Learning Low Degree Hypergraphs (Extended Abstract)

Eric Balkanski EB3224@COLUMBIA.EDU
Oussama Hanguir OH2204@COLUMBIA.EDU
Shatian Wang SW3219@COLUMBIA.EDU
Columbia University

Editors: Po-Ling Loh and Maxim Raginsky

Abstract

We study the problem of learning a hypergraph via edge detecting queries. In this problem, a learner queries subsets of vertices of a hidden hypergraph and observes whether these subsets contain an edge or not. This problem has relevant applications in computational chemistry and molecular biology where hypergraphs are used to represent groups of chemicals and molecules that cause a reaction. Such chemical reaction networks are often unknown a priori, and the goal is to learn which subsets of chemicals react using a small number of experiments. This central problem of interest has motivated a long line of work on hypergraph learning in the edge-detecting queries model, e.g., (Torney, 1999; Angluin and Chen, 2006, 2004; Abasi et al., 2014, 2018; Abasi, 2018). In general, learning a hypergraph with m edges of maximum size d requires $\Omega((2m/d)^{d/2})$ queries (Angluin and Chen, 2006). In this paper, we aim to identify families of hypergraphs that can be learned without suffering from a query complexity that grows exponentially in the size of the edges.

We show that hypermatchings and low-degree near-uniform hypergraphs with n vertices are learnable with $\text{poly}(n)$ queries. For learning hypermatchings (hypergraphs of maximum degree 1), we give an $\mathcal{O}(\log^3 n)$-round algorithm with $\mathcal{O}(n \log^5 n)$ queries. We complement this upper bound by showing that there are no algorithms with $\text{poly}(n)$ queries that learn hypermatchings in $o(\log \log n)$ adaptive rounds. For hypergraphs with maximum degree Δ and edge size ratio ρ, we give a non-adaptive algorithm with $\mathcal{O}((2n)^{\rho\Delta+1} \log^2 n)$ queries. To the best of our knowledge, these are the first algorithms with $\text{poly}(n, m)$ query complexity for learning non-trivial families of hypergraphs that have a super-constant number of edges of super-constant size. Our learning algorithms rely on constructions of collections of queries that satisfy a simple property that we call unique-edge covering, which is a novel and general approach for constructing hypergraph learning algorithms.[1]

Keywords: query learning, hypergraph, edge detecting queries, adaptive complexity, chemical reaction network

1. Extended abstract. Full version appears as [https://arxiv.org/abs/2202.09989, v2]

Acknowledgments

We thank Meghan Pantalia and Matthew Ulgherait for helpful discussions on finding drug combinations that reduce cancer cell viability.

References

Hasan Abasi. Error-tolerant non-adaptive learning of a hidden hypergraph. In *43rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

Hasan Abasi, Nader H Bshouty, and Hanna Mazzawi. On exact learning monotone DNF from membership queries. In *International Conference on Algorithmic Learning Theory*, pages 111–124. Springer, 2014.

Hasan Abasi, Nader H. Bshouty, and Hanna Mazzawi. Non-adaptive learning of a hidden hypergraph. *Theoretical Computer Science*, 716:15 – 27, 2018. ISSN 0304-3975. doi: https://doi.org/10.1016/j.tcs.2017.11.019. URL http://www.sciencedirect.com/science/article/pii/S0304397517308496. Special Issue on ALT 2015.

Dana Angluin and Jiang Chen. Learning a hidden graph using o (log n) queries per edge. In *International Conference on Computational Learning Theory*, pages 210–223. Springer, 2004.

Dana Angluin and Jiang Chen. Learning a hidden hypergraph. *Journal of Machine Learning Research*, 7:2215–2236, December 2006. ISSN 1532-4435.

David C. Torney. Sets pooling designs. *Annals of Combinatorics*, 3(1):95–101, 1999. doi: 10.1007/BF01609879. URL https://doi.org/10.1007/BF01609879.