The Planted Spanning Tree Problem

Jiaming Xu

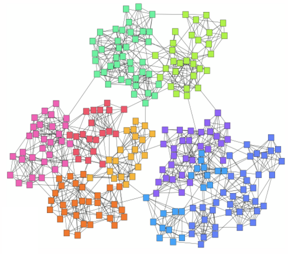
The Fuqua School of Business
Duke University

Joint work with
Mehrdad Moharrami (UIowa) and Cristopher Moore (Santa Fe Institute)

June 30, 2025
Applied Probability Society Conference

Network model with planted structure

Question: How to recover latent structure from noisy network data?



Classical examples

- Community detection under Stochastic Block Model
- Recovery of planted clique in Erdős-Rényi graphs

An expanding zoo of planted subgraph problems...

- Planted bipartite matching [Chertkov-Kroc-Krzakala-Vergassola-Zdeborová '10, Moharrami-Moore-X. '21, Ding-Wu-X.-Yang ' 23]
- Planted Hamiltonian cycle problem (TSP) [Bagaria-Ding-Tse-Wu-X. '20]
- Planted trees [Massoulié-Stephan-Towsley '18]
- Planted k-factors [Sicuro-Zdeborová '20, Gaudio-Sandon-X.-Yang '25]
- Planted k-nearest-neighbor graph [Ding-Wu-X.-Yang '21]
- Planted dense cycles [Mao-Wein-Zhang '23]
- Planted general subgraphs [Mossel-Niles-Weed-Sohn-Sun-Zadik '23, Lee-Pernice-Rajaraman-Zadik '25]

An expanding zoo of planted subgraph problems...

- Planted bipartite matching [Chertkov-Kroc-Krzakala-Vergassola-Zdeborová '10, Moharrami-Moore-X. '21, Ding-Wu-X.-Yang ' 23]
- Planted Hamiltonian cycle problem (TSP) [Bagaria-Ding-Tse-Wu-X. '20]
- Planted trees [Massoulié-Stephan-Towsley '18]
- Planted k-factors [Sicuro-Zdeborová '20, Gaudio-Sandon-X.-Yang '25]
- Planted k-nearest-neighbor graph [Ding-Wu-X.-Yang '21]
- Planted dense cycles [Mao-Wein-Zhang '23]
- Planted general subgraphs [Mossel-Niles-Weed-Sohn-Sun-Zadik '23, Lee-Pernice-Rajaraman-Zadik '25]

Many fascinating results on phase transitions, statistical vs. computational gaps, optimal algorithms; however, characterizing exact value of asymptotic overlap remains formidable mathematical challenge

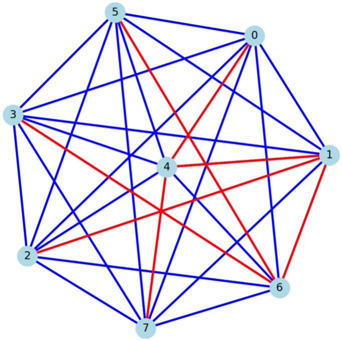
An expanding zoo of planted subgraph problems...

- Planted bipartite matching [Chertkov-Kroc-Krzakala-Vergassola-Zdeborová '10, Moharrami-Moore-X. '21, Ding-Wu-X.-Yang ' 23]
- Planted Hamiltonian cycle problem (TSP) [Bagaria-Ding-Tse-Wu-X. '20]
- Planted trees [Massoulié-Stephan-Towsley '18]
- Planted k-factors [Sicuro-Zdeborová '20, Gaudio-Sandon-X.-Yang '25]
- Planted k-nearest-neighbor graph [Ding-Wu-X.-Yang '21]
- Planted dense cycles [Mao-Wein-Zhang '23]
- Planted general subgraphs [Mossel-Niles-Weed-Sohn-Sun-Zadik '23, Lee-Pernice-Rajaraman-Zadik '25]

Many fascinating results on phase transitions, statistical vs. computational gaps, optimal algorithms; however, characterizing exact value of asymptotic overlap remains formidable mathematical challenge

