THEORY OF COMPUTING, Volume 4 (2008), pp. 137–168
http://theoryofcomputing.org

Norms, XOR lemmas, and lower bounds for polynomials and protocols

Emanuele Viola[*] Avi Wigderson[†]

Received: July 24, 2007; published: November 18, 2008.

Abstract: This paper presents a unified and simple treatment of basic questions concerning two computational models: multiparty communication complexity and polynomials over $GF(2)$. The key is the use of (known) norms on Boolean functions, which capture their proximity to each of these models (and are closely related to property testers of this proximity).

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

- For polynomials over $GF(2)$ 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 $2^c \cdot \varepsilon^{m/2^k}$. No XOR lemma was known for $k \geq 3$ parties.

[*] Supported by NSF grant CCR-0324906. 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, and ONR grant N-00014-04-1-0478.

[†] Supported by NSF grant CCR-0324906.

ACM Classification: 68Q17

AMS Classification: F.2.3

Key words and phrases: XOR lemma, direct product, lower bound, polynomial over GF(2), multiparty protocol, communication complexity, correlation, norm, degree-d norm, generalized inner product, small-bias, mod-m.

DOI: 10.4086/toc.2008.v004a007

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 $\varepsilon \leq 1/2$ with either of the above models, we obtain the following bounds on the probability of computing m independent instances of f correctly:

- For polynomials over $GF(2)$ 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 $\varepsilon \leq \exp\left(-c \cdot 2^k\right)$.

We also use the norms to give improved lower bounds or simplified proofs of known 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 polynomials over $GF(2)$.

1 Introduction

1.1 Background

A natural measure of agreement between two functions is their "correlation."

Definition 1.1. We define the *correlation*[1] between two functions $f, p : D \to \mathbb{C}$ with respect to a probability distribution Q on D as

$$\mathrm{Cor}_Q(f,p) := |E_{x\sim Q}[f(x)\cdot p(x)]|\,.$$

For a class C of functions (e. g., polynomials of degree d on any number of variables) and Q a family of distributions, one for every domain $D = \mathrm{dom}(f)$ for $f \in C$, we denote by $\mathrm{Cor}_Q(f,C)$ the maximum of $\mathrm{Cor}_Q(f,p)$ over all functions $p \in C$ whose domain is $D := \mathrm{dom}(f)$. Unless specified otherwise, Q is the family of uniform distributions. In this case, we simply write $\mathrm{Cor}(f,p)$. If our functions are $\{-1,1\}$-valued, the correlation can be written as

$$\mathrm{Cor}(f,p) = \left|\Pr_x[f(x) = p(x)] - \Pr_x[f(x) \neq p(x)]\right| \in [0,1]\,,$$

where the probabilities are over the uniform distribution.

For functions that are $\{-1,1\}$-valued and nearly balanced, $\mathrm{Cor}(f,C)$ captures how well we can approximate f by a function from C.

Correlation bounds are fundamental in computational complexity. Proving that $\mathrm{Cor}(f,C) < 1$ is equivalent to establishing that $\pm f \notin C$, but what is far more desirable is to prove that $\mathrm{Cor}(f,C)$ is very close to zero, for natural functions f and classes C. Such bounds yield pseudorandom generators that "fool" the class C (e. g. [30, 32, 40, 28, 44]), and they also imply lower bounds for richer classes related to C (e. g., if $\mathrm{Cor}_Q(f,C) < 1/t$ for some distribution Q then f is not equal to any function which is the

[1] Our notion of "correlation" differs from the standard notion in that we do not balance and do not normalize our functions. However, most of our functions of interest will be nearly balanced and automatically normalized (as Boolean functions), so we stay close to the standard concept.

majority of t functions from C [20]). For such applications, we would like to prove correlation bounds as close to zero as possible.

A celebrated way of decreasing correlation (a.k.a. amplifying hardness) is via an *XOR lemma*, first suggested by Yao in his seminal paper [46] (cf. [13]). One starts with a function f of nontrivial correlation with C, and constructs a new function $f^{\times m}$ (on $n \cdot m$ bits), which is the exclusive-OR of the value of f on m independent inputs. (For functions with range in $\{-1,1\}$ exclusive-OR amounts to multiplication.) The hope is that the correlation with C will decay exponentially with m. This idea is best demonstrated in the information-theoretic setting, in which we try to compute the value of a biased coin. In our language, take C to be the class of constant functions (in any number of variables), and f any function with $|E_x[f(x)]| = \mathrm{Cor}(f,C) = \varepsilon$. Then it is easy to see that $\mathrm{Cor}(f^{\times m},C) = \varepsilon^m$ for every m. So the decay of the correlation in this trivial scenario is purely exponential in m, the number of copies.

Yao's XOR lemma deals with the most studied combinatorial model of computation, namely polynomial-size circuits, and goes as follows. Let C be the set of Boolean circuits of size s on n bits, and let f be any function on n bits with $\mathrm{Cor}(f,C) \le \varepsilon$. Then for any m and any $\alpha > 0$, if C' is the set of circuits of size $s \cdot (\alpha/nm)^2$ on $n \cdot m$ bits then $\mathrm{Cor}(f^{\times m},C') \le \varepsilon^m + \alpha$.

Many proofs of this XOR lemma have been given, starting with Levin [27, 23, 13, 24]. All in fact show that this lemma holds under more restrictive circumstances, namely for any C and C' as long as C includes the majority of about $1/\varepsilon$ functions that are in C' (up to complementing the output). However, none of these proofs can be applied to the computational models for which we actually can establish the existence of functions with non-trivial correlation bounds (i. e., prove lower bounds on complexity), such as low-degree polynomials over $GF(2)$, multiparty protocols, or constant-depth circuits (cf. [41]). Specifically, none of the above proofs can be applied to obtain a correlation bound of $1/n$ for a function on n bits. Another weakness of the results in [27, 23, 13, 24] is their loss in resources (e. g., circuit size) in C' compared to C (cf. [13]).

1.2 Our results

In this paper we prove new XOR lemmas for two models: low-degree polynomials over $GF(2)$, and low-communication multiparty protocols.

Both proofs of our XOR lemmas use a common approach, very different from the one used for circuits. With each of these classes C we associate a real *norm* N on all Boolean functions which has the following properties (informally stated):[2]

1. N CAPTURES CORRELATION WITH C. For every function f, $N(f) \approx \mathrm{Cor}(f,C)$.
2. N IS MULTIPLICATIVE WITH RESPECT TO XOR. If f,g are two functions on *disjoint* inputs then $N(f \cdot g) = N(f) \cdot N(g)$. In particular, $N(f^{\times m}) = N(f)^m$.

Given such a norm N, the proof of an XOR lemma for C is almost straightforward:

$$\mathrm{Cor}(f^{\times m},C) \approx N(f^{\times m}) = N(f)^m \approx \mathrm{Cor}(f,C)^m\,.$$

Of course, the challenge is to find the appropriate norms and prove their properties. As it turns out, much of this work has already been done. Specifically, we will see that if the functions in C (of a fixed

[2] As we discuss later, N will not quite be a norm but rather "close" to a norm.

input length) form a linear code, as is the case with polynomials over $GF(2)$, a norm can be viewed as arising from a local tester for proximity to this code (cf. [2]). And when the functions in C (of a fixed input length) do not form a linear code, as is the case with multi-party protocols, it may be useful first to approximate them by a linear code, for which such a norm exists by the foregoing.

The proofs of some of the central lemmas in this area (notably Lemmas 2.3 and 3.4 in this paper) follow a certain iterated Cauchy-Schwarz scheme through the segregation of some variables in each round. This method was first introduced into the subject by Babai, Nisan, and Szegedy [5]. A strikingly similar method was employed later by Gowers, Bourgain, Green-Tao [14, 15, 6, 16] in various contexts, some of it closely related to our subject and used in this paper.

1.2.1 Polynomials over $GF(2)$

Let P_d be the class of all polynomials of degree at most d (in any number of variables) over $GF(2)$. This class has been studied in many contexts in computational complexity. First, it is a natural class that arises in other settings, like error-correcting codes. Second, it is related to important computational models. For example, it is not hard to see that every Boolean decision tree of depth d is in this class. Another, far less obvious connection was proved by Razborov [36] in his lower bound for unbounded fan-in polynomial-size constant-depth circuits over $GF(2)$. Razborov proved that any function $f : \{0,1\}^n \to \{0,1\}$ computable by such circuits satisfies $\mathrm{Cor}(f,P_d) \geq 1-1/n^{\omega(1)}$ for some $d = \mathrm{poly}(\log n)$. That same paper of Razborov exhibits a symmetric function f satisfying $\mathrm{Cor}(f,P_d) \leq O(1/\sqrt{n})$ for such d, and the quest to find functions of smaller correlation with that class continues. Specifically, no explicit function is known which has correlation at most $1/n$ with polynomials of degree $\log_2 n$. The XOR lemma we prove falls short of meeting this challenge: it gives meaningful amplification only if the degree d is below $\log n$. In particular, we prove that the correlation of the XOR of m copies decays exponentially with $m/2^d$.

Theorem 1.2 (XOR lemma for polynomials over $GF(2)$)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$ be a function such that $\mathrm{Cor}(f,P_d) \leq 1-1/2^d$. Then $\mathrm{Cor}(f^{\times m},P_d) \leq \exp\left(-\Omega\left(m/\left(4^d \cdot d\right)\right)\right)$.*

The implied constants in all occurrences of the Ω notation in this paper are absolute.

No XOR lemma was previously known even for $d = 2$.

The norm we use for the proof of this XOR lemma is the so-called "Gowers norm," or "degree-d norm," introduced by Gowers [14, 15] and independently by Alon et al. [2]. We note that its relationship to the class P_d has already been applied in a variety of contexts. Gowers [14, 15] used it to give sharper bounds in Szemerédi's Theorem on arithmetic progressions in subsets of the integers. Green and Tao [16] found further applications to arithmetic combinatorics. Alon et al. [2] used it for property testing of low-degree polynomials. Finally, Samorodnitsky and Trevisan [37, 38] used it to give optimal results on the free-bit complexity of PCPs. These papers contain various inequalities relating these norms to low-degree polynomials; we use the ones in [16], [2], and in [37].

1.2.2 Multiparty protocols

In Yao's standard 2-party communication complexity model [45], each party holds a separate input, and they attempt to compute (or approximate) a given function of these two inputs by exchanging at most c bits of communication (cf. the excellent monograph [26]). This model has been one of the most extensively studied in complexity theory, and captures essential features of diverse computational settings, from Turing machines, VLSI, and distributed computation, to linear programming and auctions. A variety of techniques for proving strong lower bounds and correlation bounds have been developed.

This model was generalized by Chandra, Furst, and Lipton [7] to the *multiparty model* (often called "number-on-forehead" or NOF model). In k-party communication complexity each party is assigned a separate input again. However, that input (figuratively) resides on that party's forehead, and so (formally) each party knows *all but* its own input. Again, the parties have to compute (or approximate) a function on all k inputs by exchanging c bits of communication. The overlapping information of the parties allows this model to capture more complex settings, like multi-tape Turing machines, branching programs, constant-depth circuits with modular gates and more. Here, lower bounds and even correlation bounds are known as long as k is below $\log n$ (where n is the total input length). These bounds were proven in the seminal work of Babai, Nisan, and Szegedy [5], and remain the state-of-the-art after 18 years of intense work; no explicit function is known to require communication $c = \omega(\log n)$ for $k = \log_2 n$ parties.

The fact that the $\log n$ barrier in our knowledge appears in both our models is no coincidence; a beautiful observation of Håstad and Goldmann [21, Proof of Lemma 4] shows that any degree-d polynomial over $GF(2)$ can be computed by $k = d+1$ parties, exchanging only $c = d+1$ communication bits.[3] Thus, breaking the $\log n$ barrier for multiparty protocols would imply breaking the $\log n$ barrier for polynomials over $GF(2)$. Again, our XOR lemma falls short of breaking this barrier, and shows that when computing the XOR of m copies of a function in this model (with the inputs distributed among the k parties as before), the correlation decays (roughly) like $m/2^k$. More precisely, denoting by $\Pi_{k,c}$ the class of all protocols between k parties exchanging at most c bits, we obtain the following theorem.

Theorem 1.3 (XOR lemma for multiparty protocols)**.** *Let* $f : D^k \to \{-1,1\}$ *be a function such that* $\mathrm{Cor}(f, \Pi_{k,k}) \leq \varepsilon$. *Then* $\mathrm{Cor}(f^{\times m}, \Pi_{k,c}) \leq 2^c \cdot \varepsilon^{m/2^k}$.

No such result was known for $k \geq 3$ parties (although, as explained below, a related assumption was known to imply the same consequence). For $k = 2$ our result can be seen as an alternative proof of an XOR lemma by Shaltiel [39]; cf. Remark 3.12.

Note that in the hypothesis of Theorem 1.3 we only require that the function f has small correlation with k-bit protocols (as opposed to c-bit protocols). In fact, we only need that f has small correlation with a special case of k-bit protocols, cf. Section 3.1. We do not know how to exploit the stronger assumption that f has small correlation with c-bit protocols, and in general we do not know whether our XOR lemma is tight. On the other hand, in this work we prove that the "ideal" XOR lemma, i. e., replacing $2^c \cdot \varepsilon^{m/2^k}$ simply by ε^m in Theorem 1.3, is actually *false* for $k = 2$ and $c = 2$ (Claim 3.13). It would be interesting to find the correct bound.

The norm we use to prove this XOR lemma is the one supplied (indirectly or directly) in certain lower bound proofs for this model [5, 9, 35]. In particular, Chung and Tetali [9] show that this norm bounds the correlation from *above* (which proves one direction of Property 1 in Section 1.2), and they also observe that it is multiplicative with respect to XOR (which proves Property 2 in Section 1.2). With this work in place, we only need to show that this norm bounds the correlation from *below*, too (which proves

[3] We point out that the converse is false: multiparty protocols are stronger than low-degree polynomials, as exemplified by the Mod_3 function.

the other direction of Property 1 in Section 1.2). We also give a somewhat more direct proof that this norm bounds the correlation from above, and extend the norm to complex-valued functions, to obtain correlation bounds for certain unbalanced functions.

Such bounds are implicit in the works by Grolmusz [18] and Babai, Hayes, and Kimmel [4]; those papers introduce discrepancy concepts for complex-valued functions.

