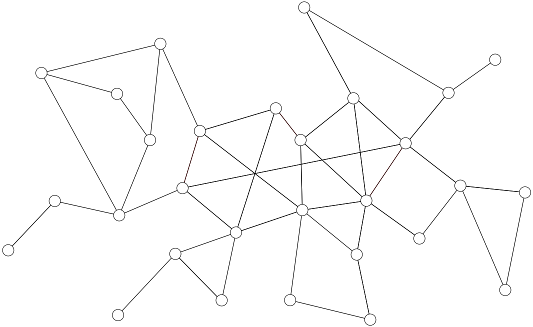
$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

- **Algorithmic implication:** Φ is efficiently approximable via local distributed algorithms, independently of network size.
- **Analytic implication:** Φ admits a limit along most sparse graph sequences. The distributional self-similarity of $\mathcal{L}$ may even allow for an explicit determination of $\Phi(\mathcal{L})$.
- **Examples:** number of spanning trees (Lyons, 2005), spectrum and rank (Bordenave-Lelarge-S, 2011), matching polynomial (idem, 2013), Ising models (Dembo-Montanari-Sun, 2013)...

The densest subgraph problem

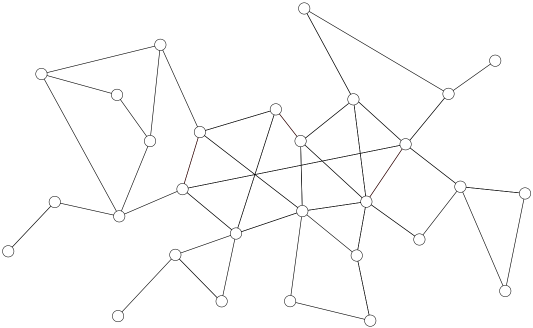
The densest subgraph problem

Fix a finite graph $G = (V, E)$



The densest subgraph problem

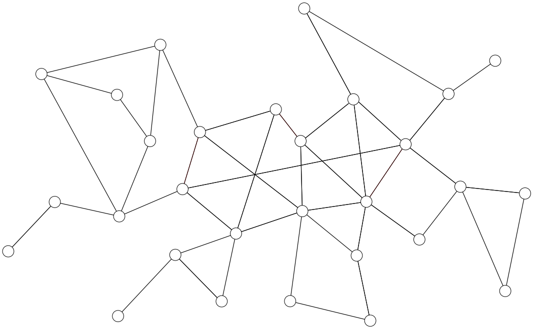
Fix a finite graph $G = (V, E)$



Densest subgraph : $H^{\star} = \mathrm{argmax}\left\{ \frac{|E(H)|}{|H|} : H \subseteq V \right\}$

The densest subgraph problem

Fix a finite graph $G = (V, E)$

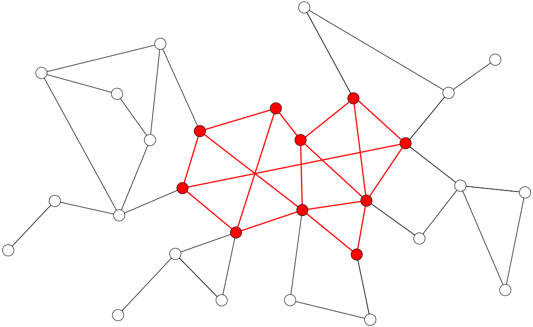


Densest subgraph : $H^\star = \operatorname{argmax}\left\{\frac{|E(H)|}{|H|} : H \subseteq V\right\}$

Maximum subgraph density : $\varrho^\star = \max\left\{\frac{|E(H)|}{|H|} : H \subseteq V\right\}$

The densest subgraph problem

Fix a finite graph $G = (V, E)$

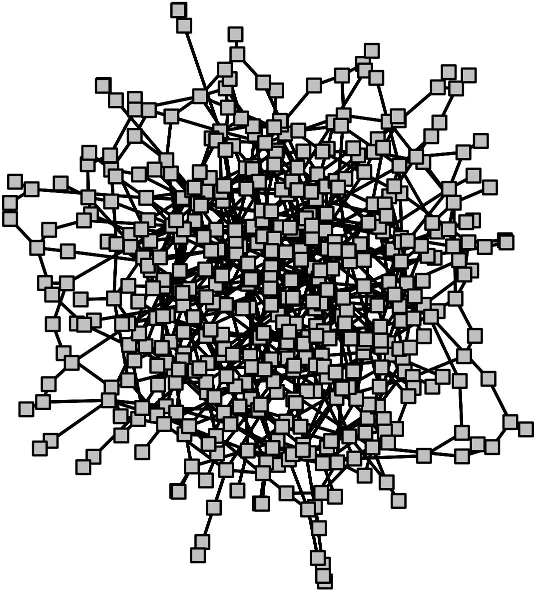


Densest subgraph : $H^\star = \operatorname{argmax}\left\{\frac{|E(H)|}{|H|} : H \subseteq V\right\}$

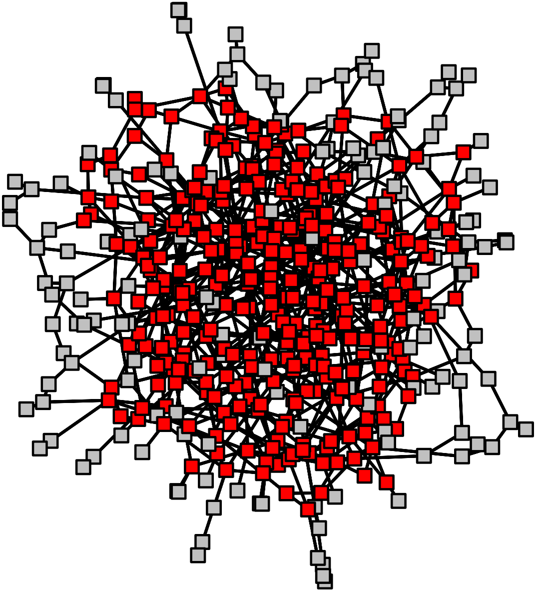
Maximum subgraph density : $\varrho^\star = \max\left\{\frac{|E(H)|}{|H|} : H \subseteq V\right\} = \frac{17}{10}$

