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

Emanuele Viola* Avi Wigderson†

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

Abstract

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

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

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

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

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

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

1. Introduction

1.1. Background

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

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

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

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

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

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

f cannot be computed exactly by any function which is the majority of any t functions from C [17]). For these 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 [42] (cf. [10]). One starts with a function f of nontrivial correlation with C, and constructs a new function $f^{\oplus m}$ (on $n \cdot m$ bits), which is the exclusive-OR of the value of f on m independent inputs. The hope is that the correlation 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, and f any function with $|E_x[f(x)]| = \mathrm{Cor}(f, C) = \epsilon$. Then it is easy to see that $\mathrm{Cor}(f^{\oplus m}, C) = \epsilon^m$ for every m.[1] So the decay of the correlation in this trivial scenario is purely exponential in the number of copies m.

Yao's XOR lemma deals with the most studied computational model, namely polynomial-size circuits, and goes as follows. Let C be the class of Boolean circuits of size s, and let f be any function on n bits with $\mathrm{Cor}(f, C) \leq \epsilon$. Then for any large m and small $\alpha > 0$, if C' is the class of circuits of size $s \cdot (\alpha/nm)^2$ then $\mathrm{Cor}(f^{\oplus m}, C') \leq \epsilon^m + \alpha$. Many proofs of this XOR lemma were given, starting with Levin [24, 20, 10, 21]. All in fact show that that this lemma holds in more general circumstances, namely as long as C can compute the majority of functions in C'. 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 (i.e. prove lower bounds), such as low-degree $GF(2)$ polynomials, multiparty protocols, or constant-depth circuits (cf. [39, Chapter 6]). 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 [24, 20, 10, 21] is their loss in resources (e.g., circuit size) in C' with respect to C (cf. [10]).

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. To each of these complexity 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 COMMUTES WITH XOR. Let f, g be two functions on *disjoint* inputs, then $N(f \cdot g) = N(f) \cdot N(g)$.

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

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

Of course, the challenge is to find the appropriate norm and prove their properties. As it turns out, much of this work was done already.

1.2.1. $GF(2)$ polynomials

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 captures important complexity classes. 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 [33] in his lower bound for unbounded fan-in polynomial-size constant-depth circuits over $GF(2)$. Razborov proved that any function f 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 proved that $\mathrm{Cor}(\mathrm{Majority}, P_d) \leq O(1/\sqrt{n})$ for such d, and the quest for finding functions of smaller correlation with that class has been a challenge that was not yet answered. Specifically, no explicit function is known which has correlation at most $1/n$ even 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 $\log n$: We prove that the correlation of the XOR of m copies decays exponentially with $m/2^d$.

Theorem 1.1 (XOR lemma for $GF(2)$ polynomials)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$ be a function such that $\mathrm{Cor}(f, P_d) \leq 1 - 1/2^d$. Then $\mathrm{Cor}(f^{\oplus m}, P_d) \leq \exp\left(-\Omega\left(m/\left(4^d \cdot d\right)\right)\right)$.*

No XOR lemma was 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 [11, 12] and independently by Alon et al. [2]. We note that its relationship to the class P_d was already used for many purposes. Gowers [11, 12] used it to give sharper bounds in Szemeredi's Theorem on arithmetic progressions in subsets of the integers. Green and Tao [13] further found more applications in arithmetic combinatorics. Alon et al. [2] used it for property testing of low-degree polynomials. Finally, Samorodnitsky and Trevisan [35, 34] 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 [13], [2], and in [34].

[1] Strictly speaking, now C denotes the constant functions on $n \cdot m$ bits.

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

1.2.2. Multiparty protocols

In Yao's standard 2-party communication complexity [43], the two parties are each holding an 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 comprehensive book [23]). This model has been one of the most extensively studied in computational complexity, and captures essential features of many diverse computational settings, from Turing machines, VLSI and distributed computation, to linear programming and auctions. Many techniques for proving various strong lower bounds and correlation bounds are known.

This model was generalized by Chandra, Furst, and Lipton [5], to the *multiparty model* (often called "number-on-forehead" or NOF model). In k-party communication complexity each party is assigned an input again. However, that input (figuratively) resides on that party's forehead, and so (formally) each party knows *all but* his 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). No explicit function is known which cannot be computed using $k = \log_2 n$ parties and $c = \log_2 n$ communication.