1.2.3 Direct product vs. XOR lemmas

XOR lemmas are intimately related to *direct product* lemmas. Here we again start with a function $f : D \to \{-1,1\}$ that does not belong to some class C, and want to amplify its hardness by taking many copies of it on independent inputs. However, rather than requiring the computation of only the XOR of all outputs, we simply require the computation of *all* outputs. In other words, the new function $f^{(m)} : D^m \to \{-1,1\}^m$ is the concatenation of m copies of f,

$$f^{(m)}(x_1,x_2,\dots,x_m) := (f(x_1), f(x_2),\dots, f(x_m))\,.$$

Here the natural measure is the success probability, denoted $\mathrm{Suc}\left(f^{(m)}, C\right)$, of giving the right answer when the m-tuple of inputs is chosen uniformly at random. In this setting it makes sense to allow every output to be computed by a function from C (thus, in a sense, allowing a factor m more resources for this solution), and the results in this section indeed hold in this strong form: we define $\mathrm{Suc}\left(f^{(m)}, C\right)$ to be the maximum, over functions $p_1,\dots,p_m \in C$ with domain D^m and range $\{-1,1\}$, of the probability over $x \in D^m$ that $f^{(m)}(x) = (p_1(x), p_2(x),\dots, p_m(x))$.

As for XOR lemmas, one expects exponential decay of the probability $\mathrm{Suc}\left(f^{(m)}, C\right)$ with m, and in fact such direct product lemmas are known for several models. For Boolean decision trees, Nisan et al. [31] show that the success probability of computing $f^{(m)}$ using decision trees of depth d decays purely exponentially with m (independently of d). For c-bit 2-party protocols, Parnafes et al. [33] prove a decay of the form $\varepsilon \to (1/2+\varepsilon/2)^{\Omega(m/c)}$, which mildly deteriorates with the communication complexity c. This bound is proved using (and somewhat extending and strengthening) the celebrated parallel repetition theorem of Raz [34].

We now discuss the connection between XOR lemmas and direct product lemmas and highlight our contributions.

From XOR to direct product. Intuitively, computing all the m f-outputs for $f^{(m)}$ seems like a much harder task than computing only their exclusive-or for $f^{\times m}$. However, a formal connection of this sort does not seem to have been known. We observe that one can indeed formalize such a connection.

We need the following notation: for a set D, let $F(D) = \bigcup_{k\geq 0}\{-1,1\}^{D^k}$ denote the set of all $\{-1,1\}$-valued functions of any number of variables where each variable ranges over D.

Proposition 1.4 (XOR lemma implies direct product lemma)**.** *Let $T(m,m') := 2^{-m}\sum_{k<m'}\binom{m}{k}$ be the tail of the sum of the binomial coefficients. For every m and $0 < m' < m$, function $f : D \to \{-1,1\}$ and class $C \subseteq F(D)$ of $\{-1,1\}$-valued functions that is closed under projections (i. e., under fixing some of the input variables), we have:*

$$\mathrm{Suc}\left(f^{(m)}, C\right) \leq \mathrm{Cor}\left(f^{\times m'}, C'\right) + T(m,m')\,,$$

where C' consists of products of m' functions from C.

In particular, $\mathrm{Suc}\left(f^{(m)},C\right) \leq \mathrm{Cor}\left(f^{\times m/3},C'\right) + \alpha^m$ *for an absolute constant* $\alpha \approx 0.945$.

Proof. Let $p_1,\ldots,p_m \in C$, $p_i : D^m \to \{-1,1\}$ for every i, be such that with probability ε over $X = (X_1,\ldots,X_m) \in D^m$ we have $f(X_i) = p_i(X)$ for every i. For $x = (x_1,\ldots,x_m)$ and $z = (z_1,\ldots,z_m) \in \{0,1\}^m$ let $P(z,x)$ denote the quantity $\prod_{i\leq m}(f(x_i)\cdot p_i(x))^{z_i}$. Let us choose $Z = (Z_1,\cdots,Z_m)$ uniformly in $\{0,1\}^m$. Observe that the expectation of $(f(x_i)\cdot p_i(x))^{Z_i}$, over the choice of Z_i, is 0 if $f(x_i)\cdot p_i(x) = -1$, which is equivalent to $f(x_i) \neq p_i(x)$; otherwise the expectation is 1. Therefore,

$$\varepsilon = \mathop{\mathrm{E}}_{Z,X}[P(Z,X)] \leq \mathop{\mathrm{E}}_{Z,X}[P(Z,X)\,|\,\mathrm{wt}(Z) \geq m'] + \Pr_{Z}[\mathrm{wt}(Z) < m'] = \mathop{\mathrm{E}}_{Z,X}[P(Z,X)\,|\,\mathrm{wt}(Z) \geq m'] + T(m,m'),$$

where wt denotes Hamming weight. Therefore for some fixed z with $\mathrm{wt}(z) \geq m'$ we have

$$\mathop{\mathrm{E}}_{X}[P(z,X)] \geq \varepsilon - T(m,m').$$

The result now follows by fixing the values of all x_i except exactly m' of them corresponding to $z_i = 1$ so as to maximize the expectation; which shows that the XOR of the function in the non-fixed m' inputs has correlation at least $\varepsilon - T(m,m')$ with an XOR of m' functions in C with some inputs fixed.

The "in particular part" follows from the standard estimate $T(m,m/3) < 2^{(H(1/3)-1)m}$, where H is the binary entropy function. □

Remark 1.5. Proposition 1.4 strengthens a result by Impagliazzo and Wigderson [24, Theorem 11] which is about the special case $m' = 1$ (i. e., computing f), and simplifies its proof: in Proposition 1.4, setting $m' = 1$ gives $\mathrm{Suc}\left(f^{(m)},C\right) \leq \mathrm{Cor}(f,C') + 2^{-m}$, whereas in [24] they obtain $\mathrm{Suc}\left(f^{(m)},C\right) \leq \mathrm{Cor}(f,C') + O(\sqrt{m}\cdot 2^{-m})$.

Combining Proposition 1.4 with our XOR lemma for polynomials over $GF(2)$ (Theorem 1.2) we obtain a direct product lemma for polynomials over $GF(2)$. We note that there is no loss in the degree because, although the reduction given by Proposition 1.4 requires taking products of functions from C, recall that in our $\{-1,1\}$ notation multiplication corresponds to exclusive-OR, an operation which does not increase the degree.

Corollary 1.6 (Direct product lemma for polynomials over $GF(2)$)**.** *Let* $f : \{0,1\}^n \to \{-1,1\}$ *be a function such that* $\mathrm{Cor}(f,P_d) \leq 1 - 1/2^d$. *Then* $\mathrm{Suc}\left(f^{(m)},P_d\right) \leq \exp\left(-\Omega\left(m/\left(4^d\cdot d\right)\right)\right)$.

Similarly, we obtain a direct product lemma for multiparty protocols. As discussed above, we allow each of the m protocols to use c bits of communication (i. e., c represents the amount of communication per instance). However, in the reduction in Proposition 1.4, the protocol for the XOR needs to run $\Omega(m)$ of the protocols for the direct product, and this increases the communication by a factor of m, making the result only meaningful when $\varepsilon \ll 2^{-c\cdot 2^k}$.

Corollary 1.7 (Direct product lemma for multiparty protocols)**.** *Let* $f : D \to \{-1,1\}$ *be a function such that* $\mathrm{Cor}(f,\Pi_{k,k}) \leq \varepsilon \leq 2^{-(c+1)\cdot 2^k}$. *Then* $\mathrm{Suc}\left(f^{(m)},\Pi_{k,c}\right) \leq 2^{-\Omega(m)}$.

The above corollary, in its range of parameter $\varepsilon \ll 2^{-c \cdot 2^k}$, beats the bound for 2-party protocols in [33] discussed above, because the latter never gives success probability smaller than $\exp(-\Omega(m/c))$, no matter what ε is. Also, the proof of our bound is simpler. Moreover, the above corollary is the first direct product result for $k \geq 3$ parties. We stress again that to apply the above corollary we only require that f has small correlation with a special case of k-bit protocols (cf. Section 3.1). Finally, we note that the rightmost quantity $2^{-\Omega(m)}$ in the corollary does not depend on c or k; intuitively, this is possible because the bound only holds when $\varepsilon \ll 2^{-c \cdot 2^k}$.

From direct product to XOR. Connections are also known in the other direction: The seminal Goldreich-Levin theorem [12] shows that if a *circuit* has correlation ε with $f^{\times m}$, then a slightly larger circuit will succeed in computing $f^{(m)}$ correctly with probability $\mathrm{poly}(\varepsilon)$ (cf. [13]). However, this reduction suffers again from the problems discussed at the end of Section 1.1: it usually cannot be implemented in the models for which we can currently prove lower bounds, as it needs to compute majority on inputs of length about $1/\varepsilon$ (cf. [41]). Because of this fact, the direct product lemma for 2-party protocols in [33] does not yield an XOR lemma.

Another important computational model where the direct product problem has been studied is that of *k-prover one-round proof systems*, which are often viewed as *games* between a verifier and k provers who cannot communicate with each other (cf. [10]). The problem was first formulated by Fortnow [11, Sec. 4.5] and answered by Raz's celebrated "Parallel Repetition Theorem" [34] which is an essentially tight direct product lemma for two provers.

In this work we show that the XOR lemma for games is false in a strong sense. Specifically, we exhibit a very simple game G for which any prover strategy has correlation at most $1/2$, but there is a prover strategy that has correlation $1 - 1/2^m$ with $G^{\times m}$ (see Section 3.2.2).

Equivalence of direct product and XOR lemmas for circuits. Although in this paper we mainly apply Proposition 1.4 to the models C of low-degree polynomials over $GF(2)$ and multiparty protocols, the proposition is very general and in particular applies to the model of polynomial-size circuits. For this latter model, using the Goldreich-Levin theorem discussed above, we now have the following equivalence.

Corollary 1.8 (Equivalence of direct product and XOR lemmas for circuits)**.** *Let $C(s)$ denote the class of Boolean circuits of size s, and let $f : \{0,1\}^n \to \{-1,1\}$ be any function. We have:*

1. *(Proposition 1.4)* $\mathrm{Suc}\left(f^{(m)}, C(s)\right) \leq \mathrm{Cor}\left(f^{\times m'}, C(s')\right) + 2^{-\Omega(m)}$*, where $m' = m/3$ and $s' = O(s \cdot m')$, and*

2. *(Goldreich and Levin [12])* $\mathrm{Cor}\left(f^{\times m}, C(s)\right) \leq \left(n \cdot \mathrm{Suc}\left(f^{(m)}, C(s')\right)\right)^{\Omega(1)}$,
where $s' = s \cdot \mathrm{poly}(n/\mathrm{Cor}\left(f^{\times m}, C(s)\right))$.

In particular, let C be the set of all $\mathrm{poly}(n)$*-size circuits, and let $m = m(n)$ be any function such that $m(n) = \omega(\log n)$. Then we have that*

$$\mathrm{Suc}\left(f^{(m(n))}, C\right) \leq 1/n^{\omega(1)} \text{ \textit{if and only if} } \mathrm{Cor}\left(f^{\times m(n)}, C\right) \leq 1/n^{\omega(1)}.$$

1.2.4 Lower bounds

The intimate connection of the norms used above to correlation bounds in these models naturally invites their use for proving lower bounds. Indeed, as mentioned earlier, this is exactly what was done in the case of multiparty protocols. We apply this connection to polynomials over $GF(2)$, obtaining a number of new bounds which somewhat improve and considerably simplify correlation bounds for some natural functions. Our bounds rely on the fact that the correlation of any function f with a degree-d polynomial over $GF(2)$ can (essentially) be bounded from above by the degree-d norm of the function f raised to the power of 2^{-d} (Lemma 2.3). Using this fact we obtain the following results.

(1) We consider the Mod_m function on n bits, defined as $\mathrm{Mod}_m(x_1, x_2, \ldots, x_n) = 1$ iff $\sum_i x_i \equiv 0 \pmod{m}$, for a fixed odd integer m. We prove that this function has correlation at most $\exp\left(-\Omega\left(n/4^d\right)\right)$ with any polynomial over $GF(2)$ of degree d with respect to a certain distribution Q on $\{0,1\}^n$ (the distribution Q is defined in Section 2.3). A correlation bound of $\exp\left(-\Omega\left(n/8^d\right)\right)$ was first proved in a breakthrough result by Bourgain [6][4].

After our work [43], Chattopadhyay [8] showed how to modify Bourgain's proof to obtain the same $\exp\left(-\Omega\left(n/4^d\right)\right)$ bound we obtain. Our proof appears to be more modular than the proofs in [6, 17, 8]. It proceeds by again relating the correlation to the degree norm, and then giving an exact calculation of the degree norm of the Mod_m function, yielding $\exp\left(-\Theta\left(n/2^d\right)\right)$. However, the techniques of [6, 17, 8] generalize to polynomials modulo q for arbitrary q relatively prime to m, while our methods appear to be limited to $q = 2$.

(2) We exhibit a polynomial-time computable function on n bits whose correlation with any polynomial of degree d over $GF(2)$ is at most $\exp\left(-\Omega\left(n/2^d\right)\right)$. Prior to our work, in the range $d \ll \log n$ the best correlation bound for an explicit function was $\exp\left(-\Omega\left(n/\left(d \cdot 2^d\right)\right)\right)$, which follows from the multiparty communication complexity lower bound by Babai, Nisan, and Szegedy [5] and the connection between such multiparty protocols and low-degree polynomials discussed in Section 1.2.2. To obtain this result, we note that (for any $d \leq n/2$) a random function $F : \{0,1\}^n \to \{-1,1\}$ has degree-d norm that is exponentially small (i. e., $\exp(-\Omega(n))$) with high probability. We derandomize this probabilistic construction by showing that the same holds when the truth-table of F (of length 2^n) is selected at random from a *small-bias space* [29, 1]. A function F_s from such a sample space can be generated using only an $O(n)$-bit random string s, which we can include as part of the input to our function. Thus, we see that the function $f(s,x) := F_s(x)$ has correlation at most $\exp\left(-\Omega\left(n/2^d\right)\right)$ with any polynomial over $GF(2)$ of degree d. In particular, using a construction by Alon et al. [1], we obtain the result that this correlation bound holds for the function $(\alpha, \beta, x) \mapsto \langle \alpha^x, \beta \rangle$, where α is an element of $GF(2^n)$ and $\langle \cdot, \cdot \rangle$ denotes inner product modulo 2.

Organization of the paper. This paper is organized as follows. In Section 2 we discuss polynomials over $GF(2)$, while in Section 3 we discuss multiparty protocols. For each of these models, we first describe the associated norm, then use it to prove the XOR and direct product lemmas, and finally to prove lower bounds.