The densest subgraph problem on large sparse graphs

The densest subgraph problem on large sparse graphs



The densest subgraph problem on large sparse graphs



Load balancing

Load balancing

An **allocation** on G is a function $\theta \colon \vec{E} \to [0,1]$ satisfying

$$\theta(i,j) + \theta(j,i) = 1$$

Load balancing

An **allocation** on G is a function $\theta\colon \vec{E} \to [0,1]$ satisfying

$$\theta(i,j) + \theta(j,i) = 1$$

The **induced load** at $i \in V$ is

$$\partial\theta(i) = \sum_{j \sim i} \theta(j,i)$$

Load balancing

An **allocation** on G is a function $\theta\colon \vec{E} \to [0,1]$ satisfying

$$\theta(i,j) + \theta(j,i) = 1$$

The **induced load** at $i \in V$ is

$$\partial\theta(i) = \sum_{j \sim i} \theta(j,i)$$

The allocation is **balanced** if for each $(i,j) \in \vec{E}$

$$\partial\theta(i) < \partial\theta(j) \implies \theta(i,j) = 0$$

From local to global optimality

From local to global optimality

Claim. For an allocation θ, the following are equivalent:

From local to global optimality

Claim. For an allocation θ, the following are equivalent:

1. θ is balanced

From local to global optimality

Claim. For an allocation θ, the following are equivalent:

1. θ is balanced
2. θ minimizes $\sum_i (\partial\theta(i))^2$.

From local to global optimality

Claim. For an allocation θ, the following are equivalent:

1. θ is balanced
2. θ minimizes $\sum_i (\partial\theta(i))^2$.
3. θ minimizes $\sum_i f(\partial\theta(i))$ for any convex function $f : \mathbb{R} \to \mathbb{R}$.

From local to global optimality

Claim. For an allocation θ, the following are equivalent:

1. θ is balanced
2. θ minimizes $\sum_i (\partial\theta(i))^2$.
3. θ minimizes $\sum_i f(\partial\theta(i))$ for any convex function $f : \mathbb{R} \to \mathbb{R}$.

Corollary 1. Balanced allocations always exist.

From local to global optimality

Claim. For an allocation θ, the following are equivalent:

1. θ is balanced
2. θ minimizes $\sum_i (\partial\theta(i))^2$.
3. θ minimizes $\sum_i f(\partial\theta(i))$ for any convex function $f\colon \mathbb{R} \to \mathbb{R}$.

Corollary 1. Balanced allocations always exist.

Corollary 2. They all induce the same loads $\partial\theta\colon V \to [0,\infty)$.

From local to global optimality

Claim. For an allocation θ, the following are equivalent:

1. θ is balanced
2. θ minimizes $\sum_i (\partial\theta(i))^2$.
3. θ minimizes $\sum_i f(\partial\theta(i))$ for any convex function $f\colon \mathbb{R} \to \mathbb{R}$.

Corollary 1. Balanced allocations always exist.

Corollary 2. They all induce the same loads $\partial\theta\colon V \to [0, \infty)$.

Corollary 3. Balanced loads solve the densest subgraph problem:

From local to global optimality

Claim. For an allocation θ, the following are equivalent:

1. θ is balanced
2. θ minimizes $\sum_i (\partial\theta(i))^2$.
3. θ minimizes $\sum_i f(\partial\theta(i))$ for any convex function $f\colon \mathbb{R} \to \mathbb{R}$.

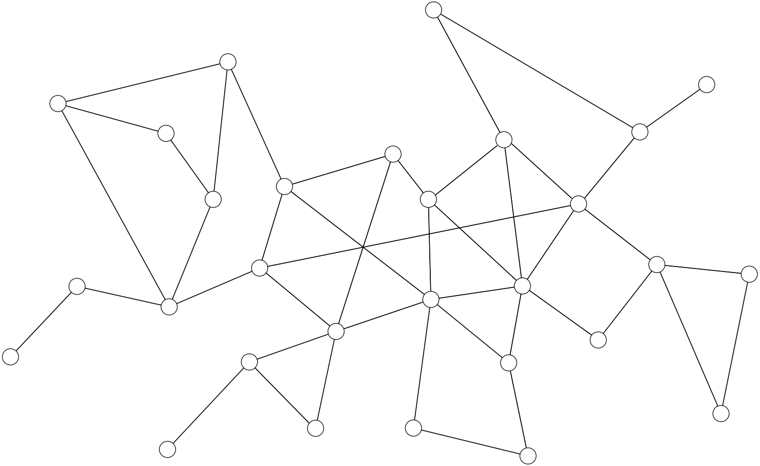
Corollary 1. Balanced allocations always exist.

Corollary 2. They all induce the same loads $\partial\theta\colon V \to [0,\infty)$.

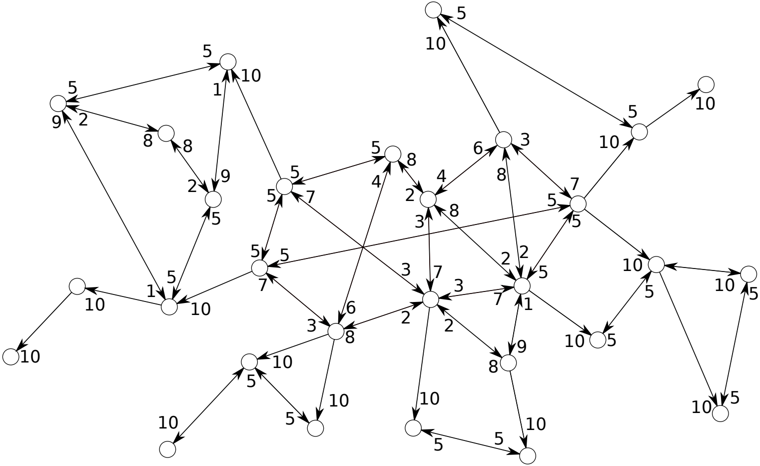
Corollary 3. Balanced loads solve the densest subgraph problem:

$$\max_{i\in V} \partial\theta(i) = \varrho^\star \quad \text{and} \quad \operatorname*{argmax}_{i\in V} \partial\theta(i) = H^\star$$