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 [18, Proof of Lemma 4] shows that any degree-d $GF(2)$ polynomial 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 implies breaking the $\log n$ barrier for $GF(2)$ polynomials. 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 1.2 (XOR lemma for multiparty protocols)**.** *Let $f : D^k \to \{-1, 1\}$ be a function such that $\mathrm{Cor}\left(f, \Pi_{k,k}\right) \leq \epsilon$. Then $\mathrm{Cor}\left(f^{\oplus m}, \Pi_{k,c}\right) \leq 2^c \cdot \epsilon^{m/2^k}$.*

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 [36]; cf. Remark 3.6.

Note that in the hypothesis of Theorem 1.2 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. Remark 3.5. 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. taking correlation ϵ to correlation ϵ^m, is actually *false* for $k = 2$ and $c = 2$ (Claim 3.7). We believe it is an interesting question to understand what the correct bound is.

The norm we use to prove this XOR lemma is the one supplied (indirectly or directly) in the lower bound proofs for this model [3, 7, 32]. In particular, [7] shows that this norm *upper* bounds the correlation (which proves one direction of Property 1 in Section 1.2), and they also observe that it commutes with XOR (which proves Property 2 in Section 1.2). With this work in place, we need only show that this norm *lower* bounds the correlation too (which proves the other direction of Property 1 in Section 1.2). We also give a somewhat more direct proof that this norm upper bounds the correlation, and extend the norm to complex-valued functions, obtaining new correlation bounds for unbalanced functions.

1.2.3. Direct product vs. XOR lemmas

XOR lemmas are intimately related to *direct product* lemmas. Here we start again with a function $f : D \to \{-1, 1\}$ that cannot be perfectly computed by some complexity class C, and want to amplify its hardness by taking many copies of it on independent inputs. However, rather than requiring to compute only the XOR of all outputs, we simply require to compute *all* outputs. In other words, the new function $f^{(m)} : D^m \to \{-1, 1\}$ is the concatenation of m copies of f, $f^{(m)}(x_1, x_2, \ldots, x_m) := (f(x_1), f(x_2), \ldots, f(x_m))$. 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 $p_1, \ldots, p_m \in C$, of the probability over $x \in D^m$ that $f^{(m)}(x) = (p_1(x), p_2(x), \ldots, p_m(x))$.

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 products lemmas are known for several models. For Boolean decision trees, [28] show that the success proba-

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

bility of computing $f^{(m)}$ using decision trees of depth d decays purely exponentially with m (independently of d). For c-bit 2-party protocols, [30] prove a decay of the form $\epsilon \to (1/2+\epsilon/2)^{\Omega(m/c)}$, which mildly deteriorates with the communication complexity c. This bound is proved using (and somewhat extending and strengthening) the celebrated parallel repetition of Raz [31].

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

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

Proposition 1.1 (XOR lemma implies direct product lemma)**.** *Let f be any $\{-1,1\}$-valued function and C any class of $\{-1,1\}$-valued functions that is closed under projections (i.e. under fixing some of the inputs). Then* $\mathrm{Suc}\left(f^{(m)}, C\right) \leq \mathrm{Cor}\left(f^{\oplus m'}, C'\right) + 2^{-\Omega(m)}$*, where $m' = m/3$ and C' consists of products of m' functions from C.*

Proof. Let $p_1, \ldots, p_m \in C$ be such that with probability ϵ over $x = (x_1, \ldots, x_m)$ we have $f(x_i) = p_i(x)$ for every i. Let us choose Z uniformly in $\{0,1\}^m$, and note that

$$\epsilon = E_{Z,x_1,x_2,\ldots,x_m}\left[\prod_{i \leq m} \left(f(x_i) \cdot p_i(x)\right)^{Z_i}\right],$$

where Z_i is the i-th bit of Z. To see this last equality, observe that if even for one i we have $f(x_i) \neq p_i(x)$, then $f(x_i) \cdot p_i(x) = -1$ and the contribution to the expectation (over the choice of Z) is zero. By a Chernoff bound we can fix $z = Z$ of Hamming weight at least $m/3$ such that $E_{x_1,x_2,\ldots,x_m}\left[\prod_{i \leq m} \left(f(x_i) \cdot p_i(x)\right)^{z_i}\right] \geq \epsilon - 2^{-\Omega(m)}$. The result follows by fixing the values of the x_i's corresponding to $z_i = 0$ so as to maximize the expectation, showing that the XOR of the function in the remaining $m/3$ inputs has the same correlation with a XOR of $m/3$ functions in C. □

We note that the argument in the proof of Proposition 1.1 simplifies and strengthens a result by Impagliazzo and Wigderson [21, Theorem 11] which is about the special case $m' = 1$ (i.e., computing f): In the proof of Proposition 1.1, fixing any $z \in \{0,1\}^m, z \neq 0$, gives correlation $\epsilon - 2^{-m}$ with f, whereas in [21] they obtain correlation $\epsilon - O(\sqrt{m} \cdot 2^{-m})$.