[4] Bourgain's proof [6] contains all the main ideas but has a slight error. A correct proof is given by F. Green et al. [17].

2 Polynomials over $GF(2)$

In this section we present our results on polynomials over $GF(2)$. It is convenient to think of a polynomial p over $GF(2)$ as a function from $\{0,1\}^n$ to $\{-1,1\}$. For example, $p(x_1,x_2,x_3) := (-1)^{x_1\cdot x_2+x_3}$, where $x_i \in \{0,1\}$, is a polynomial over $GF(2)$ mapping $\{0,1\}^3$ to $\{-1,1\}$. In this notation, a product of functions is equivalent to their exclusive-or in the $0/1$ notation.

2.1 Degree-k norm

It is convenient to use the following notation.

Notation 2.1. *For a complex number z and an integer j, we denote by $z^{\underline{j}}$ the complex number z if j is even, and the complex conjugate $\bar{z}$ if j is odd.*

We now define the degree-k norm of a function. Although this is syntactically defined as the expectation of a complex-valued random variable, it is always a non-negative real number (cf. [38]).

Definition 2.2 (Degree-k norm[5])**.** Let $f : \{0,1\}^n \to \mathbb{C}$ be a function and $k \geq 1$ an integer. The *degree-k norm* of f is defined as

$$U_k(f) := \mathop{\mathrm{E}}_{y_1,y_2,\ldots,y_k,x\in\{0,1\}^n}\left[\prod_{S\subseteq[k]} f\left(x \oplus \bigoplus_{j\in S} y_j\right)^{\underline{|S|}}\right],$$

where $\oplus$ denotes bitwise XOR.

Degree-d polynomials form a linear code, known as the Reed-Muller code. A "parity check" of a code is a vector in the dual code. In the above definition, we focus on parity checks of low Hamming weight, pick a random one among these (corresponding to the choice of $y_1,\ldots,y_k,x$) and essentially check if it is orthogonal to the given function f by computing $\prod_{S\subseteq[k]} f\left(x\oplus\bigoplus_{j\in S} y_j\right)^{\underline{|S|}}$. The same is done in many property testers, see, e. g., [2]. For a Boolean function, the above norm equals the probability that a random parity check succeeds, minus the probability that it fails. It can be shown that a function f belongs to the class of polynomials of degree $k-1$ if and only if every parity check is 1, in which case the norm is 1 as well. So the norm being 1 captures membership in the class. Now we turn to the study of how smaller values of the norm capture proximity to (or correlation with) the class.

The following lemma shows that the degree norm provides an upper bound on the correlation of a function with polynomials of low degree. This lemma is implicit in the works by Gowers [15] and Green and Tao [16].

Lemma 2.3 (Cf. [15, 16])**.** *For every function $f : \{0,1\}^n \to \mathbb{C}$, $\mathrm{Cor}(f,P_d) \leq U_{d+1}(f)^{1/2^{d+1}}$.*

We need the following lemma for the proof of Lemma 2.3.

Lemma 2.4. *For every function $h : \{0,1\}^n \to \mathbb{C}$, and every k, $U_k(h) \leq \sqrt{U_{k+1}(h)}$.*

[5] The degree-k norm is indeed a norm when raised to the power of $1/2^k$; see, e. g., [16].

Proof of Lemma 2.4. We have:

$$\begin{aligned}
U_k(h) &= \mathop{\mathrm{E}}_{y_1,y_2,\dots,y_{k-1}}\left[\mathop{\mathrm{E}}_{x,y_k}\left[\prod_{S\subseteq[k-1]} h\left(x\oplus\bigoplus_{j\in S}y_j\right)^{\overline{|S|}}\cdot h\left(x\oplus\bigoplus_{j\in S}y_j\oplus y_k\right)^{\overline{|S|+1}}\right]\right]\\
&= \mathop{\mathrm{E}}_{y_1,y_2,\dots,y_{k-1}}\left[\mathop{\mathrm{E}}_{x}\left[\prod_{S\subseteq[k-1]} h\left(x\oplus\bigoplus_{j\in S}y_j\right)^{\overline{|S|}}\right]\cdot\overline{\mathop{\mathrm{E}}_{x}\left[\prod_{S\subseteq[k-1]} h\left(x\oplus\bigoplus_{j\in S}y_j\right)^{\overline{|S|}}\right]}\right]\\
&= \mathop{\mathrm{E}}_{y_1,y_2,\dots,y_{k-1}}\left[\left|\mathop{\mathrm{E}}_{x}\left[\prod_{S\subseteq[k-1]} h\left(x\oplus\bigoplus_{j\in S}y_j\right)^{\overline{|S|}}\right]\right|^2\right]\\
&\geq \left|\mathop{\mathrm{E}}_{y_1,y_2,\dots,y_{k-1},x}\left[\prod_{S\subseteq[k-1]} h\left(x\oplus\bigoplus_{j\in S}y_j\right)^{\overline{|S|}}\right]\right|^2 = U_{k-1}(h)^2. \qquad \left(\text{Because } \mathrm{E}\left[|Z|^2\right]\geq|\mathrm{E}[Z]|^2.\right)
\end{aligned}$$

□

Proof of Lemma 2.3. The lemma follows readily from the following claims, which hold for every function $h:\{0,1\}^n\to\mathbb{C}$:

1. $\left|\mathrm{E}_{x\in\{0,1\}^n}[h(x)]\right| = \sqrt{U_1(h)}$,
2. for every k, $U_k(h)\leq\sqrt{U_{k+1}(h)}$ (Lemma 2.4),
3. for every polynomial over $GF(2)$ p of degree at most d, $U_{d+1}(f\cdot p) = U_{d+1}(f)$.

To see that the above claims imply the lemma, let $p\in P_d$ maximize $\mathrm{Cor}(f,P_d)$, let $h := f\cdot p$, and write

$$\mathrm{Cor}(f,P_d) = |\mathop{\mathrm{E}}_x[h(x)]| = \sqrt{U_1(h)} \leq U_2(h)^{1/2^2}\leq\ldots\leq U_{d+1}(h)^{1/2^{d+1}} = U_{d+1}(f)^{1/2^{d+1}}.$$

We now explain how one obtains the above claims. Claim (1) follows from the definition:

$$|\mathop{\mathrm{E}}_x[h(x)]| = \sqrt{\mathop{\mathrm{E}}_{x,y}\left[h(x)\cdot\overline{h(x\oplus y)}\right]} = \sqrt{U_1(h)}.$$

Claim (2) is Lemma 2.4.

Claim (3) follows from the fact that for every polynomial over $GF(2)$ $p(x)$ of degree d and every fixed $y\in\{0,1\}^n$, the polynomial $q(x) := p(x)\cdot p(x+y)$ has degree $d-1$. For example, consider the polynomial p of degree $d=2$ defined as $p(x) = (-1)^{x_1\cdot x_2}$ for $x = x_1x_2\in\{0,1\}^2$. Then

$$q(x) = p(x)\cdot p(x+y) = (-1)^{x_1\cdot x_2+(x_1+y_1)\cdot(x_2+y_2)} = (-1)^{x_1\cdot y_2+y_1\cdot x_2+y_1\cdot y_2},$$

which is a polynomial of degree 1.

The same three claims above are stated in [16, Equations 1.1, 1.2, and 2.1] for $U_k(h)^{1/2^k}$. □

We now discuss the other direction, namely lower bounds on the correlation in terms of the degree norm. Such bounds arose from the study of property testing of low-degree polynomials. Specifically, Alon et al. [2] define, for a given function $f : \{0,1\}^n \to \{-1,1\}$, a probabilistic procedure and essentially show that if the function satisfies $\mathrm{E}_x[f(x)\cdot p(x)] \le \varepsilon$ for every degree-d polynomial $p : \{0,1\}^n \to \{-1,1\}$ then their procedure rejects with probability $\Omega\left(\min\left\{2^d(1-\varepsilon), 1/(d\cdot 2^d)\right\}\right)$. As noted in [37], the rejection probability of their procedure is $(1-U_{d+1}(f))/2$. Thus we have the following lemma (stated in [25, Theorem 4.1] but essentially proved in [2]).

Lemma 2.5 ([2, 25]). *Let* $f : \{0,1\}^n \to \{-1,1\}$ *be a function such that* $\mathrm{Cor}(f,P_d) \le \varepsilon$. *Then*

$$U_{d+1}(f) \le 1 - \Omega\left(\min\left\{2^d(1-\varepsilon), 1/(d\cdot 2^d)\right\}\right).$$

The above lemma does not bound $U_{d+1}(f)$ by less than $1-\Omega(1/(d\cdot 2^d))$, no matter how small the correlation ε is. Samorodnitsky [37] improved this dependence in the special case of quadratic polynomials (i. e., $d = 2$).

Lemma 2.6 ([37]). *Let* $f : \{0,1\}^n \to \{-1,1\}$ *be a function such that* $\mathrm{Cor}(f,P_2) \le \varepsilon$. *Then* $U_3(f) \le \varepsilon'$, *where* $\varepsilon' \le \log^{-\Omega(1)}(1/\varepsilon)$.

Next, we state the important observation that the norm is multiplicative for functions over disjoint sets of input variables.

Fact 2.7. For functions $f : \{0,1\}^n \to \mathbb{C}$ and $f' : \{0,1\}^{n'} \to \mathbb{C}$, define the function $(f\cdot f') : \{0,1\}^n \times \{0,1\}^{n'} \to \mathbb{C}$ by $(f\cdot f')(x,y) := f(x)\cdot f'(y)$. Then $U_k(f\cdot f') = U_k(f)\cdot U_k(f')$.

2.2 XOR and direct product lemmas for low-degree polynomials over $GF(2)$

In this section we show how the degree norm can be used to obtain XOR lemmas for low-degree polynomials over $GF(2)$. Then we derive a direct product lemma as a corollary.

We repeat our XOR lemma for polynomials for the reader's convenience.

Theorem 1.2 (XOR lemma for polynomials over $GF(2)$, restated). Let $f : \{0,1\}^n \to \{-1,1\}$ be a function such that $\mathrm{Cor}(f,P_d) \le 1-1/2^d$. Then $\mathrm{Cor}(f^{\times m},P_d) \le \exp\left(-\Omega\left(m/\left(4^d\cdot d\right)\right)\right)$.

Proof. Letting $k := d+1$ we have, for $p \in P_d$:

$$\mathop{\mathrm{E}}_x\left[f^{\times m}(x)\cdot p(x)\right] \le U_k\left(f^{\times m}\right)^{1/2^k} = U_k(f)^{m/2^k} \le \left(1-\Omega\left(1/(2^d\cdot d)\right)\right)^{m/2^k} \le 2^{-\Omega\left(m/\left(4^d\cdot d\right)\right)},$$

where the first inequality holds by Lemma 2.3, the next equality by Fact 2.7, and the next inequality by Lemma 2.5. □

Note that if the initial correlation is $\varepsilon \ge 1-1/\left(d\cdot 4^d\right)$, then in fact we can obtain an XOR lemma with the 'correct' dependence on ε, namely $\exp(-\Omega(m\cdot(1-\varepsilon))) \approx \varepsilon^m$ (for simplicity, we did not state this in the theorem). However, if the initial correlation is $\varepsilon \le 1-1/\left(d\cdot 4^d\right)$, we only obtain the stated bound of $\exp\left(-\Omega\left(m\cdot/(d\cdot 4^d)\right)\right)$. This latter dependence can be improved in the special case of quadratic polynomials (i. e., $d = 2$). Specifically, using Lemma 2.6 and reasoning as in the proof Theorem 1.2, we obtain the following XOR lemma for quadratic polynomials over $GF(2)$.

Theorem 2.8 (XOR lemma for quadratic polynomials over $GF(2)$)**.** *Let* $f : \{0,1\}^n \to \{-1,1\}$ *be a function such that* $\mathrm{Cor}(f, P_2) \le \varepsilon$*. Then* $\mathrm{Cor}(f^{\times m}, P_2) \le (\varepsilon')^m$*, where* $\varepsilon' \le \log^{-\Omega(1)}(1/\varepsilon)$*.*

As discussed in Section 1.2.3, combining Theorem 1.2 with Proposition 1.4 we immediately obtain our direct product lemma for low-degree polynomials over $GF(2)$ (Corollary 1.6).

Similarly, one can obtain a direct product lemma for quadratic polynomials over $GF(2)$ by combining Proposition 1.4 with Theorem 2.8.