Today: Planted spanning tree model $\rightarrow$ Exact overlap characterization via local weak convergence theory

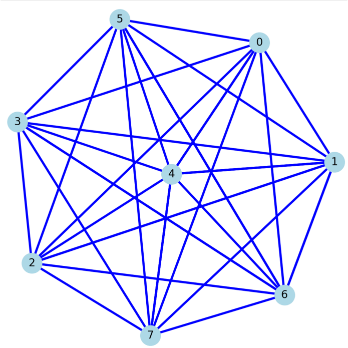
The planted spanning tree model



- A complete graph on n vertices
- A uniform spanning tree T^*
- Non-negative edge weight

$$W_e \overset{\text{ind.}}{\sim} \begin{cases} P & e \in T^* \\ Q_n & e \notin T^* \end{cases}$$

The planted spanning tree model

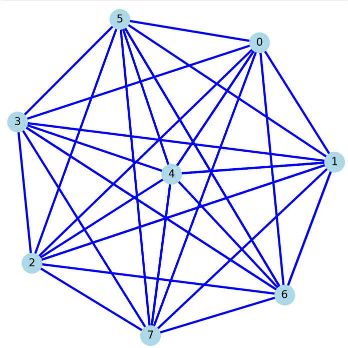


- A complete graph on n vertices
- A uniform spanning tree T^*
- Non-negative edge weight

$$W_e \overset{\text{ind.}}{\sim} \begin{cases} \color{red}{P} & e \in T^* \\ \color{blue}{Q_n} & e \notin T^* \end{cases}$$

- Goal: recover T^* from W

The planted spanning tree model



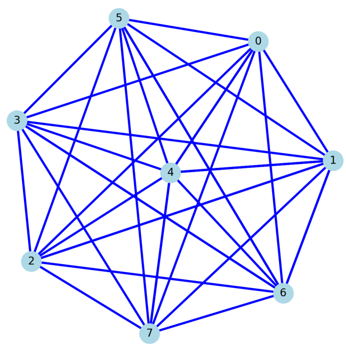
- A complete graph on n vertices
- A uniform spanning tree T^*
- Non-negative edge weight

$$W_e \overset{\text{ind.}}{\sim} \begin{cases} P & e \in T^* \\ Q_n & e \notin T^* \end{cases}$$

- Goal: recover T^* from W

- Assume P is a fixed distribution and $\lim_{n\to\infty} nQ_n'(0) = 1$

The planted spanning tree model



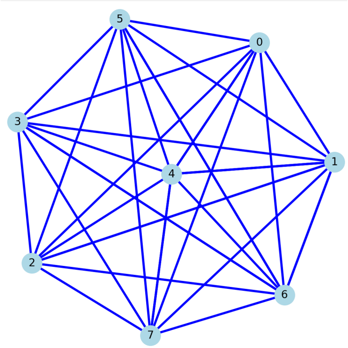
- A complete graph on n vertices
- A uniform spanning tree T^*
- Non-negative edge weight

$$W_e \overset{\text{ind.}}{\sim} \begin{cases} {\color{red}P} & e \in T^* \\ {\color{blue}Q_n} & e \notin T^* \end{cases}$$

- Goal: recover T^* from W

- Assume P is a fixed distribution and $\lim_{n\to\infty} nQ_n'(0) = 1$
- When $P = \text{Exp}(1/\mu)$ and $Q_n = \text{Exp}(1/n)$ (mean μ vs n), min-weight spanning tree $T_{\min}$ is Maximum Likelihood Estimator

The planted spanning tree model



- A complete graph on n vertices
- A uniform spanning tree T^*
- Non-negative edge weight

$$W_e \overset{\text{ind.}}{\sim} \begin{cases} P & e \in T^* \\ Q_n & e \notin T^* \end{cases}$$

- Goal: recover T^* from W

- Assume P is a fixed distribution and $\lim_{n\to\infty} nQ_n'(0) = 1$
- When $P = \mathrm{Exp}(1/\mu)$ and $Q_n = \mathrm{Exp}(1/n)$ (mean μ vs n), min-weight spanning tree $T_{\min}$ is Maximum Likelihood Estimator
- How much does $T_{\min}$ have in common with T^*?