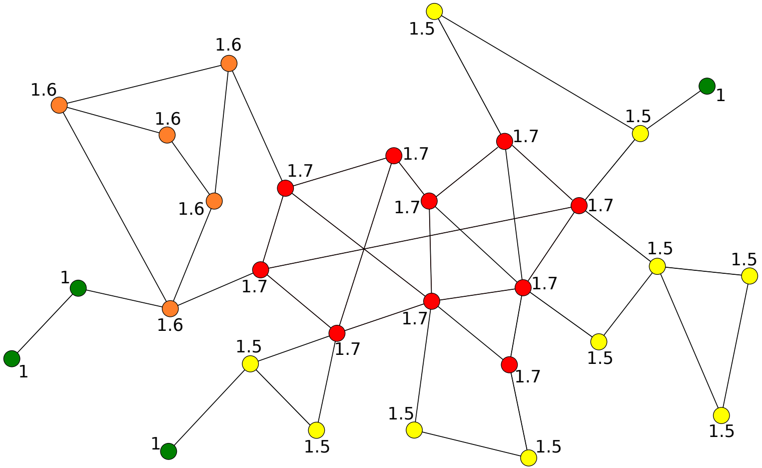
Example



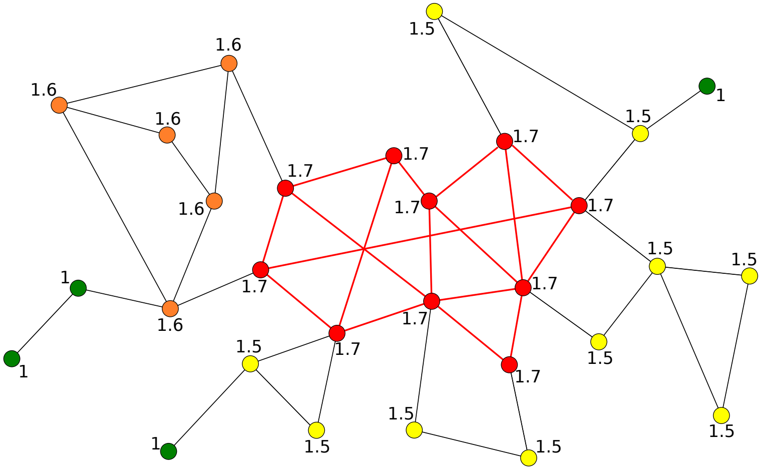
Example



Example

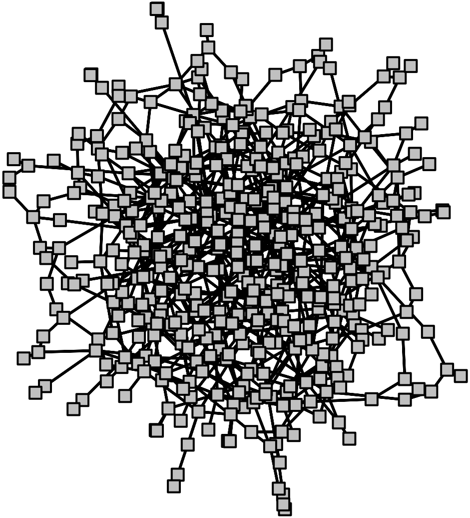

1.5
1.6
1.6
1
1.5
1.6
1.7
1.7
1.7
1.6
1.7
1.7
1.5
1.5
1
1.7
1.7
1.6
1.7
1.7
1.5
1
1.5
1.7
1.7
1.5
1.5
1.5
1
1.5
1.5

Example

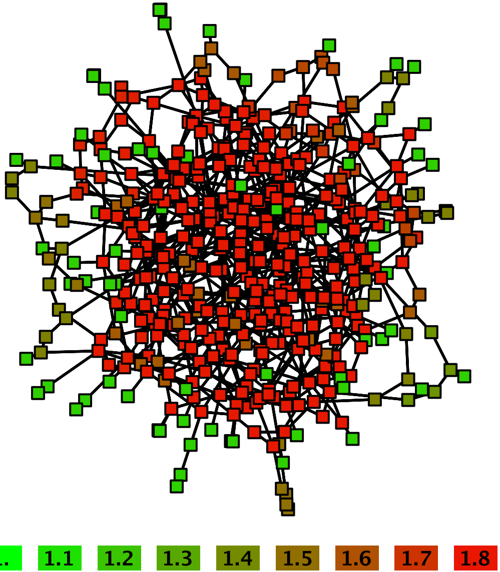

1.5
1.6
1.6
1
1.5
1.6
1.7
1.7
1.7
1.7
1.7
1.6
1.5
1.5
1
1.7
1.7
1.7
1.6
1.5
1.7
1.5
1
1.7
1.5
1.5
1
1.5
1.5

How do those ‘densities” look on a large sparse graph ?

How do those ‘densities” look on a large sparse graph ?

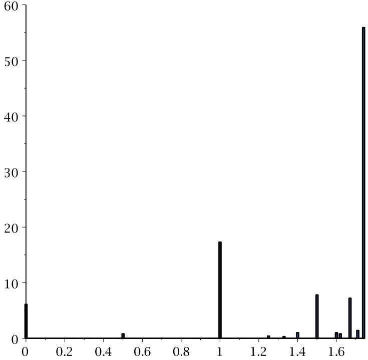


How do those ‘densities” look on a large sparse graph ?



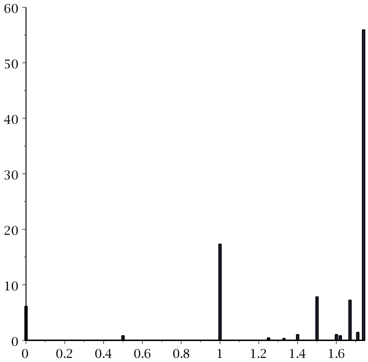
Density profile of a random graph with average degree 3

Density profile of a random graph with average degree 3

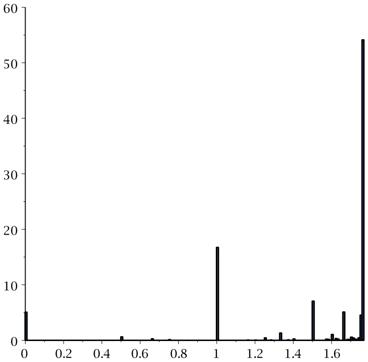


$|V| = 100$

Density profile of a random graph with average degree 3



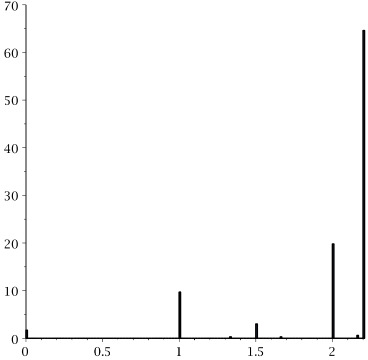
$|V| = 100$



$|V| = 10000$

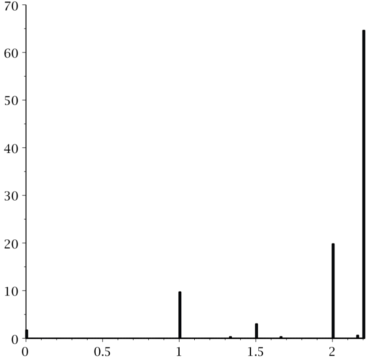
Density profile of a random graph with average degree 4