2.3 The correlation of the Mod_m function with polynomials over $GF(2)$

In this section we study the correlation of low-degree polynomials over $GF(2)$ with the function $\mathrm{Mod}_m : \{0,1\}^n \to \{-1,1\}$, for odd $m \ge 3$, where $\mathrm{Mod}_m(x_1, x_2, \ldots, x_n)$ equals -1 if and only if $\sum_i x_i$ is divisible by m. When working with unbalanced functions like Mod_m, i. e., functions f such that $\Pr_x[f(x) = 1]$ is far from $1/2$, one needs to use a non-uniform distribution Q in the definition of correlation. For fixed n, m, we define Q as follows: with probability $1/2$, Q is uniform over the inputs x such that $\mathrm{Mod}_m(x) = 1$; with probability $1/2$, Q is uniform over the inputs such that $\mathrm{Mod}_m(x) = -1$. Although we will not use this directly, the reader may find it useful to note that

$$\mathrm{Cor}_Q(\mathrm{Mod}_m, p) = \left| \Pr_{x:\mathrm{Mod}_m(x)=1}[p(x) = 1] - \Pr_{x:\mathrm{Mod}_m=-1}[p(x) = 1] \right| .$$

Theorem 2.9. *For any odd m,* $\mathrm{Cor}_Q(\mathit{Mod}_m, P_d) \le \exp\left(-\alpha \cdot n/4^d\right)$*, where* $\alpha = \alpha(m) > 0$ *depends on m only.*

Proof. To model the Mod_m function, define $f : \{0,1\}^n \to \mathbb{C}$ as $f(x_1, \ldots, x_n) := e_m\left(\sum_j x_j\right) = \prod_j e_m(x_j)$, where, denoting by $\mathbf{i}$ the imaginary unit, $e_m(y) := e^{2\pi \cdot \mathbf{i} \cdot y/m}$. We prove below (Lemma 2.10) that there is a constant $\alpha = \alpha(m) > 0$, depending only on m, such that the correlation between any function $p(x) : \{0,1\}^n \to \{-1,1\}$ and the Mod_m function can be bounded as follows:

$$\mathrm{Cor}_Q(\mathrm{Mod}_m, p) \le (1/\alpha) \cdot \max_{a \in \{1,\ldots,m-1\}} \left| \mathop{\mathrm{E}}_{x \in \{0,1\}^n} [f(x)^a \cdot p(x)] \right| + 2^{-\alpha \cdot n} . \tag{2.1}$$

We now focus on bounding the quantity

$$\left| \mathop{\mathrm{E}}_{x \in \{0,1\}^n} [f(x)^a \cdot p(x)] \right|$$

for any fixed a, in the case that p is a polynomial of degree d. For this, we use Lemma 2.3 to relate the quantity to the degree-$(d+1)$ norm of f, and then we use the fact that the norm of the product of functions on disjoint input bits multiplies (Fact 2.7). Formally, letting $k := d+1$, we obtain:

$$\left| \mathop{\mathrm{E}}_{x \in \{0,1\}^n} [f(x)^a \cdot p(x)] \right| \le U_k(f^a)^{1/2^k} = U_k(e_m^a)^{n/2^k} .$$

Thus, we are left with the task of bounding the norm of the 1-bit function e_m^a. We have:

$$U_k(e_m^a) = \mathop{\mathrm{E}}_{y_1,\ldots,y_k,x \in \{0,1\}} \left[e_m \left(a \cdot \sum_{S \subseteq [k]} (-1)^{|S|} \cdot \left(x \oplus \left(\bigoplus_{j \in S} y_j \right) \right) \right) \right] .$$

To bound $U_k\left(e_m^a\right)$, note that whenever $y_1 = y_2 = \cdots = y_k = 1$, we have that

$$\begin{aligned}
&\mathop{\mathrm{E}}_{x\in\{0,1\}}\left[e_m\left(a\cdot\sum_{S\subseteq[k]}(-1)^{|S|}\cdot\left(x\oplus\left(\bigoplus_{j\in S}y_j\right)\right)\right)\right]\\
&= \mathop{\mathrm{E}}_{x\in\{0,1\}}\left[e_m\left(a\cdot\sum_{S\subseteq[k]}(-1)^{|S|}\cdot\left(x\oplus\left(\bigoplus_{j\in S}1\right)\right)\right)\right]\\
&= \frac{e_m\left(a\cdot 2^{k-1}\right)+e_m\left(-a\cdot 2^{k-1}\right)}{2}\\
&= \Re\left(e_m\left(a\cdot 2^{k-1}\right)\right) < 1,
\end{aligned}$$

where $\Re(\cdot)$ denotes the real part, and the last inequality holds because m is odd and $a \in \{1,\ldots,m-1\}$. It is also easy to see that the expectation is 0 whenever $y_j = 0$ for some j (though we do not need this for the upper bound). Since it is the case that $y_1 = y_2 = \cdots = y_k = 1$ with probability 2^{-k}, we have, letting $\delta := \Re\left(e_m\left(a\cdot 2^{k-1}\right)\right)$:

$$U_k\left(e_m^a\right) = \delta\cdot 2^{-k} + 1 - 2^{-k}.$$

Putting everything together, we obtain

$$\left|\mathop{\mathrm{E}}_{x\in\{0,1\}^n}\left[f(x)^a\cdot p(x)\right]\right| \le \left(1-\frac{1-\delta}{2^k}\right)^{n/2^k} < e^{-(1-\delta)n/2^{2k}},$$

which concludes our proof. (Recall that $\delta < 1$ and that $k = d+1$.) □

We conclude this section with a proof of Equation (2.1). Because of later needs, we actually prove a slightly more general claim that holds for the function $GIP_m : \left(\{0,1\}^n\right)^k \to \{-1,1\}$. The inputs to GIP_m are k-tuples $(x_1,\ldots,x_k) \in \left(\{0,1\}^n\right)^k$, and we denote by $(x_i)_j$ the j-th bit of the i-th coordinate $x_i \in \{0,1\}^n$ of $x = (x_1,\ldots,x_k)$. $GIP_m(x)$ equals -1 iff $\sum_{i\le n}\prod_{j\le k}(x_j)_i$ is divisible by m. Note that the Mod_m function is a special case of GIP_m for $k = 1$.

Again, we consider the following non-uniform distribution Q: with probability $1/2$, Q is uniform on the inputs x such that $GIP_m(x) = 1$; with probability $1/2$, Q is uniform on the inputs x such that $GIP_m(x) = -1$. Let now $f : \left(\{0,1\}^n\right)^k \to \mathbb{C}$ be defined as $f(x) := e_m\left(\sum_{\ell\le n}\prod_{j\le k}(x_j)_\ell\right)$, where $e_m(y) := e^{2\pi\cdot\mathbf{i}\cdot y/m}$. Note this coincides with our previous definition for $k = 1$.

Lemma 2.10. *For any $m \ge 2$ there is a constant $\alpha > 0$ such that for any n, k and any function $p : \left(\{0,1\}^n\right)^k \to \{-1,1\}$, the function $GIP_m : \left(\{0,1\}^n\right)^k \to \{-1,1\}$ satisfies*

$$\mathrm{Cor}(p, GIP_m) \le (1/\alpha)\cdot\max_{a\in\{1,\ldots,m-1\}}\left|\mathop{\mathrm{E}}_{x\in(\{0,1\}^n)^k}\left[f(x)^a\cdot p(x)\right]\right| + 2^{-\alpha\cdot n/2^k}.$$

The proof of the above lemma uses "relatively standard" techniques, but a self-contained proof does not seem to have appeared in the literature (but see, e. g., [6, Equation (4)]).

Proof. For a given input $x \in (\{0,1\}^n)^k$, let $\delta(GIP_m(x) = 1)$ denote 1 if $GIP_m(x) = 1$ and 0 otherwise. Similarly, let $\delta(GIP_m(x) \neq 1)$ denote 1 if $GIP_m(x) \neq 1$ and 0 otherwise. Observe that

$$\delta(GIP_m(x) = -1) = \frac{1}{m} \cdot \sum_{b=0}^{m-1} f(x)^b, \qquad \delta(GIP_m(x) = 1) = 1 - \frac{1}{m} \cdot \sum_{b=0}^{m-1} f(x)^b.$$

We need the following claim.

Claim 2.11. *For every $m \geq 2$ there is a constant $\varepsilon > 0$ such that for all n, k we have:*

$$\left| \Pr_{x \in \{0,1\}^n} [GIP_m(x) = -1] - 1/m \right| \leq 2^{-\varepsilon \cdot n/2^k}.$$

Also, for every $m \geq 2$ there is a constant $\varepsilon > 0$ such that for all n, k where $n/2^k$ is sufficiently large we have:

$$\max \left\{ \left| \frac{1}{|\{x : GIP_m(x) = -1\}|} - \frac{m}{2^n} \right|, \left| \frac{1}{|\{x : GIP_m(x) = 1\}|} - \frac{m}{m-1} \cdot \frac{1}{2^n} \right| \right\} \leq 2m^2 \cdot 2^{-n} \cdot 2^{-\varepsilon \cdot n/2^k}.$$

We can assume that $n/2^k$ is sufficiently large by picking α sufficiently small in the statement of the lemma, and thus we can apply the above claim. Recall that the distribution Q with probability $1/2$ is uniformly distributed over the inputs x such that $GIP_m(x) = -1$, and with probability $1/2$ is uniformly distributed over the inputs x such that $GIP_m(x) = 1$. Therefore we can write

$$\begin{aligned}
\mathrm{Cor}_Q(p, GIP_m) &= \left| \mathop{\mathrm{E}}_{x \sim Q} [p(x) \cdot GIP_m(x)] \right| = \left| \sum_x \Pr[Q = x] \cdot p(x) \cdot GIP_m(x) \right| \\
&= \left| \sum_x \Pr[Q = x] \cdot p(x) \cdot \chi(GIP_m(x) = -1) - \sum_x \Pr[Q = x] \cdot p(x) \cdot \chi(GIP_m(x) = 1) \right| \\
&\leq \left| 2^{-n-1} \sum_x p(x) \left(m \cdot \chi(GIP_m(x) = -1) - \frac{m}{m-1} \cdot \chi(GIP_m(x) = 1) \right) \right| + m^{O(1)} \cdot 2^{-\varepsilon \cdot n/2^k} \\
&= \left| 2^{-n-1} \sum_x p(x) \left(m \left(\frac{1}{m} \sum_{b=0}^{m-1} f(x)^b \right) - \frac{m}{m-1} \left(1 - \frac{1}{m} \sum_{b=0}^{m-1} f(x)^b \right) \right) \right| + m^{O(1)} \cdot 2^{-\varepsilon \cdot n/2^k} \\
&= \left| 2^{-n-1} \sum_x p(x) \left(1 + \sum_{b>0}^{m-1} f(x)^b - \frac{m}{m-1} + \frac{1}{m-1} \left(1 + \sum_{b>0}^{m-1} f(x)^b \right) \right) \right| + m^{O(1)} \cdot 2^{-\varepsilon \cdot n/2^k} \\
&= \left| 2^{-n-1} \sum_x p(x) \left(\frac{m}{m-1} \sum_{b>0}^{m-1} f(x)^b \right) \right| + m^{O(1)} \cdot 2^{-\varepsilon \cdot n/2^k} \\
&\leq m^{O(1)} \cdot \max_{b>0} \left| \mathop{\mathrm{E}}_x [p(x) \cdot f(x)^b] \right| + m^{O(1)} \cdot 2^{-\varepsilon \cdot n/2^k} \\
&\leq (1/\alpha) \cdot \max_{b>0} \left| \mathop{\mathrm{E}}_x [p(x) \cdot f(x)^b] \right| + 2^{-\alpha \cdot n/2^k}.
\end{aligned}$$

The last inequality holds for a suitable choice of α, again using that $n/2^k$ is sufficiently large, and proves the lemma.

Proof of Claim 2.11. Let $Z_m = \{0,1,\ldots,m-1\}$ be the additive group with m elements and consider i.i.d. random variables $z_1,\ldots,z_n \in Z_m$ where $z_i = 1$ with probability $\beta := 2^{-k}$ and $z_i = 0$ with probability $\bar{\beta} := 1-\beta$. Let $S := \sum_{i \le n} z_i$, where the sum is modulo m. Note that

$$\Pr_{x\in\{0,1\}^n}[GIP_m(x) = -1] = \Pr_{z_1,\ldots,z_m}[S=0].$$

Let $t(a) := \Pr_{z_1,\ldots,z_m}[S=a]$. By the Fourier Inversion formula (or direct verification),

$$t(0) = \mathop{\mathrm{E}}_{i\in Z_m}\left[\left(\sum_{x\in Z_m} t(x)\cdot e_m(-i\cdot x)\right)\cdot e_m(i\cdot 0)\right] = \mathop{\mathrm{E}}_{i\in Z_m}\left[\sum_{x\in Z_m} t(x)\cdot e_m(-i\cdot x)\right] = \mathop{\mathrm{E}}_{i\in Z_m}[\mathop{\mathrm{E}}_{S}[e_m(-i\cdot S)]].$$

Note that $\mathrm{E}_S[e_m(-i\cdot S)] = 1$ for $i=0$. Now fix any $i \neq 0$, and note that

$$|\mathop{\mathrm{E}}_{S}[e_m(-i\cdot S)]| = |\mathop{\mathrm{E}}_{z_1}[e_m(-i\cdot z_1)]|^n \le (1-\delta\cdot\beta)^n \le 2^{-\varepsilon\cdot n\cdot\beta},$$

for constants δ and ε that depend on m only. To verify the second to last inequality, write $e_m(-i) = (u,v) \in \mathbb{R}^2$, where u is bounded away from 1 by $\gamma > 0$ that depends only on m, and $u^2+v^2 = 1$. Then

$$\begin{aligned}|\mathop{\mathrm{E}}_{z_1}[e_m(-i\cdot z_1)]| &= |(u\beta + 1\cdot\bar{\beta}, v\beta + 0\bar{\beta})| = \sqrt{u^2\beta^2 + 2u\beta\bar{\beta} + \bar{\beta}^2 + v^2\beta^2} \\ &= \sqrt{\beta^2 + 2u\beta\bar{\beta} + \bar{\beta}^2} = \sqrt{1 + 2\beta\bar{\beta}(u-1)} \le 1 + \beta\bar{\beta}(u-1) = (1-\delta\cdot\beta)\end{aligned}$$

for $\delta := (1-u)\bar{\beta}$. Therefore,

$$t(0) = \mathop{\mathrm{E}}_{i\in Z_m}[\mathop{\mathrm{E}}_{S}[e_m(-i\cdot S)]|i=0]\cdot\Pr_i[i=0] + \mathop{\mathrm{E}}_{i\in Z_m}[\mathop{\mathrm{E}}_{S}[e_m(-i\cdot S)]|i\neq 0]\cdot\Pr_{i\in Z_m}[i\neq 0] = 1\cdot\frac{1}{m} + A,$$

where $|A| \le 2^{-\varepsilon\cdot n\cdot\beta}$. This proves the first part of the claim.

The "also" part follows from the following general fact: For all strictly positive real numbers ϕ,γ,ρ, such that $\rho \le \gamma/2$, we have that

$$\phi \in [\gamma-\rho, \gamma+\rho]$$

implies

$$\phi^{-1} \in [\gamma^{-1} - c\cdot\rho, \gamma^{-1} + c\cdot\rho],$$

where $c = 2\gamma^{-2}$. To see this, note that by hypothesis

$$\phi^{-1} \in [1/(\gamma+\rho), 1/(\gamma-\rho)].$$

To conclude, note that

$$\gamma^{-1} - c\cdot\rho \le 1/(\gamma+\rho) \Leftrightarrow 1 + \gamma^{-1}\rho - c\rho\gamma - c\rho^2 \le 1,$$

which is true for $c \ge \gamma^{-2}$, and

$$1/(\gamma-\rho) \le \gamma^{-1} + c\cdot\rho \Leftrightarrow 1 \le 1 + \gamma c\rho - \rho\gamma^{-1} - \rho^2 c \Leftrightarrow \gamma^{-1} + \rho c \le \gamma c,$$

which using that $\rho \leq \gamma/2$ is true for $c \geq 2\gamma^{-2}$.