Combining the above Proposition 1.1 with our XOR lemma for $GF(2)$ polynomials we obtain a direct product lemma for $GF(2)$ polynomials (with no loss in the degree).

Corollary 1.1 (Direct product lemma for $GF(2)$ polynomials)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$ be a function such that* $\mathrm{Cor}\left(f, P_d\right) \leq 1 - 1/2^d$*. 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 1.1, 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 $\epsilon \ll 2^{-c \cdot 2^k}$.

Corollary 1.2 (Direct product lemma for multiparty protocols)**.** *Let $f : D \to \{-1,1\}$ be a function such that* $\mathrm{Cor}\left(f, \Pi_{k,k}\right) \leq \epsilon \leq 2^{-(c+1)\cdot 2^k}$*. Then* $\mathrm{Suc}\left(f^{(m)}, \Pi_{k,c}\right) \leq 2^{-\Omega(m)}$.

The above corollary, in its range of parameter $\epsilon \ll 2^{-c \cdot 2^k}$, beats the bound for 2-party protocols in [30] 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. Remark 3.5).

From direct product to XOR. Connections are also known in the other direction: The seminal Goldreich-Levin theorem [9] shows that if a *circuit* can obtain correlation ϵ with $f^{\oplus m}$, then a slightly larger circuit can succeed in computing $f^{(m)}$ correctly with probability $\mathrm{poly}(\epsilon)$ (cf. [10]). However, this reduction suffers again from the problems discussed at the end of Section 1.1: It cannot be applied to any model for which we can currently prove lower bounds (cf., [39, Chapter 6]). Because of this fact, the direct product lemma for 2-party protocols in [30] 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 cooperating provers (see, e.g., [8]). For 2 parties, Raz [31] proved an essentially tight direct product lemma. In this work we show that the XOR lemma for games is false in a strong sense. Specifically, we show a very simple game G for which any prover strategy has correlation at most $1/2$, but on the other hand there is a prover strategy that has correlation $1 - 1/2^m$ with $G^{\oplus m}$.

Equivalence of direct product and XOR lemmas for circuits. Although in this paper we mainly apply Proposition 1.1 to the models C of low-degree $GF(2)$ polynomials 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 [9] discussed above, we now have the following equivalence.

Corollary 1.2 (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. $\mathrm{Suc}\left(f^{(m)}, C(s)\right) \leq \mathrm{Cor}\left(f^{\oplus m'}, C(s')\right) + 2^{-\Omega(m)}$, *where $m' = m/3$ and $s' = O(s \cdot m')$ (Prop. 1.1),*
2. *and* $\mathrm{Cor}\left(f^{\oplus m}, C(s)\right) \leq \left(n \cdot \mathrm{Suc}\left(f^{(m)}, C(s')\right)\right)^{\Omega(1)}$, *where $s' = s \cdot \mathrm{poly}(n/\mathrm{Cor}\left(f^{\oplus m}, C(s)\right))$ ([9]).*

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

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

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 $GF(2)$ polynomials, 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 $GF(2)$ polynomial can be (essentially) upper bounded 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 prove that the Mod_m function on n bits, defined as $Mod_m(x_1 x_2 \ldots x_n) = 1$ iff $\sum_i x_i \equiv 0 \pmod m$, has correlation[4] at most $\exp\left(-\Omega\left(n/4^d\right)\right)$ with any $GF(2)$ polynomial of degree d, for any fixed odd integer m. A correlation bound of $\exp\left(-\Omega\left(n/8^d\right)\right)$ was first proved in a breakthrough result by Bourgain [4][5]. Subsequently to our results [40], Chattopadhyay [6] 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 [4, 14, 6]. 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 in [4, 14, 6] generalize to polynomials modulo q for arbitrary q relatively prime to m, as opposed to $q = 2$ in this work. It is not clear to us how to generalize the techniques in this work to any $q \neq 2$.