Density profile of a random graph with average degree 4

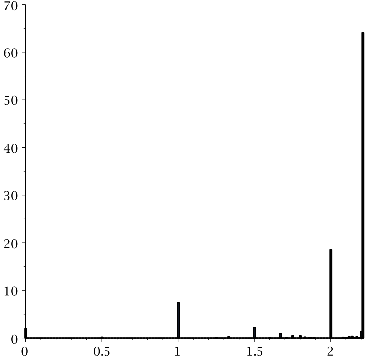


$|V| = 100$

Density profile of a random graph with average degree 4



$|V| = 100$



$|V| = 10000$

The conjecture (Hajek, 1990)

The conjecture (Hajek, 1990)

$\partial\Theta(G, o)$: load induced at o by any balanced allocation on G.

The conjecture (Hajek, 1990)

$\partial\Theta(G,o)$: load induced at o by any balanced allocation on G.

Define the **density profile** of $G=(V,E)$ as

$$\Lambda_G = \frac{1}{|V|}\sum_{o\in V}\delta_{\partial\Theta(G,o)} \in \mathcal{P}(\mathbb{R}).$$

The conjecture (Hajek, 1990)

$\partial\Theta(G, o)$: load induced at o by any balanced allocation on G.

Define the **density profile** of $G = (V, E)$ as

$$\Lambda_G = \frac{1}{|V|} \sum_{o \in V} \delta_{\partial\Theta(G,o)} \in \mathcal{P}(\mathbb{R}).$$

Conjecture: G_n Erdős-Rényi $\left(n, \frac{c}{n}\right)$; c fixed, $n \to \infty$

The conjecture (Hajek, 1990)

$\partial\Theta(G,o)$: load induced at o by any balanced allocation on G.

Define the **density profile** of $G=(V,E)$ as

$$\Lambda_G = \frac{1}{|V|}\sum_{o\in V}\delta_{\partial\Theta(G,o)} \in \mathcal{P}(\mathbb{R}).$$

Conjecture: G_n Erdős-Rényi $\left(n,\frac{c}{n}\right)$; c fixed, $n\to\infty$

1. Λ_{G_n} concentrates around a deterministic $\Lambda\in\mathcal{P}(\mathbb{R})$

The conjecture (Hajek, 1990)

$\partial\Theta(G,o)$: load induced at o by any balanced allocation on G.

Define the **density profile** of $G=(V,E)$ as

$$\Lambda_G = \frac{1}{|V|}\sum_{o\in V}\delta_{\partial\Theta(G,o)} \ \in \ \mathcal{P}(\mathbb{R}).$$

Conjecture: G_n Erdős-Rényi $\left(n,\frac{c}{n}\right)$; c fixed, $n\to\infty$

1. Λ_{G_n} concentrates around a deterministic $\Lambda\in\mathcal{P}(\mathbb{R})$
2. $\varrho^\star(G_n)\xrightarrow[n\to\infty]{\mathbb{P}}\sup\{t\in\mathbb{R}\colon \Lambda(t,+\infty)>0\}$

Result 1 : the density profile of sparse graphs

Result 1 : the density profile of sparse graphs

Theorem. Assume that $\mathcal{L}\left[\deg(G, o)\right] < \infty$.

Result 1 : the density profile of sparse graphs

Theorem. Assume that $\mathcal{L}\left[\deg(G,o)\right] < \infty$. Then,

$$G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Lambda_{G_n} \xrightarrow[n\to\infty]{\mathcal{P}(\mathbb{R})} \Lambda_{\mathcal{L}}$$

Result 1 : the density profile of sparse graphs

Theorem. Assume that $\mathcal{L}\left[\deg(G, o)\right] < \infty$. Then,

$$G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Lambda_{G_n} \xrightarrow[n\to\infty]{\mathcal{P}(\mathbb{R})} \Lambda_{\mathcal{L}}$$

where $\Lambda_{\mathcal{L}}$ is the solution to a certain optimization problem on $\mathcal{L}$.

Result 1 : the density profile of sparse graphs

Theorem. Assume that $\mathcal{L}\left[\deg(G,o)\right] < \infty$. Then,

$$G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Lambda_{G_n} \xrightarrow[n\to\infty]{\mathcal{P}(\mathbb{R})} \Lambda_{\mathcal{L}}$$

where $\Lambda_{\mathcal{L}}$ is the solution to a certain optimization problem on $\mathcal{L}$.

Specifically, the excess function $\Phi\colon t \mapsto \int_{\mathbb{R}}(x-t)^+\Lambda_{\mathcal{L}}(dx)$

Result 1 : the density profile of sparse graphs

Theorem. Assume that $\mathcal{L}\left[\deg(G,o)\right] < \infty$. Then,

$$G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Lambda_{G_n} \xrightarrow[n\to\infty]{\mathcal{P}(\mathbb{R})} \Lambda_{\mathcal{L}}$$

where $\Lambda_{\mathcal{L}}$ is the solution to a certain optimization problem on $\mathcal{L}$.