Main result: Asymptotic overlap

Theorem (Moharrami-Moore-X. '25)

Let F denote the CDF of planted weight distribution P. Then

$$\lim_{n\to\infty} \frac{1}{n-1} \mathbb{E}[|T_{\min} \cap T^*|] = \int_0^\infty (1 - p_U(s) p_B(s))\, dF(s)$$

Main result: Asymptotic overlap

Theorem (Moharrami-Moore-X. '25)

Let F denote the CDF of planted weight distribution P. Then

$$\lim_{n\to\infty} \frac{1}{n-1} \mathbb{E}[|T_{\min} \cap T^*|] = \int_0^\infty (1 - p_U(s)p_B(s))\, dF(s)$$

where $p_U(s)$ and $p_B(s)$ are the largest fixed point of

$$1 - p_U(s) = \exp\left(-sp_U(s) - p_B(s)\right)\left(1 - F(s)p_U(s)\right)$$
$$1 - p_B(s) = \exp\left(-sp_U(s) - p_B(s)\right)$$

Main result: Asymptotic overlap

Theorem (Moharrami-Moore-X. '25)

Let F denote the CDF of planted weight distribution P. Then

$$\lim_{n\to\infty} \frac{1}{n-1}\mathbb{E}[|T_{\min} \cap T^*|] = \int_0^\infty (1 - p_U(s)p_B(s))\, dF(s)$$

where $p_U(s)$ and $p_B(s)$ are the largest fixed point of

$$1 - p_U(s) = \exp\left(-sp_U(s) - p_B(s)\right)\left(1 - F(s)p_U(s)\right)$$
$$1 - p_B(s) = \exp\left(-sp_U(s) - p_B(s)\right)$$

Remark

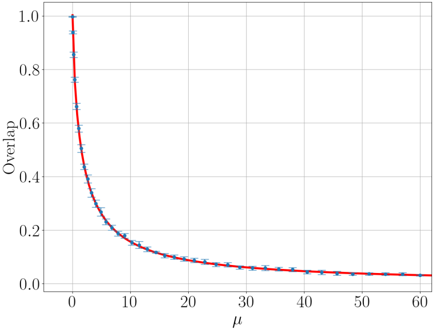
- $p_U(s)$ and $p_B(s)$ are probabilities of certain branching process growing to infinity
- The fixed-point equations have at most two solutions, one of which is zero

Main result: Asymptotic overlap

Theorem (Moharrami-Moore-X. '25)

Let F denote the CDF of planted weight distribution P. Then

$$\lim_{n\to\infty} \frac{1}{n-1}\mathbb{E}[|T_{\min} \cap T^*|] = \int_0^\infty (1 - p_U(s)p_B(s))\, dF(s)$$



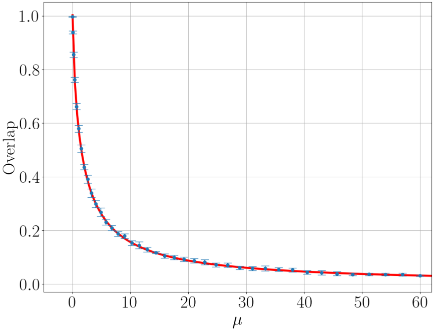
$P = \exp(1/\mu)$
$Q = \exp(1/n)$
mean: μ vs n

Main result: Asymptotic overlap

Theorem (Moharrami-Moore-X. '25)

Let F denote the CDF of planted weight distribution P. Then

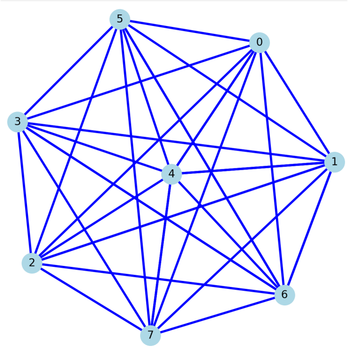
$$\lim_{n\to\infty} \frac{1}{n-1}\mathbb{E}[|T_{\min} \cap T^*|] = \int_0^\infty (1 - p_U(s)p_B(s))\, dF(s)$$



