Proceedings of Machine Learning Research vol 247:1–3, 2024 37th Annual Conference on Learning Theory

Counting Stars is Constant-Degree Optimal For Detecting Any Planted Subgraph: Extended Abstract

Xifan Yu XIFAN.YU@YALE.EDU
Ilias Zadik ILIAS.ZADIK@YALE.EDU
Peiyuan Zhang PEIYUAN.ZHANG@YALE.EDU
Yale University

Editors: Shipra Agrawal and Aaron Roth

Abstract

We prove that whenever $p = \Omega(1)$ and for any graph H, counting $O(1)$-stars is optimal among all constant degree polynomial tests in terms of strongly separating an instance of $G(n, p)$, from the union of a random copy of H with an instance of $G(n, p)$. Our work generalizes and extends multiple previous results on the inference abilities of $O(1)$-degree polynomials in the literature. [1]

Keywords: planted subgraph, low-degree polynomial, computational hardness

We study the computational limits of low-degree polynomial algorithms for the planted subgraph detection problem. Let $H = H_n$ be an *arbitary* undirected graph. We wish to distinguish a "null" Erdős-Rényi random graph $G(n, p)$ from the planted instance where one observes the union of a random copy of $H = H_n$ and an instance of $G(n, p)$. This is a generalization of a plethora of recently studied models initiated with the study of the planted clique model (Jerrum 1992), which corresponds to the special case where H is a k-clique and $p = 1/2$.

Over the last decade, several papers have studied the power of low-degree polynomials for detecting specific choices of H's in planted subgraph model. In this work, we adopt a unifying perspective and characterize the power of *constant degree* polynomials for *all* graphs H and for *any* $p = \Omega(1)$ at the same time. Perhaps surprisingly, we prove that the optimal constant degree polynomial is always given by simply *counting signed stars* in the input random graph. As a direct corollary, we conclude that the class of constant-degree polynomials is only able to "sense" the degree distribution of the planted graph H, and no other graph theoretic property of it.

To establish our characterization of the power of constant degree polynomials, we show that whenever the degree-D advantage tends to infinity, thresholding a signed star counting polynomial of degree at most D strongly distinguishes the planted model and the null model. At the core of our argument is the analysis of the second moment of the signed star counting polynomials in the planted model. We remark that this completes the picture of constant degree testing for the planted subgraph detection problem, as it is well established in the low-degree literature that a bounded degree-D advantage implies the failure of degree-D polynomials (see e.g. Coja-Oghlan et al. 2022).

1. Extended abstract. The full arXiv version of our paper appears in https://arxiv.org/abs/2403.17766v1.

References

Noga Alon, Michael Krivelevich, and Benny Sudakov. Finding a large hidden clique in a random graph. *Random Structures & Algorithms*, 13(3-4):457–466, 1998.

Vivek Bagaria, Jian Ding, David Tse, Yihong Wu, and Jiaming Xu. Hidden hamiltonian cycle recovery via linear programming. *Operations research*, 68(1):53–70, 2020.

Boaz Barak, Samuel Hopkins, Jonathan Kelner, Pravesh K Kothari, Ankur Moitra, and Aaron Potechin. A nearly tight sum-of-squares lower bound for the planted clique problem. *SIAM Journal on Computing*, 48(2):687–735, 2019.

Matthew Brennan, Guy Bresler, and Wasim Huleihel. Reducibility and computational lower bounds for problems with planted sparse structure. In *Conference On Learning Theory*, pages 48–166. PMLR, 2018.

Amin Coja-Oghlan, Oliver Gebhard, Max Hahn-Klimroth, Alexander S Wein, and Ilias Zadik. Statistical and computational phase transitions in group testing. In *Conference on Learning Theory*, pages 4764–4781. PMLR, 2022.

Yael Dekel, Ori Gurel-Gurevich, and Yuval Peres. Finding hidden cliques in linear time with high probability. *Combinatorics, Probability and Computing*, 23(1):29–49, 2014.

Yash Deshpande and Andrea Montanari. Finding hidden cliques of size n/e n/e in nearly linear time. *Foundations of Computational Mathematics*, 15:1069–1128, 2015.

Abhishek Dhawan, Cheng Mao, and Alexander S Wein. Detection of dense subhypergraphs by low-degree polynomials. *arXiv preprint arXiv:2304.08135*, 2023.

David Gamarnik, Aukosh Jagannath, and Alexander S Wein. Low-degree hardness of random optimization problems. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 131–140. IEEE, 2020.

Bruce Hajek, Yihong Wu, and Jiaming Xu. Computational lower bounds for community detection on random graphs. In *Conference on Learning Theory*, pages 899–928. PMLR, 2015.

Samuel Hopkins. *Statistical inference and the sum of squares method*. PhD thesis, Cornell University, 2018.

Wasim Huleihel. Inferring hidden structures in random graphs. *IEEE Transactions on Signal and Information Processing over Networks*, 8:855–867, 2022.

Mark Jerrum. Large cliques elude the metropolis process. *Random Structures & Algorithms*, 3(4):347–359, 1992. doi: https://doi.org/10.1002/rsa.3240030402. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.3240030402.

Chris Jones, Aaron Potechin, Goutham Rajendran, Madhur Tulsiani, and Jeff Xu. Sum-of-squares lower bounds for sparse independent set. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 406–416. IEEE, 2022.

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

Dmitriy Kunisky, Alexander S Wein, and Afonso S Bandeira. Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio. In *ISAAC Congress (International Society for Analysis, its Applications and Computation)*, pages 1–50. Springer, 2019.

Laurent Massoulié, Ludovic Stephan, and Don Towsley. Planting trees in graphs, and finding them back. In *Conference on Learning Theory*, pages 2341–2371. PMLR, 2019.

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.

Andrea Montanari and Alexander S Wein. Equivalence of approximate message passing and low-degree polynomials in rank-one matrix estimation. *arXiv preprint arXiv:2212.06996*, 2022.

Richard Montgomery. Spanning trees in random graphs. *Advances in Mathematics*, 356:106793, 2019.

Elchanan Mossel, Jonathan Niles-Weed, Nike Sun, and Ilias Zadik. On the second kahn–kalai conjecture. *arXiv preprint arXiv:2209.03326*, 2022.

Jinyoung Park and Huy Tuan Pham. A proof of the Kahn–Kalai conjecture. arXiv:2203.17207, 2022.

Lajos Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

Asaf Rotenberg, Wasim Huleihel, and Ofer Shayevitz. Planted bipartite graph detection. *IEEE Transactions on Information Theory*, 2024.

Alexander S Wein. Optimal low-degree hardness of maximum independent set. *Mathematical Statistics and Learning*, 4(3):221–251, 2022.