To obtain the "also" part, let $\phi := \Pr_x[GIP_m(x) = -1]$ (respectively, $\Pr_x[GIP_m(x) = 1]$), $\gamma := 1/m$ (respectively, $1-1/m$), and $\rho := 2^{-\varepsilon \cdot n/2^k}$. We have $\rho \leq \gamma/2$ by our assumption that $n/2^k$ is sufficiently large, and thus we conclude the proof by applying the first part of the claim and the above general fact. □

□

2.4 A function with correlation $\exp\left(-\Omega\left(n/2^d\right)\right)$

In this section we exhibit a polynomial-time computable function on n bits whose correlation with any polynomial over $GF(2)$ of degree d is at most $\exp\left(-\Omega\left(n/2^d\right)\right)$.

Theorem 2.12. *There is a polynomial-time computable function $f : \{0,1\}^n \to \{-1,1\}$ such that for every $d < n/2$ we have $\mathrm{Cor}(f, P_d) \leq \exp\left(-\alpha \cdot n/2^d\right)$, where $\alpha > 0$ is a universal constant.*

As mentioned in the Introduction, previously the best correlation bound in the range $d \ll \log n$ was the one implicit in BNS [5] via the Håstad-Goldmann argument [21], namely, an $\exp\left(-\alpha \cdot n/(d \cdot 2^d)\right)$ bound in the stronger computational model of $(d+1)$-party protocols. Our proof is similar to theirs; it exploits a property of the target function which is captured in Lemma 2.13 below. Our main contribution is to show that using the degree-norm one obtains a slightly better bound for the special case of P_d.

Proof. It is sufficient and more convenient to prove the theorem for a function with input length $O(n)$ rather than n. We prove that the theorem holds for the function that on input (σ, x) equals the xth output bit of a small-bias generator on seed σ. The following lemma summarizes the definition and the existence of small-bias generators.

Lemma 2.13 ([29, 1][6])**.** *There is a polynomial-time computable function $f : \{0,1\}^{O(n)} \times \{0,1\}^n \to \{-1,1\}$ such that for every $\emptyset \neq T \subseteq \{0,1\}^n$, we have:*

$$\mathop{\mathrm{E}}_{\sigma}\left[\prod_{x \in T} f(\sigma, x)\right] \leq 2^{-n}.$$

Let f be the function in Lemma 2.13 and write f_σ for the function that maps x to $f(\sigma, x)$. We now show that, over the choice of σ, we expect f_σ to have small degree norm.

Claim 2.14. $\mathrm{E}_\sigma\left[U_k\left(f_\sigma\right)\right] \leq 2^{-\alpha \cdot n}$*, for every $k \leq n/2$, where $\alpha > 0$ is a universal constant.*

[6] Our presentation is syntactically different from the one in [1], which is in terms of sample spaces. The lemma stated here follows from the results in [1] by considering a small-bias sample space over $\{0,1\}^N$, where $N := 2^n$, and defining $f(\alpha, x)$ to be the xth bit of the sample that corresponds to α.

Proof. Let D be the event (over the choice of $y_1,\ldots,y_k$) that the dimension of the vector space generated by the $y_i's$ is k, i. e., that for every $S,S' \subseteq [k]$ we have $\sum_{j\in S} y_j \neq \sum_{j\in S'} y_j$. We have:

$$\begin{aligned}\mathop{\mathrm{E}}_{\sigma}[U_k(f_\sigma)] &= \mathop{\mathrm{E}}_{x,y_1,\ldots,y_k}\left[\mathop{\mathrm{E}}_{\sigma}\left[\prod_{S\subseteq[k]} f_\sigma\left(x+\sum_{j\in S} y_j\right)\right]\right] \\ &\leq E_{x,y_1,\ldots,y_k}\left[\mathop{\mathrm{E}}_{\sigma}\left[\prod_{S\subseteq[k]} f_\sigma\left(x+\sum_{j\in S} y_j\right)\right]\middle| D\right] + \Pr[\neg D] \leq 2^{-\alpha\cdot n}.\end{aligned}$$

The last inequality above is obtained by bounding each term separately. For the first term, we observe that, conditioned on D, $\prod_{S\subseteq[k]} f_\sigma\left(x+\sum_{j\in S} y_j\right) = \prod_{z\in T} f_\sigma(z)$ where T consists of the 2^k distinct values $x+\sum_{j\in S} y_j$ for $S\subseteq[k]$, and then we apply Lemma 2.13. As for the second term, we note that D is the event: "$y_1 \notin \mathrm{Span}(0)$ and $y_2 \notin \mathrm{Span}(y_1)$ and ... and $y_k \notin \mathrm{Span}(y_1,y_2,\ldots,y_{k-1})$". Thus we obtain

$$\Pr[\neg D] = 1-\left(1-2^{-n}\right)\left(1-2^{-n+1}\right)\cdots\left(1-2^{-n+k-1}\right) \leq 1-\left(1-2^{-n+k-1}\right)^{k-1} \leq 2^{-\alpha\cdot n}$$

for a universal constant $\alpha > 0$, using that $k \leq n/2$. □

To conclude the proof of the theorem, let $p : \{0,1\}^n \to \{-1,1\}$ be any polynomial over $GF(2)$ of degree d, and notice that

$$\begin{aligned}\mathop{\mathrm{E}}_{\sigma,x}[f(\sigma,x)\cdot p(\sigma,x)] &= \mathop{\mathrm{E}}_{\sigma}\left[\mathop{\mathrm{E}}_{x}[f_\sigma(x)\cdot p(\sigma,x)]\right] \\ &\leq \mathop{\mathrm{E}}_{\sigma}\left[U_{d+1}(f_\sigma)^{1/2^{d+1}}\right] \leq \mathop{\mathrm{E}}_{\sigma}[U_{d+1}(f_\sigma)]^{1/2^{d+1}} \leq 2^{-\alpha\cdot n/2^d},\end{aligned}$$

where $\alpha > 0$ is a universal constant, the first inequality holds by Lemma 2.3, the second is Jensen's inequality, and the last holds by Claim 2.14. □

Remark 2.15 (On the tightness of Theorem 2.12)**.** It is natural to ask whether the $\exp\left(-\Omega\left(n/2^d\right)\right)$ correlation bound is tight for the particular function f given by Theorem 2.12, which (recall) computes the xth bit of a small-bias generator, given the seed and x. We observe that this bound is somewhat tight in the sense that, for some small-bias generator, the associated function f has correlation $1-o(1)$ with some polynomial over $GF(2)$ of degree $d = \log^{O(1)} n$. This follows from the fact that, for some small-bias generator, the associated function f is computable by polynomial-size constant-depth circuits with parity gates [19, 22][7] and the well-known fact that any such function has correlation at least $1-o(1)$ with some polynomial over $GF(2)$ of degree $\log^{O(1)} n$ [36, 42].

3 Multiparty protocols

In this section we discuss our results on multiparty protocols. Rather than working directly with multiparty protocols, in the next section we introduce a simple subclass Π_k^* of such protocols, which happens

[7] These works give *uniform* circuits, while for the point made here, non-uniform circuits would suffice. However, we do not know of a simpler proof of existence of such circuits.

to be a linear code. We introduce a norm capturing proximity with this class in a way analogous to what we did with polynomials over $GF(2)$. Then in Section 3.2 we discuss the general multiparty model, and show that it can be reasonably well approximated by simple protocols, and hence the same norm yields an XOR lemma and lower bounds for it as well. (When dealing with linear codes such as Π_k^*, recall that addition modulo 2 becomes multiplication with our $\{-1,1\}$ notation.)

3.1 k-party norm and Π_k^* protocols

In this section we discuss the relationship between the model Π_k^* and the k-party norm, both of which are defined next, and then we prove an XOR lemma for Π_k^*.

Definition 3.1 (The model Π_k^*)**.** We say that a function $g_j : D^k \to \{-1,1\}$ is *cylindrical in dimension* j if it does not depend on the jth coordinate. The class Π_k^* consists of the functions $f : D^k \to \{-1,1\}$ that are products of cylindrical functions over all dimensions. Equivalently, Π_k^* is the class of functions $f : D^k \to \{-1,1\}$ such that $f(x_1,\dots,x_k) = \prod_{j\le k} g_j(x_1,\dots,x_k)$ for some functions $g_1,\dots,g_k$ such that g_j does not depend on the input x_j.

This definition is motivated by the concept of “cylinder intersections” introduced by Babai, Nisan, and Szegedy.

Definition 3.2 (Cylinder intersection [5])**.** A subset of D^k is called a *cylinder* in dimension j if its characteristic function is cylindrical in dimension j. A subset of D^k is a *cylinder intersection* if it is the intersection of cylinders in all dimensions.

It is not hard to see that the model Π_k^* above is a linear code. We now define the k-party norm (recall Notation 2.1). This norm is implicit in [5] and explicit in [9, 35]. As we will see (cf. Remark 3.12), this quantity is closely related to the discrepancy over the family of cylinder intersections, the central concept studied in [5] (cf. [26]). The discrepancy of a function $f : D^k \to \{-1,1\}$ is

$$\max_S |E_x[f(x)|x\in S]| \cdot \Pr[x\in S]\,,$$

where the maximum is over all cylinder intersections S.

Definition 3.3 (k-party norm)**.** Let $f : D^k \to \mathbb{C}$ be a function. The *k-party norm* of f is defined as

$$R_k(f) := \mathop{\mathrm{E}}_{\substack{x_1^0,x_2^0,\dots,x_k^0\in D\\ x_1^1,x_2^1,\dots,x_k^1\in D}} \left[\prod_{\varepsilon_1,\dots,\varepsilon_k\in\{0,1\}} f\left(x_1^{\varepsilon_1},x_2^{\varepsilon_2},\dots,x_k^{\varepsilon_k}\right)^{\underline{\sum_{j\le k}\varepsilon_j}}\right].$$

Similarly to Section 2, the norm $R_k(f)$ can be seen as computing random short “parity checks” of the linear code Π_k^*. For a Boolean function, the above expectation equals the probability that a random parity check succeeds, minus the probability that it fails. It can be shown that a function f belongs to the class (which is a linear code) Π_k^* if and only if every parity check is 1, in which case the norm is 1 as well. So the norm being 1 captures membership in the class. Now we turn to study how smaller values of the norm capture proximity to (or correlation with) the class.

First, we have the following lemma that shows that the norm bounds the correlation with Π_k^* from above. The same lemma (for real-valued functions) is implicit in [9, 35].

Lemma 3.4 ([9, 35])**.** *For every function* $f : D^k \to \mathbb{C}$, $\mathrm{Cor}(f,\Pi_k^*) \le R_k(f)^{1/2^k}$.

The proof of Lemma 3.4 is very similar to that of Lemma 2.3, and makes use of the following lemma.

Lemma 3.5 ([9, 35])**.** *For any function* $f : D^k \to \mathbb{C}$, $|\,\mathrm{E}_{x\in D^k}[f(x)]\,| \le R(f)^{1/2^k}$.

Proof of Lemma 3.5. We have:

$$
\begin{aligned}
R_k(f) &= \mathop{\mathrm{E}}_{\substack{x_1^0,\dots,x_{k-1}^0\in D\\ x_1^1,\dots,x_{k-1}^1\in D}} \left[\mathop{\mathrm{E}}_{\substack{x_k^0\\ x_k^1}} \left[\prod_{\varepsilon_1,\dots,\varepsilon_{k-1}\in\{0,1\}} f\left(x_1^{\varepsilon_1},\dots,x_{k-1}^{\varepsilon_{k-1}},x_k^0\right)^{\overline{\sum_{\ell<k}\varepsilon_\ell}} \cdot f\left(x_1^{\varepsilon_1},\dots,x_{k-1}^{\varepsilon_{k-1}},x_k^1\right)^{\overline{1+\sum_{\ell<k}\varepsilon_\ell}} \right]\right] \\
&= E_{\substack{x_1^0,\dots,x_{k-1}^0\in D\\ x_1^1,\dots,x_{k-1}^1\in D}} \left[\mathop{\mathrm{E}}_{x_k\in D} \left[\prod_{\varepsilon_1,\dots,\varepsilon_{k-1}\in\{0,1\}} f\left(x_1^{\varepsilon_1},\dots,x_{k-1}^{\varepsilon_{k-1}},x_k\right)^{\overline{\sum_{\ell<k}\varepsilon_\ell}} \right] \right. \\
&\qquad\qquad \left. \overline{\cdot \mathop{\mathrm{E}}_{x_k\in D} \left[\prod_{\varepsilon_1,\dots,\varepsilon_{k-1}\in\{0,1\}} f\left(x_1^{\varepsilon_1},\dots,x_{k-1}^{\varepsilon_{k-1}},x_k\right)^{\overline{\sum_{\ell<k}\varepsilon_\ell}} \right]} \right] \\
&= \mathop{\mathrm{E}}_{\substack{x_1^0,\dots,x_{k-1}^0\in D\\ x_1^1,\dots,x_{k-1}^1\in D}} \left[\left| \mathop{\mathrm{E}}_{x_k\in D} \left[\prod_{\varepsilon_1,\dots,\varepsilon_{k-1}\in\{0,1\}} f\left(x_1^{\varepsilon_1},\dots,x_{k-1}^{\varepsilon_{k-1}},x_k\right)^{\overline{\sum_{\ell<k}\varepsilon_\ell}} \right] \right|^2 \right] \\
&\ge \left| \mathop{\mathrm{E}}_{\substack{x_1^0,\dots,x_{k-1}^0\in D\\ x_1^1,\dots,x_{k-1}^1\in D}, x_k\in D} \left[\prod_{\varepsilon_1,\dots,\varepsilon_{k-1}\in\{0,1\}} f\left(x_1^{\varepsilon_1},\dots,x_{k-1}^{\varepsilon_{k-1}},x_k\right)^{\overline{\sum_{\ell<k}\varepsilon_\ell}} \right] \right|^2 . \qquad \left(\text{Because } \mathrm{E}\left[|Z|^2\right] \ge |\mathrm{E}[Z]|^2.\right)
\end{aligned}
$$

Repeating the above argument k times one obtains the lemma. □

Proof of Lemma 3.4. We have

$$
\begin{aligned}
\mathrm{Cor}(f,\Pi_k^*) = \max_{g\in\Pi_k^*,\ g:D^k\to\{-1,1\}} \left| \mathop{\mathrm{E}}_{x\in D^k}[(f\cdot g)(x)] \right| &\le R_k(f\cdot g)^{1/2^k} && (3.1)\\
&= R_k(f)^{1/2^k}, && (3.2)
\end{aligned}
$$

where Inequality (3.1) holds by Lemma 3.5, and Equation (3.2) is justified as follows. Recall from Definition 3.1 that $g = g_1 \cdot g_2 \cdots g_k$, where each g_j is a cylindrical function in dimension j, i. e., does not depend on the jth coordinate. It is enough to prove Equation (3.2) for every such g_j (since f is arbitrary). We prove it for g_k without loss of generality. Note that for every fixed $x_1^0,x_2^0,\dots,x_k^0,x_1^1,x_2^1,\dots,x_k^1$, we have

$$
\begin{aligned}
&\prod_{\varepsilon_1,\dots,\varepsilon_k\in\{0,1\}} (f\cdot g_k)\left(x_1^{\varepsilon_1},x_2^{\varepsilon_2},\dots,x_k^{\varepsilon_k}\right)^{\overline{\sum_{j\le k}\varepsilon_j}} \\
&= \prod_{\varepsilon_1,\dots,\varepsilon_{k-1}} \left(g_k\left(x_1^{\varepsilon_1},x_2^{\varepsilon_2},\dots,x_{k-1}^{\varepsilon_{k-1}},0\right)^2 \cdot \prod_{\varepsilon_k} f\left(x_1^{\varepsilon_1},x_2^{\varepsilon_2},\dots,x_k^{\varepsilon_k}\right)^{\overline{\sum_{j\le k}\varepsilon_j}} \right) \\
&= \prod_{\varepsilon_1,\dots,\varepsilon_k} f\left(x_1^{\varepsilon_1},x_2^{\varepsilon_2},\dots,x_k^{\varepsilon_k}\right)^{\overline{\sum_{j\le k}\varepsilon_j}},
\end{aligned}
$$

using $g_k^2 \equiv 1$ because g_k takes values in $\{-1,1\}$. □

We now state and prove a new lemma that shows that the k-party norm also bounds from below the correlation. We note that in this model we have a much tighter connection between norm and correlation than for polynomials over $GF(2)$: here we have positive correlation as soon as the norm is positive, whereas for polynomials over $GF(2)$ we needed the norm to be very close to 1 to infer positive correlation (a notable exception is the result in Lemma 2.6 which gives a tighter connection for the special case of quadratic polynomials).