$P = \exp(1/\mu)$
$Q = \exp(1/n)$
mean: μ vs n

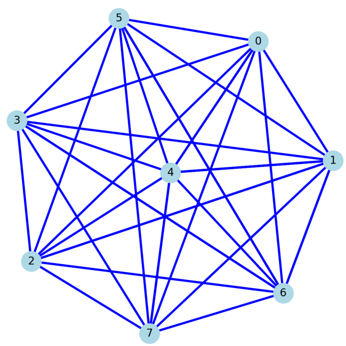
No phase transition!

Warmup: (un-planted) spanning tree problem



- A complete graph on n vertices
- Weights are i.i.d. Q_n with $nQ_n'(0) \to 1$
- What is the mean weight of the minimum spanning tree, $w(T_{\min}) \triangleq \frac{1}{n-1} \sum_{e \in T_{\min}} W_e$?

Warmup: (un-planted) spanning tree problem



- A complete graph on n vertices
- Weights are i.i.d. Q_n with $nQ'_n(0) \to 1$
- What is the mean weight of the minimum spanning tree, $w(T_{\min}) \triangleq \frac{1}{n-1} \sum_{e \in T_{\min}} W_e$?

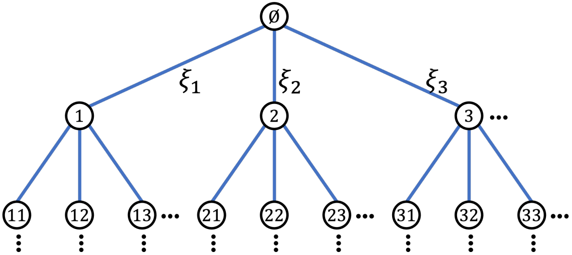
[Frieze '85, Aldous-Steele '04, Addario-Berry'15]

$$\mathbb{E}[w(T_{\min})] \to \zeta(3) \triangleq \sum_{i=1}^{\infty} i^{-3} \approx 1.202$$

Proved in several distinct and elegant ways

Poisson-weighted infinite tree approximation

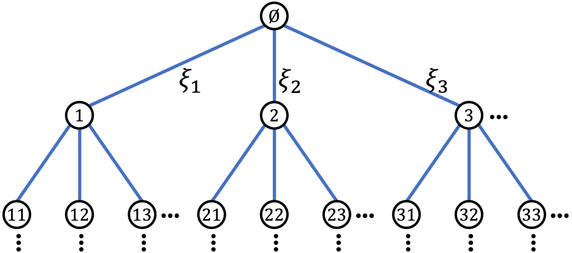
Treat edge weight as distance and model "local" neighborhood as a tree
[Aldous'00, Aldous-Steele '04]



Sort iid Q_n edge weights $W_{\varnothing,1}, W_{\varnothing,2}, \ldots$ from smallest to largest:
$\xrightarrow{n\to\infty}$ arrival times $\xi_1, \xi_2, \ldots$ of a Poisson process with rate 1

Poisson-weighted infinite tree approximation

Treat edge weight as distance and model "local" neighborhood as a tree
[Aldous'00, Aldous-Steele '04]

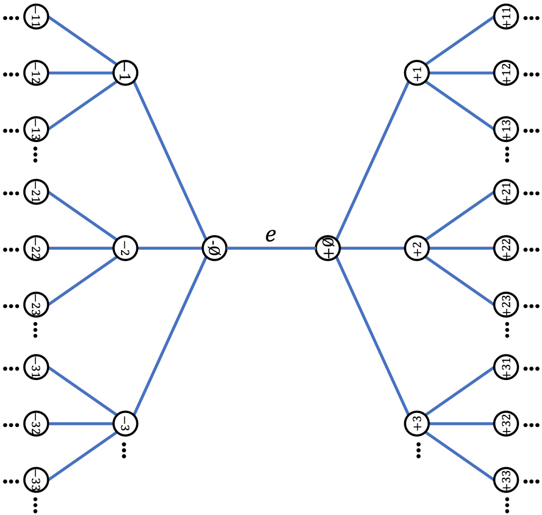


Sort iid Q_n edge weights $W_{\varnothing,1}, W_{\varnothing,2}, \ldots$ from smallest to largest:
$\xrightarrow{n\to\infty}$ arrival times $\xi_1, \xi_2, \ldots$ of a Poisson process with rate 1