Specifically, the excess function $\Phi\colon t \mapsto \int_{\mathbb{R}}(x-t)^+\Lambda_{\mathcal{L}}(dx)$ solves

$$\Phi(t) = \max_{f\colon \mathcal{G}_\star \to [0,1]} \left\{ \frac{1}{2}\mathcal{L}\left[\sum_{i\sim o} f(G,i) \wedge f(G,o)\right] - t\mathcal{L}[f(G,o)] \right\}$$

Result 2 : maximum subgraph density of sparse graphs

Result 2 : maximum subgraph density of sparse graphs

Extend the definition of $\varrho^\star$ to local weak limits by

$$\varrho^\star(\mathcal{L}) := \sup \mathrm{ess}\, \Lambda_{\mathcal{L}} = \sup\{t\colon \Phi(t) > 0\}$$

Result 2 : maximum subgraph density of sparse graphs

Extend the definition of $\varrho^\star$ to local weak limits by

$$\varrho^\star(\mathcal{L}) := \sup \mathrm{ess}\, \Lambda_{\mathcal{L}} = \sup\{t \colon \Phi(t) > 0\}$$

In light of previous result, one expects a continuity principle :

Result 2 : maximum subgraph density of sparse graphs

Extend the definition of $\varrho^\star$ to local weak limits by

$$\varrho^\star(\mathcal{L}) := \sup\mathrm{ess}\,\Lambda_{\mathcal{L}} = \sup\{t\colon \Phi(t) > 0\}$$

In light of previous result, one expects a continuity principle :

$$G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \varrho^\star(G_n) \xrightarrow[n\to\infty]{} \varrho^\star(\mathcal{L})$$

Result 2 : maximum subgraph density of sparse graphs

Extend the definition of $\varrho^\star$ to local weak limits by

$$\varrho^\star(\mathcal{L}) := \operatorname{sup\,ess} \Lambda_{\mathcal{L}} = \sup\{t \colon \Phi(t) > 0\}$$

In light of previous result, one expects a continuity principle :

$$G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \varrho^\star(G_n) \xrightarrow[n\to\infty]{} \varrho^\star(\mathcal{L})$$

Counter-example:

Result 2 : maximum subgraph density of sparse graphs

Extend the definition of $\varrho^\star$ to local weak limits by

$$\varrho^\star(\mathcal{L}) := \sup\operatorname{ess}\Lambda_{\mathcal{L}} = \sup\{t \colon \Phi(t) > 0\}$$

In light of previous result, one expects a continuity principle :

$$G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \varrho^\star(G_n) \xrightarrow[n\to\infty]{} \varrho^\star(\mathcal{L})$$

Counter-example: adding a large but fixed clique to G_n will arbitrarily boost $\varrho^\star(G_n)$ without affecting convergence $G_n \to \mathcal{L}$.

Result 2 : maximum subgraph density of sparse graphs

Extend the definition of $\varrho^\star$ to local weak limits by

$$\varrho^\star(\mathcal{L}) := \sup \operatorname{ess} \Lambda_{\mathcal{L}} = \sup\{t \colon \Phi(t) > 0\}$$

In light of previous result, one expects a continuity principle :

$$G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \varrho^\star(G_n) \xrightarrow[n\to\infty]{} \varrho^\star(\mathcal{L})$$

Counter-example: adding a large but fixed clique to G_n will arbitrarily boost $\varrho^\star(G_n)$ without affecting convergence $G_n \to \mathcal{L}$.

Theorem. G_n uniform with degree distribution $\{\pi_k\}_{k\in\mathbb{N}}$.

Result 2 : maximum subgraph density of sparse graphs

Extend the definition of $\varrho^\star$ to local weak limits by

$$\varrho^\star(\mathcal{L}) := \operatorname{sup\,ess} \Lambda_{\mathcal{L}} = \sup\{t \colon \Phi(t) > 0\}$$

In light of previous result, one expects a continuity principle :

$$G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \varrho^\star(G_n) \xrightarrow[n\to\infty]{} \varrho^\star(\mathcal{L})$$

Counter-example: adding a large but fixed clique to G_n will arbitrarily boost $\varrho^\star(G_n)$ without affecting convergence $G_n \to \mathcal{L}$.

Theorem. G_n uniform with degree distribution $\{\pi_k\}_{k\in\mathbb{N}}$. Assume degrees have light tail, i.e. $\limsup_{k\to\infty} \pi_k^{1/k} < 1$.

Result 2 : maximum subgraph density of sparse graphs

Extend the definition of $\varrho^\star$ to local weak limits by

$$\varrho^\star(\mathcal{L}) := \operatorname{sup\,ess} \Lambda_{\mathcal{L}} = \sup\{t \colon \Phi(t) > 0\}$$

In light of previous result, one expects a continuity principle :

$$G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \varrho^\star(G_n) \xrightarrow[n\to\infty]{} \varrho^\star(\mathcal{L})$$

Counter-example: adding a large but fixed clique to G_n will arbitrarily boost $\varrho^\star(G_n)$ without affecting convergence $G_n \to \mathcal{L}$.

Theorem. G_n uniform with degree distribution $\{\pi_k\}_{k\in\mathbb{N}}$. Assume degrees have light tail, i.e. $\limsup_{k\to\infty} \pi_k^{1/k} < 1$. Then,

$$\varrho^\star(G_n) \xrightarrow[n\to\infty]{} \varrho^\star(\mathcal{L}), \text{ where } \mathcal{L} = \text{Galton-Watson}(\pi).$$

Result 3 : the case of Galton-Watson trees

Result 3 : the case of Galton-Watson trees

Theorem. In the case where $\mathcal{L} = \textsc{Galton-Watson}(\pi)$,

$$\Phi(t) = \max_{Q \in \mathcal{P}([0,1])} \left\{ \frac{\mathbb{E}[D]}{2} \mathbb{P}\left(\xi_1 + \xi_2 > 1\right) - t\mathbb{P}\left(\xi_1 + \cdots + \xi_D > t\right) \right\}$$

Result 3 : the case of Galton-Watson trees

Theorem. In the case where $\mathcal{L} = \textsc{Galton-Watson}(\pi)$,

$$\Phi(t) = \max_{Q \in \mathcal{P}([0,1])} \left\{ \frac{\mathbb{E}[D]}{2} \mathbb{P}\left(\xi_1 + \xi_2 > 1\right) - t\mathbb{P}\left(\xi_1 + \cdots + \xi_D > t\right) \right\}$$

where $D \sim \pi$ and $\{\xi_k\}_{k \geq 1}$ are IID with law Q, independent of D.

Result 3 : the case of Galton-Watson trees

Theorem. In the case where $\mathcal{L} = \textsc{Galton-Watson}(\pi)$,

$$\Phi(t) = \max_{Q \in \mathcal{P}([0,1])} \left\{ \frac{\mathbb{E}[D]}{2} \mathbb{P}\left(\xi_1 + \xi_2 > 1\right) - t\mathbb{P}\left(\xi_1 + \cdots + \xi_D > t\right) \right\}$$

where $D \sim \pi$ and $\{\xi_k\}_{k \geq 1}$ are IID with law Q, independent of D.