Lemma 3.6. *For every function $f : D^k \to \{-1,1\}$, $\mathrm{Cor}(f, \Pi_k^*) \geq R_k(f)$.*

Proof. For $x_1^1, x_2^1, \ldots, x_k^1 \in D$, consider the function $g_{x_1^1,\ldots,x_k^1} : D^k \to \{-1,1\}$ defined as

$$g_{x_1^1,\ldots,x_k^1}(x_1^0,\ldots,x_k^0) := \prod_{(\varepsilon_1,\ldots,\varepsilon_k)\in\{0,1\}^k\setminus 0^k} f\left(x_1^{\varepsilon_1},\ldots,x_k^{\varepsilon_k}\right) .$$

Now observe that

$$\mathop{\mathrm{E}}_{x_1^1,\ldots,x_k^1}\left[\mathop{\mathrm{E}}_{x_1^0,\ldots,x_k^0}\left[f(x_1^0,\ldots,x_k^0)\cdot g_{x_1^1,\ldots,x_k^1}(x_1^0,\ldots,x_k^0)\right]\right] = R_k(f) .$$

Therefore we can fix a particular function $g = g_{x_1^1,\ldots,x_k^1}$ such that $\mathrm{E}_{x\in D^k}\left[f(x)\cdot g(x)\right] \geq R_k(f)$.

To conclude the proof, note that g is in Π_k^* because $g(x_1^0,\ldots,x_k^0)$ is the product of factors none of which depends on all the variables x_j^0. □

Next, we state the important observation that the norm is multiplicative for functions over disjoint sets of input variables (cf., [9]).

Fact 3.7. For functions $f : D^k \to \mathbb{C}$ and $f' : (D')^k \to \mathbb{C}$, define the function $(f\cdot f') : (D\times D')^k \to \mathbb{C}$ by

$$\left(f\cdot f'\right)\left((x_1,x_1'),(x_2,x_2'),\ldots,(x_k,x_k')\right) := f(x_1,x_2,\ldots,x_k)\cdot f'(x_1',x_2',\ldots,x_k') .$$

Then $R_k(f\cdot f') = R_k(f)\cdot R_k(f')$.

Using the above results and arguing as for Theorem 1.2 one can prove the following XOR lemma for Π_k^*.

Theorem 3.8 (XOR lemma for Π_k^*)**.** *Let $f : D^k \to \{-1,1\}$ be a function such that $\mathrm{Cor}\left(f,\Pi_k^*\right) \leq \varepsilon$. Then $\mathrm{Cor}\left(f^{\times m},\Pi_k^*\right) \leq \varepsilon^{m/2^k}$.*

3.2 Back to multiparty protocols

In this section we use the results from Section 3.1 to obtain an XOR lemma for multiparty protocols. Let us first recall the model of multiparty protocols. In the *multiparty communication model* there are k parties, each having unlimited computational power, who wish to collaboratively compute a certain function. The input bits to the function are partitioned into k blocks, and the ith party knows all the input bits except those corresponding to the ith block in the partition. The communication between the parties is by "writing on a blackboard" (broadcast): any bit sent by any party is seen by all the others. The parties exchange messages according to a fixed protocol. For each possible sequence of bits that is written on the board so far, the protocol specifies whether the run is over (as a function of the bits on the board), or else which party writes next (as a function of the bits on the board) and what the party writes (as a function of the bits on the board and the partial input seen by that party). The last bit written on the board is the output of the protocol, a value in $\{-1,1\}$. The cost measure of interest is the number of bits c exchanged by the parties. (For background, see the monograph by Kushilevitz and Nisan [26].)

A *c-bit k-party protocol* is a protocol between k parties that prescribes the exchange of at most c bits on any input. For a domain D, we denote by $\Pi_{k,c}$ the class of functions $\pi : D^k \to \{-1,1\}$ computable by c-bit k-party protocols.

We observe that the model Π_k^* can be seen as a special case of k-party k-bit protocols. Specifically, any function in Π_k^* can be computed by a simultaneous protocol (see, e. g., [3, 26]) where each party sends one bit independently from the others, and the output of the protocols is the XOR of these k bits (which, in our $\{-1,1\}$ domain, is the product); the bit sent by the ith party is the value of the function g_i in Definition 3.1. The next lemma shows that in fact the general c-bit model is only stronger than Π_k^* by a factor of 2^c. The same result (for real-valued functions) is implicit in [9, 35] where it is proved by bounding the discrepancy over cylinder intersections. We give a direct proof of the lemma.

Lemma 3.9 ([9, 35])**.** *For every function $f : D^k \to \mathbb{C}$, $\mathrm{Cor}(f,\Pi_{k,c}) \leq 2^c \cdot \mathrm{Cor}(f,\Pi_k^*)$.*

The proof of Lemma 3.9 makes use of the following lemma.

Lemma 3.10 ([5], see also Lemma 6.10 in [26])**.** *Let $\pi : D^k \to \{-1,1\}$ be a function computable by a c-bit k-party protocol. There exists a partition of D^k into 2^c cylinder intersections (see Def. 3.2) $\Gamma_1,\ldots,\Gamma_{2^c}$ such that π is constant over each Γ_ℓ.*

Proof of Lemma 3.9. Let π be a function computed by a c-bit k-party protocol, and let $\Gamma_1,\ldots,\Gamma_{2^c}$ be the cylinder intersections given in Lemma 3.10. The idea in what follows is to define appropriate $-1/1$ random functions that, via averaging, will help us convert a $0/1$ (characteristic) function into a $-1/1$ function. This is beneficial to us because π is naturally written in terms of $0/1$ functions, but our norms require $-1/1$ functions. For any ℓ, j, consider the random function $g_{\ell,j} : D^k \to \{-1,1\}$ defined as $g_{\ell,j}(x) := 1$ with probability 1 if $x \in C_{\ell,j}$, and $g_{\ell,j}(x) := 1$ with probability $1/2$ if $x \notin C_{\ell,j}$ (and consequently $g_{\ell,j}(x) := -1$ also with probability $1/2$ if $x \notin C_{\ell,j}$). Now observe that for every $\ell \leq 2^c$ and every $x \in (\{0,1\}^n)^k$, the expectation

$$\mathop{\mathrm{E}}_{g_{\ell,1},\ldots,g_{\ell,k}}[g_{\ell,1}(x)\cdot g_{\ell,2}(x)\cdots g_{\ell,k}(x)] = \prod_{j\leq k}\mathop{\mathrm{E}}_{g_{\ell,j}}\left[g_{\ell,j}(x)\right]$$

equals 1 if $x \in \Gamma_\ell = C_{\ell,1} \cap \ldots \cap C_{\ell,k}$, and 0 otherwise. Therefore, denoting by $v(\ell) \in \{-1,1\}$ the value of π on inputs in (the cylinder intersection) Γ_ℓ, we can write

$$\pi(x) = \sum_{\ell \leq 2^c} v(\ell) \cdot \mathop{\mathrm{E}}_{g_{\ell,1},\ldots,g_{\ell,k}}\left[\prod_{j\leq k} g_{\ell,j}(x)\right].$$

We now have, by linearity of expectation,

$$\mathop{\mathrm{E}}_x[f(x)\cdot\pi(x)] = \mathop{\mathrm{E}}_{g_{\ell,1},\dots,g_{\ell,k}}\left[\sum_{\ell\le 2^c}\mathop{\mathrm{E}}_x\left[f(x)\cdot v(\ell)\cdot\prod_{j\le k} g_{\ell,j}(x)\right]\right].$$

By fixing the random functions $g_{\ell,j}$ so as to maximize the outermost expectation, we have

$$\mathrm{Cor}(f,\pi) = \left|\mathop{\mathrm{E}}_x[f(x)\cdot\pi(x)]\right| \le 2^c \cdot \max_\ell\left|\mathop{\mathrm{E}}_x\left[f(x)\cdot v(\ell)\cdot\prod_{j\le k} g_{\ell,j}(x)\right]\right| \le 2^c \cdot \mathrm{Cor}(f,\Pi_k^*).$$

□

In particular, combining Lemmas 3.9 and 3.4 we obtain the following corollary.

Corollary 3.11 ([9, 35]). *For every function $f: D^k \to \mathbb{C}$, $\mathrm{Cor}(f,\Pi_{k,c}) \le 2^c \cdot R_k(f)^{1/2^k}$.*

We are now in the position to obtain the following XOR lemma for multiparty communication complexity.

Theorem 1.3 (XOR lemma for multiparty protocols, restated). Let $f: D^k \to \{-1,1\}$ be a function such that $\mathrm{Cor}(f,\Pi_{k,k}) \le \varepsilon$. Then $\mathrm{Cor}(f^{\times m},\Pi_{k,c}) \le 2^c\cdot\varepsilon^{m/2^k}$.

Proof. We have

$$\mathrm{Cor}(f^{\times m},\Pi_{k,c}) \le 2^c \cdot R_k\left(f^{\times m}\right)^{1/2^k} = 2^c \cdot R_k(f)^{m/2^k} \le 2^c\cdot\varepsilon^{m/2^k},$$

where the first inequality holds by Corollary 3.11, the next equality by Fact 3.7, and the last inequality by Lemma 3.6. □

Combining the above XOR lemma with Proposition 1.4 we immediately obtain our direct product lemma for multiparty communication complexity (Corollary 1.7). We repeat the statement for the reader's convenience.

Corollary 1.7 (Direct product lemma for multiparty protocols, restated). Let $f: D \to \{-1,1\}$ be a function such that $\mathrm{Cor}(f,\Pi_{k,k}) \le \varepsilon \le 2^{-(c+1)\cdot 2^k}$. Then $\mathrm{Suc}\left(f^{(m)},\Pi_{k,c}\right) \le 2^{-\Omega(m)}$.

Proof. Proposition 1.4 implies that $\mathrm{Suc}\left(f^{(m)},\Pi_{k,c}\right)$ can bounded from above by $\mathrm{Cor}\left(f^{\times m'},C'\right) + 2^{-\Omega(m)}$, where $m' = m/3$ and C' consists of products of m' $\{-1,1\}$-functions from $\Pi_{k,c}$. Functions in C' can be computed using $m'\cdot c$ communication, simply by computing the m' corresponding functions in $\Pi_{k,c}$ one at the time. Therefore, we obtain $\mathrm{Suc}\left(f^{(m)},\Pi_{k,c}\right) \le \mathrm{Cor}\left(f^{\times m'},\Pi_{k,m'\cdot c}\right) + 2^{-\Omega(m)}$. By Theorem 1.3, we have that $\mathrm{Cor}\left(f^{\times m'},\Pi_{k,m'\cdot c}\right) \le 2^{m'\cdot c}\cdot\varepsilon^{m'/2^k} \le 2^{-m'}$, which gives the result.

□

We conclude this section with a remark on the relative power of the models discussed so far. The observation of Håstad and Goldmann [21, Proof of Lemma 4] shows that Π_k^* is more powerful than degree-$(k-1)$ polynomials over $GF(2)$, while obviously Π_k^* is computable by k-bit k-party protocols. Together with Lemma 3.9 we have the following informal picture:

$$P_{k-1} \subseteq \Pi_k^* \subseteq \Pi_{k,k} \subseteq \Pi_{k,c} \subseteq 2^c \cdot \Pi_k^* .$$

(The first two inclusions above are formally true, as is the next for $c \geq k$, while the last is meant to informally capture Lemma 3.9.) It would be interesting to have a further upper bound in the above sequence in terms of P_d, but it is currently unclear to us if a meaningful bound of this sort exists.

3.2.1 The case of two parties

In this section we further discuss XOR lemmas for the interesting special case of $k = 2$ parties. We start by comparing our results with an XOR lemma by Shaltiel [39], and then we present a counterexample to the "ideal" setting of parameters of the XOR lemma, i. e., going from correlation ε to correlation ε^m.

For $k = 2$, the notion of "cylinder intersections" (Definition 3.2) simplifies to "rectangles," i. e., sets of the form $R = A \times B$ for some $A, B \subseteq \{0,1\}^n$.