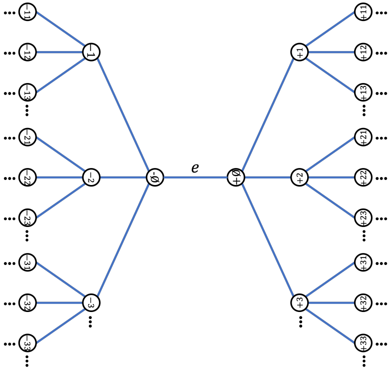
Removing all edges with weights s or greater, we get Galton-Watson tree with $\mathrm{Pois}(s)$ offspring distribution, denoted by PGW(s)

From the perspective of an edge

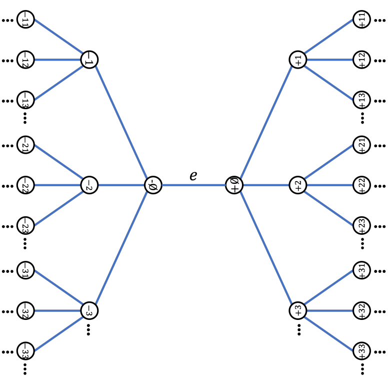
Pick an edge and model its local neighborhood as a two-sided infinite tree



From the perspective of an edge, cont'd

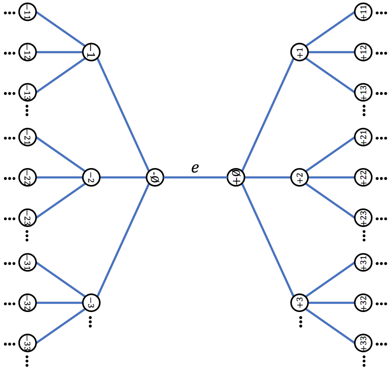


- Suppose e has weight s



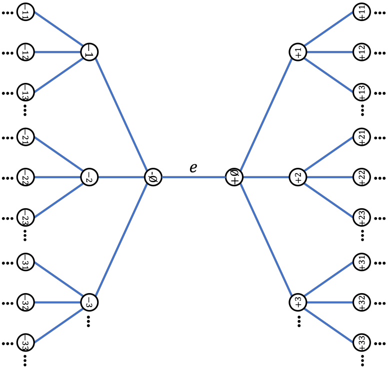
- Suppose e has weight s
- e in MST iff removing all edges with weight s or greater, no path exists between its two endpoints $\Leftrightarrow$ at least one side is finite

From the perspective of an edge, cont'd



- Suppose e has weight s
- e in MST iff removing all edges with weight s or greater, no path exists between its two endpoints
  $\Leftrightarrow$ at least one side is finite
- After removing all edges with weight s or greater, two sides are independent PGW(s)

From the perspective of an edge, cont'd

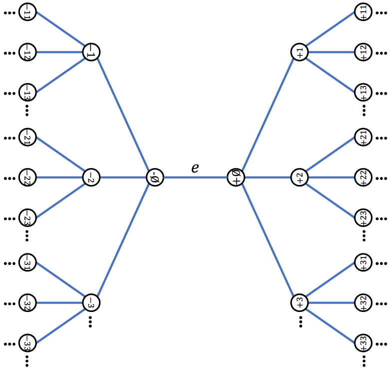


- Suppose e has weight s
- e in MST iff removing all edges with weight s or greater, no path exists between its two endpoints $\Leftrightarrow$ at least one side is finite
- After removing all edges with weight s or greater, two sides are independent $\mathrm{PGW}(s)$

