Proceedings of Machine Learning Research vol 291:1–2, 2025 38th Annual Conference on Learning Theory

The Planted Spanning Tree Problems: Exact Overlap Characterization via Local Weak Convergence

Mehrdad Moharrami MOHARAMI@UIOWA.EDU
Computer Science, University of Iowa, Iowa City, IA, USA

Cristopher Moore MOORE@SANTAFE.EDU
Santa Fe Institute, Santa Fe, NM, USA

Jiaming Xu JX77@DUKE.EDU
The Fuqua School of Business, Duke University, Durham, NC, USA

Editors: Nika Haghtalab and Ankur Moitra

We study the problem of recovering a planted spanning tree M_n^* hidden in a complete, randomly weighted graph G_n.[1] Each edge e in G_n has a non-negative weight drawn independently from P_n if $e \in M_n^*$, and from Q_n otherwise. The goal is to recover M_n^* from the observed weighted graph. We assume $P_n \equiv P$ is a fixed continuous distribution, while Q_n varies with n such that its density at zero satisfies $\lim_{n\to\infty} nQ_n'(0) = 1$. A canonical example, the exponential model, takes P and Q_n as exponential distributions with means μ and n, respectively.

We consider two representative cases: (i) when M_n^* is drawn uniformly from all n^{n-2} spanning trees, and (ii) when M_n^* is chosen uniformly from all $n!/2$ Hamiltonian paths. In the uniform spanning tree model, the planted spanning tree may belong to different isomorphism classes, whereas the uniform Hamiltonian path model is a special case restricted to the isomorphism class of n-paths.

As an estimator, we use the minimum-weight spanning tree (MST), denoted by M_n, which coincides with the maximum-likelihood estimator in the uniform spanning tree model with exponential weights. While the MST is not guaranteed to recover Hamiltonian paths (a problem that is NP-hard), it remains computationally tractable and thus a natural choice even in that setting.

To quantify recovery performance, we consider the fraction of common edges, denoted by the overlap $\mathrm{overlap}(M_n, M_n^*)$, between M_n and M_n^*. Using the local weak convergence framework of Aldous and Steele (2004); Steele (2002), we characterize $\lim_{n\to\infty} \mathbb{E}[\mathrm{overlap}(M_n, M_n^*)]$ via a fixed-point equation. Our approach follows the methodology in Moharrami et al. (2021), originally developed for planted matchings. A key distinction in our setting is the non-trivial local structure of the uniform spanning tree, which converges to a *skeleton tree* Aldous (1991); Grimmett (1980). This object consists of an infinite path from the root, with independent Poisson Galton–Watson trees of mean 1 attached to each vertex. In contrast to the planted matching model, which exhibits a sharp phase transition in overlap, we find that no such transition arises here.

We also analyze the asymptotic mean weight of the MST, $\lim_{n\to\infty} \mathbb{E}[w(M_n)]$, where $w(M_n) = \frac{1}{n-1}\sum_{e\in M_n} W_e$, extending the classical result of Frieze (1985) in the unplanted setting with $P_n = Q_n$ uniform. Leveraging this result, we propose an efficient test based on $w(M_n)$ and show that it can distinguish the planted model from the unplanted model with vanishing testing error as $n \to \infty$.

Our analysis demonstrates that local weak convergence provides a powerful framework for planted recovery problems. While we focus on spanning trees and Hamiltonian paths, our techniques extend to other planted structures with tractable local limits, suggesting new avenues for principled recovery guarantees in random graph models.

1. Extended abstract. Full version appears as [arXiv:2502.08790,v2]

References

David Aldous. Asymptotic fringe distributions for general families of random trees. *Annals of Applied Probability*, 1(2):228–266, 1991. doi: 10.1214/aoap/1177005936.

David Aldous and J. Michael Steele. *The Objective Method: Probabilistic Combinatorial Optimization and Local Weak Convergence*, pages 1–72. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. ISBN 978-3-662-09444-0. doi: 10.1007/978-3-662-09444-0_1. URL https://doi.org/10.1007/978-3-662-09444-0_1.

A.M. Frieze. On the value of a random minimum spanning tree problem. *Discrete Applied Mathematics*, 10(1):47–56, 1985. ISSN 0166-218X. doi: https://doi.org/10.1016/0166-218X(85)90058-7. URL https://www.sciencedirect.com/science/article/pii/0166218X85900587.

Geoffrey R. Grimmett. Random labelled trees and their branching networks. *Journal of the Australian Mathematical Society*, 30(2):229–237, 1980. doi: 10.1017/S1446788700016517.

Mehrdad Moharrami, Cristopher Moore, and Jiaming Xu. The planted matching problem: Phase transitions and exact results. *The Annals of Applied Probability*, 31(6):2663 – 2720, 2021. doi: 10.1214/20-AAP1660. URL https://doi.org/10.1214/20-AAP1660.

J. Michael Steele. Minimal spanning trees for graphs with random edge lengths. In Brigitte Chauvin, Philippe Flajolet, Danièle Gardy, and Abdelkader Mokkadem, editors, *Mathematics and Computer Science II*, pages 223–245, Basel, 2002. Birkhäuser Basel. ISBN 978-3-0348-8211-8.