Remark 3.12 (Comparison with the XOR lemma by Shaltiel [39])**.** For $k = 2$ parties, Shaltiel proves an XOR lemma which (up to different constants) has the same conclusion as ours (Theorem 1.3) but starts from the assumption that the original function f has bounded *discrepancy* over rectangles (as opposed to bounded correlation with 2-bit protocols in our result). Recall that the discrepancy of a function $f : D \times D \to \{-1,1\}$ is defined as the maximum, over all rectangles R, of

$$\left| \mathop{\mathrm{E}}_{x,y} [f(x,y) | (x,y) \in R] \right| \cdot \Pr[(x,y) \in R] .$$

Shaltiel suggests that the requirement that the discrepancy of f is small is stronger than the requirement that the correlation of f with low-communication protocols is small. However, the discrepancy of f in fact equals the maximum correlation of f with 2-bit protocols (up to a constant factor). To see this, first note that there is always a 2-bit protocol that achieves correlation which is the discrepancy of f. Specifically, let R be the rectangle that maximizes the discrepancy, and consider the protocol where Alice and Bob send two bits to the referee to identify whether $(x,y) \in R$, and then the referee decides according to the bias of f if $(x,y) \in R$, and chooses a random bit otherwise. The correlation of this protocol is exactly the discrepancy of f. (Although the protocol we just defined is randomized, one can obtain a deterministic protocol at least as good by fixing a choice of the random bits that maximizes the correlation.) The converse, i. e., that the discrepancy is an upper bound on the correlation with 2-bit protocols, is standard and can be found, e. g., in the proof of Lemma 2.2 in [5]. Thus, for $k = 2$, our XOR lemma (Theorem 1.3) can be seen as an alternative proof of the XOR lemma by Shaltiel.

It is natural to ask whether the parameters of our XOR lemma (Theorem 1.3) are the best possible. In particular, we would like to know whether the 2^c factor can be eliminated. Although we do not know the answer to this question, we can show a counterexample to the "ideal" setting of parameters, i. e., going from correlation ε to correlation ε^m, for $k = 2$ parties communicating $c = 2$ bits. In the rest of this section we describe this counterexample. First we exhibit a counterexample over the domain $D := \{0,1,2\}$, which was found via brute-force search, then we observe that one can extend it to a counterexample over $D := \{0,1\}^n$.

Claim 3.13. *Let $D := \{0,1,2\}$, and consider the function $f : D^2 \to \{-1,1\}$ defined as $f(x,y) := 1$ if and only if $x = y$.*

1. $\mathrm{Cor}(f,\Pi_{2,2}) \leq 5/9$. $|\mathrm{E}_{x,y}[f(x,y) \cdot \pi(x,y)]| \leq 5/9$.
2. $\mathrm{Cor}(f^{\times 2},\Pi_{2,2}) \geq 33/81 > (5/9)^2$.

Remark 3.14 (Comparison with the counterexample by Shaltiel [39])**.** Shaltiel shows that the XOR lemma for 2-party protocols is false in a strong sense *if one allows for communication $c' = m \cdot c$* to compute m copies of the function. Our result (Claim 3.13) shows that even for the "minimal choice" $c' = c$ some loss occurs (with respect to the "ideal" correlation bound of ε^m).

We now present the proof of Claim 3.13. Although the proof involves a certain amount of calculation, it is perhaps instructive to observe how a 2-bit protocol can correlate with $f^{\times 2}$ in the various cases.

Proof. It is easy to check that $5/9$ is the best correlation of 2-bit protocols with f.

For the second claim, consider the protocol $\pi(x,x',y,y') := f(x,x') \cdot f(y,y')$. Note that this is indeed a 2-bit protocol. Let us compute the probability, over the choice of x,x',y,y', of the event

$$\mathcal{E} := \pi(x,x',y,y') = f(x,y) \cdot f(x',y') .$$

Note that, by definition, $\mathcal{E}$ holds exactly when $f(x,x') \cdot f(y,y') \cdot f(x,y) \cdot f(x',y') = 1$.

Let us condition on the event that $x = x'$ and $y = y'$, which happens with probability $(1/3) \cdot (1/3)$. We have $f(x,x') \cdot f(y,y') \cdot f(x,y) \cdot f(x',y') = 1 \cdot 1 \cdot f(x,y) \cdot f(x,y) = 1$. Thus, $\Pr[\mathcal{E}|x = x' \wedge y = y'] = 1$.

Let us condition on the event that $x \neq x'$ and $y \neq y'$, which happens with probability $(2/3) \cdot (2/3)$. In this case we have

$$f(x,x') \cdot f(y,y') \cdot f(x,y) \cdot f(x',y') = -1 \cdot -1 \cdot f(x,y) \cdot f(x+b,y+b') = f(x,y) \cdot f(x+b,y+b') ,$$

where b and b' are uniform and independent in $\{1,2\}$, and the sum is modulo 3. Thus we are interested in the probability that $f(x,y) = f(x+b,y+b')$ over random x,y,b,b'. Let us now further condition on $x = y$. Then $f(x,y) = 1$ and $f(x+b,y+b') = 1$ if and only if $b = b'$ which happens with probability $1/2$ over the choice of the $b's$. Let us now condition on $x \neq y$, and let us assume in particular that $y = x+1$ (the case $y = x+2$ is analogous). Then $f(x,x+1) = -1$ and $f(x+b,x+1+b') = -1$ if and only if $b \neq 1+b'$ which happens with probability $3/4$ over the choice of the $b's$. Thus,

$$\Pr[\mathcal{E} \mid x \neq x' \wedge y \neq y'] = (1/3)(1/2) + (2/3)(3/4) = 1/6 + 1/2 = 2/3 .$$

Let us condition on the event that $x = x'$ and $y \neq y'$, which happens with probability $(1/3) \cdot (2/3)$. In this case we have $f(x,x') \cdot f(y,y') \cdot f(x,y) \cdot f(x',y') = 1 \cdot -1 \cdot f(x,y) \cdot f(x,y+b)$, where b is uniform in $\{1,2\}$. Thus we are interested in the probability that $-f(x,y) = f(x,y+b)$, which equals the probability that x equals either y or $y+b$, which is $2/3$. Thus,

$$\Pr[\mathcal{E}|x = x' \wedge y \neq y'] = 2/3 \quad \text{and, by symmetry,} \quad \Pr[\mathcal{E}|x \neq x' \wedge y = y'] = 2/3 .$$

Thus

$$\Pr[\mathcal{E}] = (1/3)(1/3)\cdot 1 + (2/3)(2/3)\cdot 2/3 + 2\cdot(1/3)(2/3)\cdot 2/3 = 1/9 + 8/27 + 8/27 = 19/27\,.$$

Therefore $|\mathrm{E}_{x,x',y,y'}\left[f^{\times 2}(x,x',y,y')\cdot\pi(x,x',y,y')\right]| = 2\cdot\Pr[\mathcal{E}]-1 = (38-27)/27 = 11/27 = 33/81.$ □

We now briefly explain how to extend the counterexample in Claim 3.13 to a counterexample in the domain $D := \{0,1\}^n$ (for sufficiently large n). First, consider any domain of the form $D = \{0,1,2,\ldots,3a-1\}$ for some integer $a \geq 1$. It is not hard to see that one can prove the analogous of Claim 3.13 for the function $f : D^2 \to \{-1,1\}$ defined as $f(x,y) := 1$ if and only if $x \equiv y \pmod 3$. Now, consider a domain of the form $\{0,1\}^n$, and let a be the biggest integer such that $3\cdot a < 2^n$. Conditioned on the event that the inputs fall in the set $\{0,\ldots,3a-1\}$, the above counterexample works. Since this event happens with probability approaching 1 (when n grows), the result over the domain $D := \{0,1\}^n$ follows.

3.2.2 The XOR lemma for games is false

In this section we argue that the XOR lemma for games is false. In a single-prover *game*, a verifier chooses a question x according to a publicly known distribution, and sends it to the prover. The prover then responds by $a(x)$, and wins if a publicly known predicate $V(x,a)$ accepts. We are interested in the value of a game, which is the maximum, over all provers, of the probability that the prover wins. For our result it is enough to consider single-prover games, but it will be clear that similar examples exist for any number of provers.

For a game G with acceptance predicate $V(x,a)$ we define the game $G^{\times m}$ as follows: the verifier asks m independent questions $x_1,\ldots,x_m$ and expects m answers $a_1,\ldots,a_m$, where each answer is allowed to depend on all questions $x_1,\ldots,x_m$. The prover wins if and only if the number of indices i such that $V(x_i,a_i)$ accepts is odd.

Claim 3.15 (The XOR lemma for games is false)**.** *There is a single-prover game G that has value at most $3/4$, but such that the value of $G^{\times m}$ approaches 1 as $m\to\infty$.*

Proof. Consider the following game G between a verifier and prover A. The verifier sends two uniform and independent bits (p,t) to A. Prover A then sends one bit $a = a(p,t)$ back to the verifier. If $p = 0$, the verifier accepts iff $a = 1$. If $p = 1$, it accepts iff $t = 1$.

The idea is that A has complete control over the game when $p = 0$, and when $p = 1$, A knows if the game is won or lost (since A knows t). Thus, whenever there is at least one game with $p = 0$, A can win the XOR of the games.

We claim that any prover A wins G with probability at most $3/4$. This is because when $p = 1$ and $t = 0$ the game is lost, no matter what A says.

Now consider the game $G^{\times m}$ and the following prover A: Upon receiving m questions

$$(p_1,t_1),(p_2,t_2),\ldots,(p_m,t_m)\,,$$

A sends back the bits $a_1,\ldots,a_m$ that are all 0 except possibly a_i where i is the least i such that $p_i = 0$, which is set to $a_i := 1\oplus\bigoplus_{i:p_i=1} t_i$. It is easy to see that the prover wins $G^{\times m}$ whenever there is an i such that $p_i = 0$, which happens with probability $1-2^{-m}$. □

3.3 Lower bounds

Using the k-party norm $R_k(\cdot)$, one can give a simple proof of the fact that the *generalized inner product function* is hard to compute with little communication. Babai, Nisan, and Szegedy [5] introduced this function and proved an $\Omega\left(n/4^k\right)$ lower bound for its k-party communication complexity. Chung and Tetali [9] and Raz [35] refined and modularized the technique of [5] and obtained an alternative proof of the same bound for the generalized inner product function.[8] This section can be seen as presenting this alternative proof in a different language. In what follows we denote by $\bigwedge_k : \{0,1\}^k \to \{0,1\}$ the AND function that outputs 1 if all its inputs bits are 1, and 0 otherwise. Let $GIP : (\{0,1\}^n)^k \to \{-1,1\}$ be the function $\left((-1)^{\wedge_k}\right)^{\times n}$, i. e., $GIP(x_1,\dots,x_k) := \prod_{i\le n}(-1)^{\wedge_{j\le k}(x_j)_i}$.

Theorem 3.16 ([5])**.** $\mathrm{Cor}(GIP,\Pi_{k,c}) \le 2^{c-\Omega\left(n/4^k\right)}$.

Proof.

$$\mathop{\mathrm{E}}_{x\in(\{0,1\}^n)^k}[GIP(x)\cdot\pi(x)] \le 2^c\cdot R_k(GIP)^{1/2^k} = 2^c\cdot R_k\left((-1)^{\wedge_k}\right)^{n/2^k} = 2^c(1-2^{-k+1})^{n/2^k},$$

where the first inequality is Corollary 3.11, the next inequality is Fact 3.7, and $R_k\left((-1)^{\wedge_k}\right) = 1-2^{-k+1}$ by straightforward calculation. □

Using the k-party norm, we can prove correlation bounds for variants GIP_m of the above GIP function where the sum is modulo m, as opposed to modulo 2. We note that Grolmusz [18] obtained the corresponding BNS-strength communication complexity lower bound by extending the methods of [5] to the discrepancy of complex-valued functions (namely, the values are m-th roots of unity).

Let $GIP_m : (\{0,1\}^n)^k \to \{-1,1\}$ be the function that equals 1 iff $\sum_{i\le n}\prod_{j\le k}(x_j)_i$ is divisible by m. Similarly to Section 2.3, in the rest of this section we work with correlation with respect to the following non-uniform distribution Q: with probability $1/2$, Q is uniform on the inputs x such that $GIP_m(x)=1$; with probability $1/2$, Q is uniform on the inputs x such that $GIP_m(x)=-1$.

Theorem 3.17. $\mathrm{Cor}_Q(GIP_m,\Pi_{k,c}) \le 2^{c-\alpha\cdot n/4^k}$*, where* $\alpha>0$ *depends on* m *only.*

Proof. Following the proof of Theorem 2.9, we consider the function $f : (\{0,1\}^n)^k \to \mathbb{C}$ defined as $f(x) := e_m\left(\sum_{\ell\le n}\wedge_{j\le k}(x_j)_\ell\right)$, where $e_m(y) := e^{2\pi\cdot\mathbf{i}\cdot y/m}$ and $\mathbf{i}$ is the imaginary unit. By Lemma 2.10, to obtain the claimed bound on the correlation it is enough to bound from above the maximum over $a\in\{1,\dots,m-1\}$ of

$$\left|\mathop{\mathrm{E}}_{x\in(\{0,1\}^n)^k}[f(x)^a\cdot\pi(x)]\right|,$$

where $\pi\in\Pi_{k,c}$.

To bound the above quantity, we use Corollary 3.11 to relate it to the k-party norm of f, and then we use the fact that the norm of the product of functions on disjoint input bits multiplies (Fact 3.7). Thus

[8] While in [9, Theorem 5] the authors claim an $\Omega\left(n/2^k\right)$ lower bound, their proof only reproduces the original $\Omega\left(n/4^k\right)$ bound, which we also obtain here. No better bound is known.

we obtain $\left|\mathrm{E}_{x\in(\{0,1\}^n)^k}\left[f(x)^a\cdot\pi(x)\right]\right| \le R_k\left(e_m\left(a\cdot\wedge_k\right)\right)^{n/2^k}$ and we are left with the task of bounding

$$R_k\left(e_m\left(a\cdot\wedge_k\right)\right) = \mathop{\mathrm{E}}_{\substack{x_1^0,x_2^0,\ldots,x_k^0\in\{0,1\}\\ x_1^1,x_2^1,\ldots,x_k^1\in\{0,1\}}}\left[e_m\left(a\cdot\sum_{\varepsilon_1,\ldots,\varepsilon_k\in\{0,1\}}(-1)^{\sum_\ell\varepsilon_\ell}\wedge_k\left(x_1^{\varepsilon_1},x_2^{\varepsilon_2},\ldots,x_k^{\varepsilon_k}\right)\right)\right].$$

Consider now the event $V :=$ “$x_\ell^0 \neq x_\ell^1$ for every ℓ.” When V happens, there is exactly one choice for the exponents ε_ℓ that gives $\bigwedge_k(x_1^{\varepsilon_1},x_2^{\varepsilon_2},\ldots,x_k^{\varepsilon_k}) = 1$, and that choice is $\varepsilon_\ell := x_\ell^1$ (since the only input that makes $\wedge_k$ equal to 1 is the all 1’s input). Therefore, conditioned on V, the above expectation becomes

$$E_{\substack{x_1^0,x_2^0,\ldots,x_k^0\in\{0,1\}\\ x_1^1,x_2^1,\ldots,x_k^1\in\{0,1\}}}\left[e_m\left(a\cdot(-1)^{\sum_\ell x_\ell^1}\right)\middle|V\right] = \frac{e_m(a)+e_m(-a)}{2} = \Re(e_m(a)) < 1\,,$$

where $\Re(\cdot)$ denotes the real part. Above, the first equality uses the fact that $\sum_\ell x_\ell^1$ is odd with probability $1/2$ (also conditioned on V), while the last inequality uses the fact that $0 < a < m$.