Let $p(s) = \mathbb{P}[\mathrm{PGW}(s) \text{ is infinite}]$. Then $1 - p(s) = \exp(-sp(s))$ and

$$\mathbb{P}[e \in T_{\min}] = 1 - p^2(s)$$

From the perspective of an edge, cont'd



- Suppose e has weight s
- e in MST iff removing all edges with weight s or greater, no path exists between its two endpoints $\Leftrightarrow$ at least one side is finite
- After removing all edges with weight s or greater, two sides are independent $\mathrm{PGW}(s)$

Let $p(s) = \mathbb{P}\left[\mathrm{PGW}(s) \text{ is infinite}\right]$. Then $1 - p(s) = \exp(-sp(s))$ and

$$\mathbb{P}\left[e \in T_{\min}\right] = 1 - p^2(s) \Rightarrow \mathbb{E}[w(T_{\min})] \to \frac{1}{2}\int_0^\infty s(1 - p^2(s))ds = \zeta(3)$$

Back to planted spanning tree

Question: How does the planted uniform spanning tree look like locally?

Back to planted spanning tree

Question: How does the planted uniform spanning tree look like locally?
Answer: Skeleton tree [Grimmett '80, Aldous '91, Aldous-Steele '04]

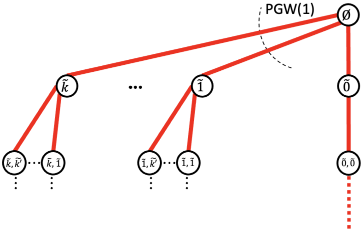
Back to planted spanning tree

Question: How does the planted uniform spanning tree look like locally?
Answer: Skeleton tree [Grimmett '80, Aldous '91, Aldous-Steele '04]

- Start with the infinite path from the root to infinity

Back to planted spanning tree

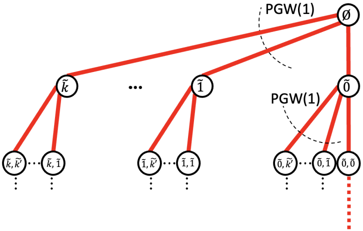
Question: How does the planted uniform spanning tree look like locally?
Answer: Skeleton tree [Grimmett '80, Aldous '91, Aldous-Steele '04]



- Start with the infinite path from the root to infinity
- To each vertex in the infinite path, attach an independent PGW(1)

Back to planted spanning tree

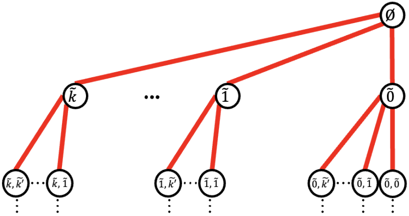
Question: How does the planted uniform spanning tree look like locally?
Answer: Skeleton tree [Grimmett '80, Aldous '91, Aldous-Steele '04]



- Start with the infinite path from the root to infinity
- To each vertex in the infinite path, attach an independent PGW(1)

Putting together: Planted Poisson-weighted infinite tree

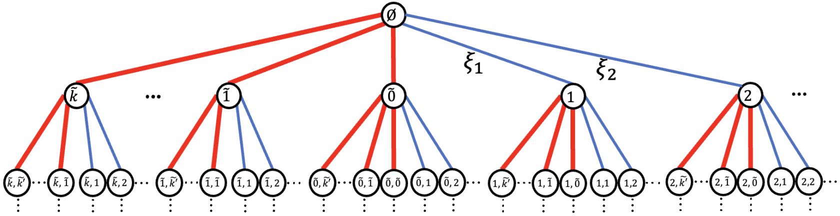
The local view of graph G with planted uniform spanning tree