The maximum is over all choices of $Q \in \mathcal{P}([0,1])$ such that

Result 3 : the case of Galton-Watson trees

Theorem. In the case where $\mathcal{L} = \textsc{Galton-Watson}(\pi)$,

$$\Phi(t) = \max_{Q \in \mathcal{P}([0,1])} \left\{ \frac{\mathbb{E}[D]}{2} \mathbb{P}\left(\xi_1 + \xi_2 > 1\right) - t\mathbb{P}\left(\xi_1 + \cdots + \xi_D > t\right) \right\}$$

where $D \sim \pi$ and $\{\xi_k\}_{k \geq 1}$ are IID with law Q, independent of D.

The maximum is over all choices of $Q \in \mathcal{P}([0,1])$ such that

$$\xi \overset{d}{=} [1 - t + \xi_1 + \cdots + \xi_D]_0^1$$

Result 3 : the case of Galton-Watson trees

Theorem. In the case where $\mathcal{L} = \textsc{Galton-Watson}(\pi)$,

$$\Phi(t) = \max_{Q \in \mathcal{P}([0,1])} \left\{ \frac{\mathbb{E}[D]}{2} \mathbb{P}\left(\xi_1 + \xi_2 > 1\right) - t\mathbb{P}\left(\xi_1 + \cdots + \xi_D > t\right) \right\}$$

where $D \sim \pi$ and $\{\xi_k\}_{k \geq 1}$ are IID with law Q, independent of D.

The maximum is over all choices of $Q \in \mathcal{P}([0,1])$ such that

$$\xi \overset{d}{=} [1 - t + \xi_1 + \cdots + \xi_D]_0^1$$

where $[\bullet]_0^1$ denotes projection onto $[0,1]$: $[x]_0^1 = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \in [0,1] \\ 1 & \text{if } x > 1 \end{cases}$

Result 3 : the case of Galton-Watson trees

Theorem. In the case where $\mathcal{L} = \textsc{Galton-Watson}(\pi)$,

$$\Phi(t) = \max_{Q \in \mathcal{P}([0,1])} \left\{ \frac{\mathbb{E}[D]}{2} \mathbb{P}\left(\xi_1 + \xi_2 > 1\right) - t\mathbb{P}\left(\xi_1 + \cdots + \xi_D > t\right) \right\}$$

where $D \sim \pi$ and $\{\xi_k\}_{k \geq 1}$ are IID with law Q, independent of D.

The maximum is over all choices of $Q \in \mathcal{P}([0,1])$ such that

$$\xi \stackrel{d}{=} [1 - t + \xi_1 + \cdots + \xi_D]_0^1$$

where $[\bullet]_0^1$ denotes projection onto $[0,1]$: $[x]_0^1 = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \in [0,1] \\ 1 & \text{if } x > 1 \end{cases}$

Distributional fixed-point equation : can be solved numerically.

An explicit formula

An explicit formula

G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$

An explicit formula

G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$ $k \in \mathbb{N}^*$ fixed

An explicit formula

G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$ $\qquad$ $k \in \mathbb{N}^*$ fixed

Question: does G_n contain a $k-$dense subgraph?

An explicit formula

G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$ $k \in \mathbb{N}^*$ fixed

Question: does G_n contain a $k-$dense subgraph?

Define $f_k(x) = e^x - \left(1 + x + \cdots + \frac{x^k}{k!}\right)$

An explicit formula

G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$ $k \in \mathbb{N}^*$ fixed

Question: does G_n contain a $k-$dense subgraph?

Define $f_k(x) = e^x - \left(1 + x + \cdots + \frac{x^k}{k!}\right)$

Set $c_* = \frac{xe^x}{f_{k-1}(x)}$, where x unique solution to $\frac{xf_{k-1}(x)}{f_k(x)} = 2k$.

An explicit formula

G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$ $\qquad$ $k \in \mathbb{N}^*$ fixed

Question: does G_n contain a $k-$dense subgraph?

Define $f_k(x) = e^x - \left(1 + x + \cdots + \frac{x^k}{k!}\right)$

Set $c_* = \frac{xe^x}{f_{k-1}(x)}$, where x unique solution to $\frac{xf_{k-1}(x)}{f_k(x)} = 2k$.

Theorem. With probability tending to one as $n \to \infty$,

An explicit formula

G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$ $k \in \mathbb{N}^*$ fixed

Question: does G_n contain a $k-$dense subgraph?

Define $f_k(x) = e^x - \left(1 + x + \cdots + \frac{x^k}{k!}\right)$

Set $c_* = \frac{xe^x}{f_{k-1}(x)}$, where x unique solution to $\frac{xf_{k-1}(x)}{f_k(x)} = 2k$.

Theorem. With probability tending to one as $n \to \infty$,

- If $c < c_*$ then G_n does not contain a $k-$dense subgraph

An explicit formula

G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$ $\qquad$ $k \in \mathbb{N}^*$ fixed

Question: does G_n contain a $k-$dense subgraph?

Define $f_k(x) = e^x - \left(1 + x + \cdots + \frac{x^k}{k!}\right)$

Set $c_* = \frac{xe^x}{f_{k-1}(x)}$, where x unique solution to $\frac{xf_{k-1}(x)}{f_k(x)} = 2k$.

Theorem. With probability tending to one as $n \to \infty$,

- If $c < c_*$ then G_n does not contain a $k-$dense subgraph
- If $c > c_*$ then G_n contains a $k-$dense subgraph

An explicit formula

G_n : Erdős-Rényi $\left(n, \frac{c}{n}\right)$ $k \in \mathbb{N}^*$ fixed

Question: does G_n contain a $k-$dense subgraph?

Define $f_k(x) = e^x - \left(1 + x + \cdots + \frac{x^k}{k!}\right)$

Set $c_* = \frac{xe^x}{f_{k-1}(x)}$, where x unique solution to $\frac{xf_{k-1}(x)}{f_k(x)} = 2k$.

