Statistical Query Algorithms and Low-Degree Tests Are Almost Equivalent: Extended Abstract

Matthew Brennan BRENNANM@MIT.EDU
Guy Bresler GUY@MIT.EDU
MIT

Samuel B. Hopkins HOPKINS@BERKELEY.EDU
UC Berkeley

Jerry Li JERRYL@MICROSOFT.COM
Microsoft Research

Tselil Schramm TSELIL@STANFORD.EDU
Stanford University

Editors: Mikhail Belkin and Samory Kpotufe

Abstract

Researchers use a number of approaches to predict and substantiate information-computation gaps in high-dimensional statistical estimation problems. A prominent approach is to characterize the limits of restricted models of computation, which on the one hand yields strong computational lower bounds for powerful classes of algorithms and on the other hand helps guide the development of efficient algorithms. In this paper, we study two of the most popular restricted computational models, the statistical query framework and low-degree polynomials, in the context of high-dimensional hypothesis testing.

We show that all in a very broad class of "nice" high-dimensional statistics problems have the following property: they are solved by computationally efficient algorithms based on low-degree polynomials if and only if they have small *statistical dimension*. Statistical dimension is closely related to *statistical query (SQ) algorithms*, and our results also show that algorithms based on low-degree polynomials are at least as powerful as SQ algorithms. As corollaries, we obtain new statistical query lower bounds for sparse PCA, tensor PCA and several variants of the planted clique problem.[1]

Keywords Hypothesis testing, statistical query lower bounds, low-degree polynomials.

Acknowledgements M.B. and G.B. are supported by MIT-IBM Watson AI Lab, NSF Career Award CCF-1940205, and ONR N00014-17-1-2147. S.H. is supported by a Miller Postdoctoral Fellowship. T.S. thanks Ankur Moitra, Alex Wein, Fred Koehler, and Adam Klivans for helpful conversations regarding the nature of statistical query algorithms and the implications of this work. Part of this work was done while T.S. was virtually visiting the Microsoft Research Machine Learning and Optimization group, and she is grateful for their hospitality.

1. Extended abstract. Full version appears as [arXiv:2009.06107, v#3].