- The planted (red) edges form skeleton trees

Putting together: Planted Poisson-weighted infinite tree

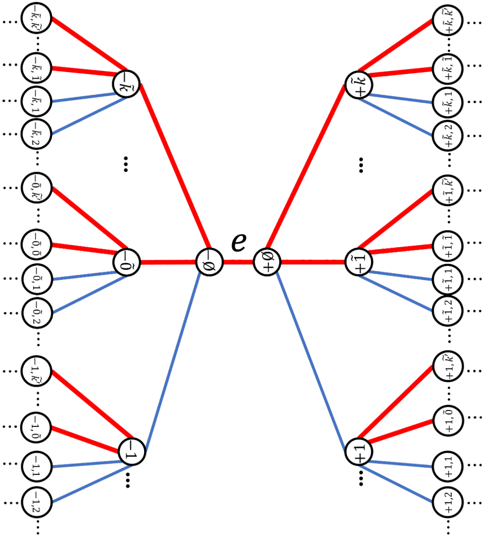
The local view of graph G with planted uniform spanning tree



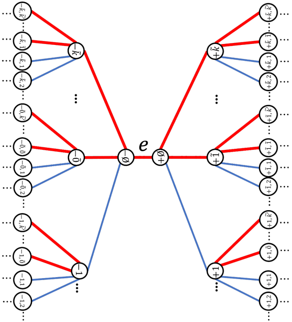
- The planted (red) edges form skeleton trees
- The unplanted (blue) edges form Poisson-weighted infinite trees

From the perspective of an edge, redux

Pick an edge and model its local neighborhood as a two-sided infinite tree

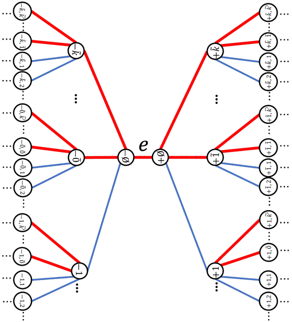


From the perspective of an edge, redux



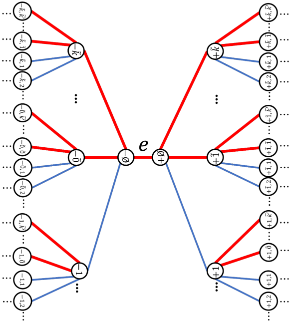
- Suppose e has weight s
- e in MST iff removing all edges with weight s or greater, at least one side is finite.

From the perspective of an edge, redux



- Suppose e has weight s
- e in MST iff removing all edges with weight s or greater, at least one side is finite.
- $p_U(s) \triangleq \mathbb{P}\left[T_{-\phi}(s) \text{ is infinite}\right]$
  $p_B(s) \triangleq \mathbb{P}\left[T_{+\phi}(s) \text{ is infinite}\right]$
  $\mathbb{P}\left[e \in T_{\min}\right] = 1 - p_U(s)p_B(s)$

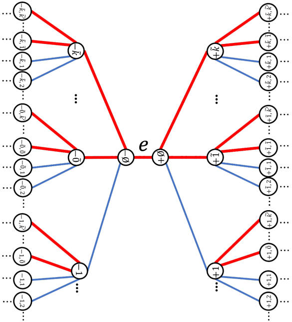
From the perspective of an edge, redux



- Suppose e has weight s
- e in MST iff removing all edges with weight s or greater, at least one side is finite.
- $p_U(s) \triangleq \mathbb{P}\left[T_{-\phi}(s) \text{ is infinite}\right]$
  $p_B(s) \triangleq \mathbb{P}\left[T_{+\phi}(s) \text{ is infinite}\right]$
  $\mathbb{P}\left[e \in T_{\min}\right] = 1 - p_U(s)p_B(s)$

$$1 - p_U(s) = \overbrace{\exp(-s p_U(s))}^{T_{-i}\text{'s are finite}} \times \overbrace{\exp(-p_B(s))}^{T_{-\tilde{k}}\text{'s are finite}} \times \overbrace{(1 - F(s)p_U(s))}^{T_{-\tilde{0}} \text{ is finite}}$$