Theorem. With probability tending to one as $n \to \infty$,

- If $c < c_*$ then G_n does not contain a $k-$dense subgraph
- If $c > c_*$ then G_n contains a $k-$dense subgraph

k	2	3	4	5	6	7	8	9	10
c_*	3.59	5.76	7.84	9.90	11.93	13.95	15.97	17.98	19.98

A few words on the proof

A few words on the proof

Microscopic contribution:

A few words on the proof

Microscopic contribution: $\partial\Theta(G, o)$

A few words on the proof

Microscopic contribution: $\partial\Theta(G, o)$

Hope: $\partial\Theta(G, o)$ is insensitive to what lies far away from o:

A few words on the proof

Microscopic contribution: $\partial\Theta(G,o)$

Hope: $\partial\Theta(G,o)$ is insensitive to what lies far away from o:

$$[G,o]_R \equiv [G',o']_R \implies \left|\partial\Theta(G,o) - \partial\Theta(G',o')\right| \le f(R),$$

where $f(R) \to 0$ as $R \to \infty$.

A few words on the proof

Microscopic contribution: $\partial\Theta(G,o)$

Hope: $\partial\Theta(G,o)$ is insensitive to what lies far away from o:

$$[G,o]_R \equiv [G',o']_R \implies \left|\partial\Theta(G,o)-\partial\Theta(G',o')\right| \le f(R),$$

where $f(R) \to 0$ as $R \to \infty$.

Counter-example: let G be a $d-$regular graph with girth $> R$

A few words on the proof

Microscopic contribution: $\partial\Theta(G,o)$

Hope: $\partial\Theta(G,o)$ is insensitive to what lies far away from o:

$$[G,o]_R \equiv [G',o']_R \implies \left|\partial\Theta(G,o) - \partial\Theta(G',o')\right| \leq f(R),$$

where $f(R) \to 0$ as $R \to \infty$.

Counter-example: let G be a $d-$regular graph with girth $> R$

- $\partial\Theta(G,o) = \frac{d}{2}$

A few words on the proof

Microscopic contribution: $\partial\Theta(G,o)$

Hope: $\partial\Theta(G,o)$ is insensitive to what lies far away from o:

$$[G,o]_R \equiv [G',o']_R \implies \left|\partial\Theta(G,o) - \partial\Theta(G',o')\right| \leq f(R),$$

where $f(R) \to 0$ as $R \to \infty$.

Counter-example: let G be a $d-$regular graph with girth $> R$

- $\partial\Theta(G,o) = \frac{d}{2}$
- $[G,o]_R$ is a tree

A few words on the proof

Microscopic contribution: $\partial\Theta(G,o)$

Hope: $\partial\Theta(G,o)$ is insensitive to what lies far away from o:

$$[G,o]_R \equiv [G',o']_R \implies \left|\partial\Theta(G,o) - \partial\Theta(G',o')\right| \le f(R),$$

where $f(R) \to 0$ as $R \to \infty$.

Counter-example: let G be a $d-$regular graph with girth $> R$

- $\partial\Theta(G,o) = \frac{d}{2}$
- $[G,o]_R$ is a tree
- $\partial\Theta \le \varrho^\star < 1$ on any tree

A few words on the proof

Microscopic contribution: $\partial\Theta(G,o)$

Hope: $\partial\Theta(G,o)$ is insensitive to what lies far away from o:

$$[G,o]_R \equiv [G',o']_R \implies \left|\partial\Theta(G,o) - \partial\Theta(G',o')\right| \leq f(R),$$

where $f(R) \to 0$ as $R \to \infty$.

Counter-example: let G be a $d-$regular graph with girth $> R$

- $\partial\Theta(G,o) = \frac{d}{2}$
- $[G,o]_R$ is a tree
- $\partial\Theta \leq \varrho^\star < 1$ on any tree

▶ **Balanced loads exhibit long-range dependences !**

Solution : relaxed load balancing

Solution : relaxed load balancing

$\varepsilon > 0$: perturbative parameter

Solution : relaxed load balancing

$\varepsilon > 0$: perturbative parameter

Definition. An allocation θ on $G = (V, E)$ is $\varepsilon-$balanced if

$$\theta(i,j) = \left[\frac{1}{2} + \frac{\partial\theta(i) - \partial\theta(j)}{2\varepsilon}\right]_0^1$$

Solution : relaxed load balancing

$\varepsilon > 0$: perturbative parameter

Definition. An allocation θ on $G = (V, E)$ is $\varepsilon-$balanced if

$$\theta(i,j) = \left[\frac{1}{2} + \frac{\partial\theta(i) - \partial\theta(j)}{2\varepsilon}\right]_0^1$$

In particular, $\partial\theta(i) \leq \partial\theta(j) - \varepsilon \implies \theta(i,j) = 0$.

Solution : relaxed load balancing

$\varepsilon > 0$: perturbative parameter

Definition. An allocation θ on $G = (V, E)$ is $\varepsilon-$balanced if

$$\theta(i,j) = \left[\frac{1}{2} + \frac{\partial\theta(i) - \partial\theta(j)}{2\varepsilon}\right]_0^1$$

In particular, $\partial\theta(i) \leq \partial\theta(j) - \varepsilon \implies \theta(i,j) = 0$.

Claim 1. There exists a unique $\varepsilon-$balanced allocation Θ_ε.

Solution : relaxed load balancing

$\varepsilon > 0$: perturbative parameter

Definition. An allocation θ on $G = (V, E)$ is $\varepsilon-$balanced if

$$\theta(i,j) = \left[\frac{1}{2} + \frac{\partial\theta(i) - \partial\theta(j)}{2\varepsilon}\right]_0^1$$

In particular, $\partial\theta(i) \leq \partial\theta(j) - \varepsilon \implies \theta(i,j) = 0$.

Claim 1. There exists a unique $\varepsilon-$balanced allocation Θ_ε.

Claim 2. If $[G, o]_R \equiv [G', o']_R$, then

Solution : relaxed load balancing

$\varepsilon > 0$: perturbative parameter

Definition. An allocation θ on $G = (V, E)$ is $\varepsilon-$balanced if

$$\theta(i,j) = \left[\frac{1}{2} + \frac{\partial\theta(i) - \partial\theta(j)}{2\varepsilon}\right]_0^1$$

In particular, $\partial\theta(i) \leq \partial\theta(j) - \varepsilon \implies \theta(i,j) = 0$.

Claim 1. There exists a unique $\varepsilon-$balanced allocation Θ_ε.