Since V happens with probability 2^{-k}, and when V does not happen the expectation is seen to be 1, we obtain

$$R_k\left(e_m\left(a\cdot\wedge_k\right)\right) = 2^{-k}\cdot\Re(e_m(a)) + 1 - 2^{-k}\,,$$

from which the result follows. □

Acknowledgments. We thank Paul Beame, Ronen Shaltiel, Vladimir Trifonov, and the anonymous referees for helpful comments. We are especially grateful to László Babai for extensive comments which greatly improved the exposition. The first author would like to thank Salil Vadhan for his helpful reading of a preliminary version of this work [43].

References

[1] * NOGA ALON, ODED GOLDREICH, JOHAN HÅSTAD, AND RENÉ PERALTA: Simple constructions of almost k-wise independent random variables. *Random Structures Algorithms*, 3(3):289–304, 1992. [Wiley:10.1002/rsa.3240030308]. 1.2.4, 2.13, 6

[2] * NOGA ALON, TALI KAUFMAN, MICHAEL KRIVELEVICH, SIMON LITSYN, AND DANA RON: Testing low-degree polynomials over GF(2). In *Approximation, randomization, and combinatorial optimization*, volume 2764 of *LNCS*, pp. 188–199. Springer-Verlag, Berlin, 2003. [Springer:5pcg1j8cfl39tmpy]. 1.2, 1.2.1, 2.1, 2.1, 2.5

[3] * LÁSZLÓ BABAI, ANNA GÁL, PETER G. KIMMEL, AND SATYANARAYANA V. LOKAM: Communication complexity of simultaneous messages. *SIAM J. Comput.*, 33(1):137–166, 2003. [SICOMP:10.1137/S0097539700375944]. 3.2

[4] * LÁSZLÓ BABAI, THOMAS P. HAYES, AND PETER G. KIMMEL: The cost of the missing bit: communication complexity with help. *Combinatorica*, 21(4):455–488, 2001. [doi:10.1007/s004930100009]. 1.2.2

[5] * LÁSZLÓ BABAI, NOAM NISAN, AND MÁRIÓ SZEGEDY: Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. Comput. System Sci.*, 45(2):204–232, 1992. [JCSS:10.1016/0022-0000(92)90047-M]. 1.2, 1.2.2, 1.2.2, 1.2.4, 2.4, 3.2, 3.1, 3.10, 3.12, 3.3, 3.16, 3.3

[6] * JEAN BOURGAIN: Estimation of certain exponential sums arising in complexity theory. *C. R. Math.*, 340(9):627–631, 2005. [Elsevier:10.1016/j.crma.2005.03.008]. 1.2, 1.2.4, 4, 2.3

[7] * ASHOK K. CHANDRA, MERRICK L. FURST, AND RICHARD J. LIPTON: Multi-party protocols. In *Proc. 15th STOC*, pp. 94–99, Boston, Massachusetts, 1983. ACM Press. [STOC:800061.808737]. 1.2.2

[8] * ARKADEV CHATTOPADHYAY: An improved bound on correlation between polynomials over Z_m and MOD_q. Technical Report TR06-107, Electronic Colloquium on Computational Complexity, 2006. [ECCC:TR06-107]. 1.2.4

[9] * FAN R. K. CHUNG AND PRASAD TETALI: Communication complexity and quasi randomness. *SIAM J. Discrete Math.*, 6(1):110–123, 1993. [SIDMA:10.1137/0406009]. 1.2.2, 3.1, 3.1, 3.4, 3.5, 3.1, 3.2, 3.9, 3.11, 3.3, 8

[10] * URI FEIGE: Error reduction by parallel repetition-the state of the art. Technical report, Weizmann Science Press of Israel, Jerusalem, Israel, 1995. 1.2.3

[11] * LANCE FORTNOW: *Complexity-theoretic aspects of interactive proof systems*. PhD thesis, Massachusetts Institute of Technology, 1989. Tech Report MIT/LCS/TR-447. 1.2.3

[12] * ODED GOLDREICH AND LEONID A. LEVIN: A hard-core predicate for all one-way functions. In *Proc. 21st STOC*, pp. 25–32, New York, 1989. ACM Press. [STOC:73007.73010]. 1.2.3, 2

[13] * ODED GOLDREICH, NOAM NISAN, AND AVI WIGDERSON: On Yao's XOR lemma. Technical Report TR95-050, Electronic Colloquium on Computational Complexity, March 1995. [ECCC:TR95-050]. 1.1, 1.2.3

[14] * W. T. GOWERS: A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998. [Springer:lg2rlw8pvtt2x0qj]. 1.2, 1.2.1

[15] * W. T. GOWERS: A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001. [Springer:00622770r8437760]. 1.2, 1.2.1, 2.1, 2.3

[16] * BEN GREEN AND TERENCE TAO: An inverse theorem for the Gowers U^3 norm, 2005. arXiv.org:math/0503014. [arXiv:math/0503014]. 1.2, 1.2.1, 2.1, 2.3, 5, 2.1

[17] * FREDERIC GREEN, AMITABHA ROY, AND HOWARD STRAUBING: Bounds on an exponential sum arising in Boolean circuit complexity. *C. R. Math.*, 341(5):279–282, 2005. [Elsevier:10.1016/j.crma.2005.07.011]. 1.2.4, 4

[18] * VINCE GROLMUSZ: Separating the communication complexities of $\bmod m$ and $\bmod p$ circuits. *J. Comput. System Sci.*, 51(2):307–313, 1995. [JCSS:10.1006/jcss.1995.1069]. 1.2.2, 3.3

[19] * DAN GUTFREUND AND EMANUELE VIOLA: Fooling parity tests with parity gates. In *Proc. 8th Intern. Workshop on Randomization and Computation (RANDOM'08)*, volume 3122 of *LNCS*, pp. 381–392. Springer-Verlag, 2004. [Springer:x9px6h8l0tb6et6b]. 2.15

[20] * ANDRÁS HAJNAL, WOLFGANG MAASS, PAVEL PUDLÁK, MÁRIÓ SZEGEDY, AND GYÖRGY TURÁN: Threshold circuits of bounded depth. *J. Comput. System Sci.*, 46(2):129–154, 1993. [JCSS:10.1016/0022-0000(93)90001-D]. 1.1

[21] * JOHAN HÅSTAD AND MIKAEL GOLDMANN: On the power of small-depth threshold circuits. *Comput. Complexity*, 1(2):113–129, 1991. [CC:r0mv45x710nn1q76]. 1.2.2, 2.4, 3.2

[22] * ALEXANDER HEALY: Randomness-efficient sampling within NC^1. In *Proceedings of the 10th International Workshop on Randomization and Computation (RANDOM'06)*, volume 4110 of *LNCS*, pp. 398–409. Springer-Verlag, 2006. [Springer:b773545612310728]. 2.15

[23] * RUSSELL IMPAGLIAZZO: Hard-core distributions for somewhat hard problems. In *Proc. 36th FOCS*, pp. 538–545, Los Alamitos, CA, USA, 1995. IEEE Computer Society. [FOCS:10.1109/SFCS.1995.492584]. 1.1

[24] * RUSSELL IMPAGLIAZZO AND AVI WIGDERSON: $P = BPP$ if E requires exponential circuits: Derandomizing the XOR lemma. In *Proc. 29th STOC*, pp. 220–229, New York, 1997. ACM Press. [STOC:258533.258590]. 1.1, 1.5

[25] * CHARANJIT S. JUTLA, ANINDYA C. PATTHAK, ATRI RUDRA, AND DAVID ZUCKERMAN: Testing low-degree polynomials over prime fields. In *Proc. 45th FOCS*, pp. 423–432, Los Alamitos, CA, USA, 2004. IEEE Computer Society. [FOCS:10.1109/FOCS.2004.64]. 2.1, 2.5

[26] * EYAL KUSHILEVITZ AND NOAM NISAN: *Communication complexity*. Cambridge University Press, Cambridge, 1997. 1.2.2, 3.1, 3.2, 3.10

[27] * LEONID A. LEVIN: One way functions and pseudorandom generators. *Combinatorica*, 7(4):357–363, 1987. [Springer:e1415188r28663m5]. 1.1

[28] * MICHAEL LUBY, BOBAN VELICKOVIC, AND AVI WIGDERSON: Deterministic approximate counting of depth-2 circuits. In *Proc. 2nd Israeli Symp. on Theoretical Computer Science (ISTCS'93)*, pp. 18–24, Los Alamitos, CA, USA, 1993. IEEE Computer Society. 1.1

[29] * J. NAOR AND M. NAOR: Small-bias probability spaces: efficient constructions and applications. In *Proc. 22nd STOC*, pp. 213–223. ACM Press, 1990. [STOC:100216.100244]. 1.2.4, 2.13

[30] * NOAM NISAN: Pseudorandom bits for constant depth circuits. *Combinatorica*, 11(1):63–70, 1991. [Springer:g79x907l52546012]. 1.1

[31] * NOAM NISAN, STEVEN RUDICH, AND MICHAEL SAKS: Products and help bits in decision trees. *SIAM J. Comput.*, 28(3):1035–1050, 1999. [SICOMP:10.1137/S0097539795282444]. 1.2.3

[32] * NOAM NISAN AND AVI WIGDERSON: Hardness vs. randomness. *J. Comput. System Sci.*, 49(2):149–167, October 1994. [JCSS:10.1016/S0022-0000(05)80043-1]. 1.1

[33] * ITZHAK PARNAFES, RAN RAZ, AND AVI WIGDERSON: Direct product results and the GCD problem, in old and new communication models. In *Proc. 29th STOC*, pp. 363–372, New York, 1997. ACM Press. [STOC:258533.258620]. 1.2.3, 1.2.3, 1.2.3

[34] * RAN RAZ: A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998. [SICOMP:10.1137/S0097539795280895]. 1.2.3, 1.2.3

[35] * RAN RAZ: The BNS-Chung criterion for multi-party communication complexity. *Comput. Complexity*, 9(2):113–122, 2000. [CC:u8q21j1ccvltrb40]. 1.2.2, 3.1, 3.1, 3.4, 3.5, 3.2, 3.9, 3.11, 3.3

[36] * ALEXANDER A. RAZBOROV: Lower bounds on the dimension of schemes of bounded depth in a complete basis containing the logical addition function. *Mat. Zametki*, 41(4):598–607, 623, 1987. 1.2.1, 2.15

[37] * ALEX SAMORODNITSKY: Low-degree tests at large distances. In *Proc. 39th STOC*, pp. 506–515, New York, 2007. ACM Press. [STOC:1250790.1250864]. 1.2.1, 2.1, 2.1, 2.6

[38] * ALEX SAMORODNITSKY AND LUCA TREVISAN: Gowers uniformity, influence of variables, and PCPs. In *Proc. 38th STOC*, pp. 11–20, New York, May 2006. ACM Press. [STOC:1132516.1132519]. 1.2.1, 2.1

[39] * RONEN SHALTIEL: Towards proving strong direct product theorems. *Comput. Complexity*, 12(1-2):1–22, 2003. [CC:ku74rl1ga9te5lpe]. 1.2.2, 3.2.1, 3.12, 3.14

[40] * RONEN SHALTIEL AND CHRISTOPHER UMANS: Simple extractors for all min-entropies and a new pseudorandom generator. *J. ACM*, 52(2):172–216, 2005. [JACM:1059513.1059516]. 1.1

[41] * RONEN SHALTIEL AND EMANUELE VIOLA: Hardness amplification proofs require majority. In *Proc. 40th STOC*, pp. 589–598, Victoria, Canada, 2008. ACM Press. [STOC:1374376.1374461]. 1.1, 1.2.3

[42] * ROMAN SMOLENSKY: Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proc. 19th STOC*, pp. 77–82, New York, 1987. ACM Press. [STOC:28395.28404]. 2.15

[43] * EMANUELE VIOLA: New correlation bounds for GF(2) polynomials using Gowers uniformity. Technical Report TR06-097, Electronic Colloquium on Computational Complexity, 2006. [ECCC:TR06-097]. 1.2.4, 3.3

[44] * EMANUELE VIOLA: Pseudorandom bits for constant-depth circuits with few arbitrary symmetric gates. *SIAM J. Comput.*, 36(5):1387–1403, 2007. [SICOMP:10.1137/050640941]. 1.1

[45] * ANDREW CHI-CHIH YAO: Some complexity questions related to distributive computing. In *Proc. 11th STOC*, pp. 209–213, New York, 1979. ACM Press. [STOC:800135.804414]. 1.2.2

[46] * ANDREW CHI-CHIH YAO: Theory and applications of trapdoor functions (extended abstract). In *Proc. 23rd FOCS*, pp. 80–91, Los Alamitos, CA, USA, 1982. IEEE Computer Society. 1.1

AUTHORS

Emanuele Viola
postdoctoral fellow
Columbia University, New York, NY
viola@cs.columbia.edu
http://www.cs.columbia.edu/~viola/

Avi Wigderson
professor
Institute for Advanced Study, Princeton, NJ
avi@ias.edu
http://www.math.ias.edu/~avi/

ABOUT THE AUTHORS

EMANUELE VIOLA graduated from Harvard University in 2006 under the supervision of Salil Vadhan. At the time of this submission he was a postdoctoral fellow at the Institute for Advanced Study in Princeton, NJ.

AVI WIGDERSON complains that it will be very difficult to be original here with every submission. “To lower the standards, let me mention the following two personal items. First, one of my favorite poems is ‘Stopping By Woods On A Snowy Evening’ by Robert Frost. Second, I recently saw a wonderful performance of ‘King Lear’ in the Kirby Shakespeare Theater in Madison, NJ, with a superb Daniel Davis as Lear.”