Norms, XOR lemmas, and lower bounds for $GF(2)$ polynomials and multiparty protocols

Emanuele Viola*

Avi Wigderson†

School of Mathematics, Institute for Advanced Study
Princeton, NJ, 08540
{viola,avi}@ias.edu

Abstract

This paper presents a unified and simple treatment of basic questions concerning two computational models: multiparty communication complexity and $GF(2)$ polynomials. The key is the use of (known) norms on Boolean functions, which capture their approximability in each of these models.

The main contributions are new XOR lemmas. We show that if a Boolean function has correlation at most $\epsilon \leq 1/2$ with any of these models, then the correlation of the parity of its values on m independent instances drops exponentially with m. More specifically:

- *For $GF(2)$ polynomials of degree d, the correlation drops to $\exp\left(-m/4^d\right)$. No XOR lemma was known even for $d = 2$.*
- *For c-bit k-party protocols, the correlation drops to $2^c \cdot \epsilon^{m/2^k}$. No XOR lemma was known for $k \geq 3$ parties.*

Another contribution in this paper is a general derivation of direct product lemmas from XOR lemmas. In particular, assuming that f has correlation at most $\epsilon \leq 1/2$ with any of the above models, we obtain the following bounds on the probability of computing m independent instances of f correctly:

- *For $GF(2)$ polynomials of degree d we again obtain a bound of $\exp\left(-m/4^d\right)$.*
- *For c-bit k-party protocols we obtain a bound of $2^{-\Omega(m)}$ in the special case when $\epsilon \leq \exp\left(-c \cdot 2^k\right)$. In this range of ϵ, our bound improves on a direct product lemma for two-parties by Parnafes, Raz, and Wigderson (STOC '97).*

We also use the norms to give improved (or just simplified) lower bounds in these models. In particular we give a new proof that the Mod_m function on n bits, for odd m, has correlation at most $\exp(-n/4^d)$ with degree-d $GF(2)$ polynomials.

1. Introduction

1.1. Background

A natural measure of agreement between two functions is their *correlation*, which measures the agreement on a random input. Formally, the correlation between two functions $f, p \in \{0,1\}^n \to \{-1,1\}$ is defined as

$$\begin{aligned}\mathrm{Cor}(f,p) &:= |E_x[f(x)\cdot p(x)]| \\ &= \left|\Pr_x[f(x)=p(x)] - \Pr_x[f(x)\neq p(x)]\right| \in [0,1].\end{aligned}$$

For a complexity class C (e.g., circuits of size s on n bits), we denote by $\mathrm{Cor}(f,C)$ the maximum of $\mathrm{Cor}(f,p)$ over all functions $p \in C$. In other words, $\mathrm{Cor}(f,C)$ captures how well on average can we compute f using a function from C.

Correlation bounds are fundamental in computational complexity. Proving that $\mathrm{Cor}(f,C) < 1$ is equivalent to establishing that $f \notin C$, but what is far more desired is proving that $\mathrm{Cor}(f,C)$ is very close to zero, for natural functions f and complexity classes C. Such bounds yield pseudorandom generators that "fool" the class C (e.g. [27, 29, 37, 25, 41]), and they also imply lower bounds for richer classes related to C (e.g., if $\mathrm{Cor}(f,C) < 1/t$ then

*The author is supported by NSF grant CCR-0324906. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundations. This research was partially done while the author was a postdoctoral fellow at Harvard University, supported by NSF grant CCR-0133096, US-Israel BSF grant 2002246, ONR grant N-00014-04-1-0478 (see [40]).

†The author is supported by NSF grant CCR-0324906.