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

The Fundamental Limits of Recovering Planted Subgraphs

Daniel Z. Lee LEE_D@MIT.EDU
Massachusetts Institute of Technology

Francisco Pernice FPERNICE@MIT.EDU
Massachusetts Institute of Technology

Amit Rajaraman AMIT_R@MIT.EDU
Massachusetts Institute of Technology

Ilias Zadik ILIAS.ZADIK@YALE.EDU
Yale University

Editors: Nika Haghtalab and Ankur Moitra

Abstract

We provide a formula for the limiting MMSE curve for recovery in the planted subgraph model for *any graph (family)* H_n, up to a mild assumption on the density of the graph.

Given an arbitrary subgraph $H = H_n$ and $p = p_n \in (0, 1)$, the planted subgraph model is defined as follows. A statistician observes the union of the "signal," which is a random "planted" copy H^* of H, together with random "noise" in the form of an instance of an Erdős–Rényi graph $G(n, p)$. Their goal is then to recover the planted H^* from the observed graph.

A recent paper Mossel et al. (2023) characterizes the graphs for which the MMSE curve undergoes a sharp phase transition from 0 to 1 as p increases, a behavior known as the All-or-Nothing phenomenon, assuming that the the average degree of H_n grows as at least $\log n$. However, their techniques fail to describe the MMSE curves for graphs that do not display such a sharp phase transition. In this paper, we provide a formula for *the limiting MMSE curve for any graph* $H = H_n$, up to the same mild density assumption. This curve is expressed in terms of a variational formula over pairs of subgraphs of H, and is inspired by the celebrated subgraph expectation thresholds from the probabilistic combinatorics literature Kahn and Kalai (2007). Furthermore, we give a polynomial-time description of the optimizers of this variational problem. This allows one to efficiently compute the MMSE curve for any given dense graph H.

Technically, the proof relies on a novel graph decomposition based on greedy iterated densest subgraph, as well as a min-max duality theorem that enables us to characterize the MMSE in terms of this decomposition.

Our results generalize to the setting of planting arbitrary monotone boolean properties. Here, the statistician observes the union of a planted minimal element $A \subseteq [N]$ of a monotone property and a random $\mathrm{Ber}(p)^{\otimes N}$ vector. In this setting, we provide a variational formula inspired by the so-called "fractional" expectation threshold Talagrand (2010), again describing the MMSE curve (in this case up to a multiplicative constant).

Keywords: Planted subgraphs, Kahn–Kalai thresholds.

1. Extended abstract. Full version appears as [arXiv:2503.15723, v1]

References

Jeff Kahn and Gil Kalai. Thresholds and expectation thresholds. *Combinatorics, Probability and Computing*, 16(3):495–502, 2007.

Elchanan Mossel, Jonathan Niles-Weed, Youngtak Sohn, Nike Sun, and Ilias Zadik. Sharp thresholds in inference of planted subgraphs. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 5573–5577. PMLR, 2023.

Michel Talagrand. Are many small sets explicitly small? In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 13–36, 2010.