From the perspective of an edge, redux

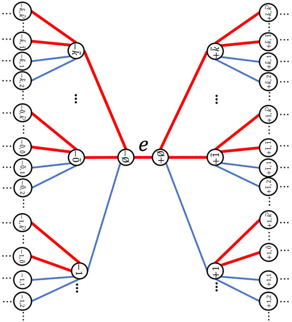


- Suppose e has weight s
- e in MST iff removing all edges with weight s or greater, at least one side is finite.
- $p_U(s) \triangleq \mathbb{P}[T_{-\phi}(s) \text{ is infinite}]$
  $p_B(s) \triangleq \mathbb{P}[T_{+\phi}(s) \text{ is infinite}]$
  $\mathbb{P}[e \in T_{\min}] = 1 - p_U(s)p_B(s)$

$$1 - p_U(s) = \overbrace{\exp(-sp_U(s))}^{T_{-i}\text{'s are finite}} \times \overbrace{\exp(-p_B(s))}^{T_{-\tilde{k}}\text{'s are finite}} \times \overbrace{(1 - F(s)p_U(s))}^{T_{-\tilde{0}} \text{ is finite}}$$

$$1 - p_B(s) = \exp(-sp_U(s)) \times \exp(-p_B(s))$$

From the perspective of an edge, redux



- Suppose e has weight s
- e in MST iff removing all edges with weight s or greater, at least one side is finite.
- $p_U(s) \triangleq \mathbb{P}[T_{-\phi}(s) \text{ is infinite}]$
  $p_B(s) \triangleq \mathbb{P}[T_{+\phi}(s) \text{ is infinite}]$
  $\mathbb{P}[e \in T_{\min}] = 1 - p_U(s)p_B(s)$

$$1 - p_U(s) = \overbrace{\exp(-sp_U(s))}^{T_{-i}\text{'s are finite}} \times \overbrace{\exp(-p_B(s))}^{T_{-\tilde{k}}\text{'s are finite}} \times \overbrace{(1 - F(s)p_U(s))}^{T_{-\tilde{0}} \text{ is finite}}$$

$$1 - p_B(s) = \exp(-sp_U(s)) \times \exp(-p_B(s))$$

$$\Rightarrow \frac{1}{n-1}\mathbb{E}[|T_{\min} \cap T^*|] \to \int_0^\infty (1 - p_U(s)p_B(s))\, dF(s)$$

Proving it: Local weak convergence (Aldous-Steele '04)

- Planted graph G_n converges to planted Poisson-weighted infinite trees T_∞
  - Local treelikeness of light edges
  - Uniform spanning tree converges to skeleton trees [Grimmett '80]
- MST of G_n converges to *minimum spanning forest* of T_∞ [Aldous '91]

Conclusion

- Exactly characterize overlap between MST and the planted one
- Establish local weak convergence of the planted model

Conclusion

- Exactly characterize overlap between MST and the planted one
- Establish local weak convergence of the planted model

Extensions

- Characterize the overlap of MST for planted Hamiltonian path/cycle model [Bagaria-Ding-Tse-Wu-X. '20]
- Determine the weight of MST under the planted model, extending Frieze's $\zeta(3)$ result
- Design an efficient test for planted vs. unplanted models

Conclusion

- Exactly characterize overlap between MST and the planted one
- Establish local weak convergence of the planted model

Extensions

- Characterize the overlap of MST for planted Hamiltonian path/cycle model [Bagaria-Ding-Tse-Wu-X. '20]
- Determine the weight of MST under the planted model, extending Frieze's $\zeta(3)$ result
- Design an efficient test for planted vs. unplanted models

Open problems

1. Determine the information-theoretically optimal overlap
2. Other planted structures, such as spanning k-regular graphs [Sicuro-Zdeborová '21, Gaudio-Sandon-X.-Yang '25]

Conclusion

- Exactly characterize overlap between MST and the planted one
- Establish local weak convergence of the planted model

Extensions

- Characterize the overlap of MST for planted Hamiltonian path/cycle model [Bagaria-Ding-Tse-Wu-X. '20]
- Determine the weight of MST under the planted model, extending Frieze's $\zeta(3)$ result
- Design an efficient test for planted vs. unplanted models

Open problems

1. Determine the information-theoretically optimal overlap
2. Other planted structures, such as spanning k-regular graphs [Sicuro-Zdeborová '21, Gaudio-Sandon-X.-Yang '25]

Reference

M. Moharrami, C. Moore, & J. X., *The planted spanning tree problem*. arXiv: 2502.08790. Conference on Learning Theory 2025