Claim 2. If $[G, o]_R \equiv [G', o']_R$, then

$$\left|\partial\Theta_\varepsilon(G, o) - \partial\Theta_\varepsilon(G', o')\right| \leq \Delta\left(1 + \frac{2\varepsilon}{\Delta}\right)^{-R}.$$

Solution : relaxed load balancing

$\varepsilon > 0$: perturbative parameter

Definition. An allocation θ on $G = (V, E)$ is $\varepsilon-$balanced if

$$\theta(i,j) = \left[\frac{1}{2} + \frac{\partial\theta(i) - \partial\theta(j)}{2\varepsilon}\right]_0^1$$

In particular, $\partial\theta(i) \leq \partial\theta(j) - \varepsilon \implies \theta(i,j) = 0$.

Claim 1. There exists a unique $\varepsilon-$balanced allocation Θ_ε.

Claim 2. If $[G, o]_R \equiv [G', o']_R$, then

$$\left|\partial\Theta_\varepsilon(G, o) - \partial\Theta_\varepsilon(G', o')\right| \leq \Delta \left(1 + \frac{2\varepsilon}{\Delta}\right)^{-R}.$$

Corollary. Θ_ε extends continuously to infinite graphs !

Proof outline

Proof outline

Assume $G_n \xrightarrow[n\to\infty]{\text{loc.}} \mathcal{L}$ with $\int_{\mathcal{G}_\star} \deg \, d\mathcal{L} < \infty$.

Proof outline

Assume $G_n \xrightarrow[n\to\infty]{\text{loc.}} \mathcal{L}$ with $\int_{\mathcal{G}_\star} \deg d\mathcal{L} < \infty$.

Consider a test function $\psi \colon \mathbb{R} \to \mathbb{R}$ (bounded, Lipschitz)

Proof outline

Assume $G_n \xrightarrow[n\to\infty]{\text{loc.}} \mathcal{L}$ with $\int_{\mathcal{G}_\star} \deg d\mathcal{L} < \infty$.

Consider a test function $\psi \colon \mathbb{R} \to \mathbb{R}$ (bounded, Lipschitz)

$$\frac{1}{|V_n|} \sum_{o \in V_n} \psi\left(\partial\Theta(G_n, o)\right) \quad \xrightarrow[n\to\infty]{???} \quad \int_{\mathcal{G}_\star} (\psi \circ \partial\Theta) d\mathcal{L}$$

Proof outline

Assume $G_n \xrightarrow[n\to\infty]{\text{loc.}} \mathcal{L}$ with $\int_{\mathcal{G}_\star} \deg d\mathcal{L} < \infty$.

Consider a test function $\psi\colon \mathbb{R} \to \mathbb{R}$ (bounded, Lipschitz)

$$\frac{1}{|V_n|}\sum_{o\in V_n} \psi\left(\partial\Theta(G_n,o)\right) \quad \xrightarrow[n\to\infty]{???} \quad \int_{\mathcal{G}_\star} (\psi \circ \partial\Theta) d\mathcal{L}$$

$$\frac{1}{|V_n|}\sum_{o\in V_n} \psi\left(\partial\Theta_\varepsilon(G_n,o)\right) \quad \xrightarrow[n\to\infty]{} \quad \int_{\mathcal{G}_\star} (\psi \circ \partial\Theta_\varepsilon)\, d\mathcal{L}$$

Proof outline

Assume $G_n \xrightarrow[n\to\infty]{\text{loc.}} \mathcal{L}$ with $\int_{\mathcal{G}_\star} \deg d\mathcal{L} < \infty$.

Consider a test function $\psi\colon \mathbb{R} \to \mathbb{R}$ (bounded, Lipschitz)

$$\begin{tikzcd}
\frac{1}{|V_n|} \sum_{o \in V_n} \psi\left(\partial\Theta(G_n, o)\right) \arrow[r, "???", "n\to\infty"'] & \int_{\mathcal{G}_\star} (\psi \circ \partial\Theta) d\mathcal{L} \\
\frac{1}{|V_n|} \sum_{o \in V_n} \psi\left(\partial\Theta_\varepsilon(G_n, o)\right) \arrow[r, "n\to\infty"'] & \int_{\mathcal{G}_\star} (\psi \circ \partial\Theta_\varepsilon)\, d\mathcal{L} \arrow[u, "\varepsilon \to 0"']
\end{tikzcd}$$

Proof outline

Assume $G_n \xrightarrow[n\to\infty]{\text{loc.}} \mathcal{L}$ with $\int_{\mathcal{G}_\star} \deg d\mathcal{L} < \infty$.

Consider a test function $\psi\colon \mathbb{R} \to \mathbb{R}$ (bounded, Lipschitz)

$$\begin{tikzcd}
\frac{1}{|V_n|}\sum_{o\in V_n} \psi\left(\partial\Theta(G_n,o)\right) \arrow[r, "n\to\infty"'] & \int_{\mathcal{G}_\star} (\psi\circ\partial\Theta) d\mathcal{L} \\
\frac{1}{|V_n|}\sum_{o\in V_n} \psi\left(\partial\Theta_\varepsilon(G_n,o)\right) \arrow[r, "n\to\infty"'] & \int_{\mathcal{G}_\star} (\psi\circ\partial\Theta_\varepsilon)\, d\mathcal{L} \arrow[u, "\varepsilon\to 0"']
\end{tikzcd}$$

Conclusion

Conclusion

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** of the graph.

Conclusion

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** of the graph.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

Conclusion

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** of the graph.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

- **Algorithmic implication:** Φ is efficiently approximable via local distributed algorithms, independently of network size.

Conclusion

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** of the graph.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

- **Algorithmic implication:** Φ is efficiently approximable via local distributed algorithms, independently of network size.
- **Analytic implication:** Φ admits a limit along most sparse graph sequences. The distributional self-similarity of $\mathcal{L}$ may sometimes even allow for an explicit determination of $\Phi(\mathcal{L})$.

Conclusion

- In the sparse regime, many important graph parameters Φ are essentially determined by the **local geometry** of the graph.
- This can be rigorously formalized by a continuity theorem:

$$\boxed{G_n \xrightarrow[n\to\infty]{loc.} \mathcal{L} \qquad \Longrightarrow \qquad \Phi(G_n) \xrightarrow[n\to\infty]{} \Phi(\mathcal{L})}$$

- **Algorithmic implication:** Φ is efficiently approximable via local distributed algorithms, independently of network size.
- **Analytic implication:** Φ admits a limit along most sparse graph sequences. The distributional self-similarity of $\mathcal{L}$ may sometimes even allow for an explicit determination of $\Phi(\mathcal{L})$.
- **Many examples:** spanning trees, spectrum and rank, matching polynomial, Ising models, dense subgraphs...

Thank you !