(2) We exhibit a polynomial-time computable function on n bits that has correlation at most $\exp\left(-\Omega\left(n/2^d\right)\right)$ with any $GF(2)$ polynomial of degree d. Previous to our work 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 [3] 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 with high probability degree-d norm that is exponentially small (i.e., $\exp(-\Omega(n))$). 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* [26, 1]. Such a sample space F_s can be generated using only $O(n)$ random bits s, which we can include as part of the input to our function. Thus, we obtain that the function $f(s,x) := F_s(x)$ has correlation at most $\exp\left(-\Omega\left(n/2^d\right)\right)$ with any $GF(2)$ polynomial of degree d. In particular, using a construction in [1], we obtain 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 $GF(2)$ polynomials, 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.

2. $GF(2)$ polynomials

2.1. Degree-k norm

In this section we discuss the degree-k norm. For our results, we need to work with both real-valued and complex-valued functions. We denote the complex conjugate of a complex number $a + ib$ by $\overline{a + ib} := a - ib$, and its norm by $|a + ib| := \sqrt{a^2 + b^2}$. It is also convenient to use the following notation.

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

We now define the degree-k norm of a function. Although this is syntactically defined as the expectation of

[4] 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 defines the correlation between f and p as $|\Pr_{x:f(x)=1}[p(x) = 1] - \Pr_{x:f(x)=-1}[p(x) = 1]|$.

[5] Bourgain's proof [4] contains all the main ideas but is slightly incorrect. A correct proof is given by Green et al. [14].

a complex-valued random variable, it is always a non-negative real number (cf. [35]).

Definition 2.2 (Degree-k norm[6]). *Let $f : \{0,1\}^n \to \mathbb{C}$ be a function. The* degree-k norm *of f is defined as $U_k(f) :=$*

$$E_{y_1,y_2,\ldots,y_k,x\in\{0,1\}^n}\left[\prod_{S\subseteq[k]} f\left(x\oplus\bigoplus_{j\in S} y_j\right)^{|S|}\right],$$

where '$\oplus$' denotes bit-wise XOR.

The following lemma lets us upper bound the correlation of f with any low-degree polynomial by the degree norm of f. Before stating the lemma, let us discuss some notation. First, we state the lemma for complex-valued functions; the correlation $\mathrm{Cor}(f,p)$ between such a function and a polynomial p is defined in the natural way, i.e. it is the (complex) norm of the complex value $E_x[f(x)\cdot p(x)]$ (cf. Section 1.1). Also, in this work it is convenient to think of a $GF(2)$ polynomial p as a function from $\{0,1\}^n$ to $\{-1,1\}$. For example, $p(x_1x_2x_3) := (-1)^{x_1\cdot x_2+x_3}$, where $x_i \in \{0,1\}$, is a $GF(2)$ polynomial mapping $\{0,1\}^3$ to $\{-1,1\}$. In this notation, a product of functions captures their exclusive-or in the $0/1$ notation.

Lemma 2.3 ([13]). *For every function $f : \{0,1\}^n \to \mathbb{C}$, $\mathrm{Cor}(f,P_d) \le U_{d+1}(f)^{1/2^{d+1}}$.*

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

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

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

$$\begin{aligned}\mathrm{Cor}(f,P_d) \le |E_x[h(x)]| &= \sqrt{U_1(h)} \le U_2(h)^{1/2^2}\\ &\le \ldots \le U_{d+1}(h)^{1/2^{d+1}} = U_{d+1}(f)^{1/2^{d+1}}.\end{aligned}$$

We now explain how one obtains the above facts. Fact (1) follows from the definition: $|E_x[h(x)]| = \sqrt{E_{x,y}\left[h(x)\cdot\overline{h(x\oplus y)}\right]} = \sqrt{U_1(h)}$.

Fact (2) amounts to writing $U_k(h)$ in an appropriate form and then using the inequality $E\left[|Z|^2\right] \ge |E[Z]|^2$. Due to space restrictions we omit the standard derivation of this inequality which can be found in [12, Lemma 3.8] and in the full version of this paper.

Fact (3) follows from the fact that for every $GF(2)$ polynomial $p(x)$ 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 facts above are stated in [13, Equations 1.1, 1.2, and 2.1] for $U_k(h)^{1/2^k}$. □

The above lemma shows that the degree norm upper bounds the correlation with low-degree $GF(2)$ polynomials. 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 $E_x[f(x)\cdot p(x)] \le \epsilon$ for every degree-d polynomial $p : \{0,1\}^n \to \{-1,1\}$ then their procedure rejects with probability $\Omega\left(\min\left\{2^d(1-\epsilon), 1/(d\cdot 2^d)\right\}\right)$. As noted in [34], the rejection probability of their procedure is $(1-U_{d+1}(f))/2$. Thus we have the following lemma (stated in [22, Theorem 4.1] but essentially proved in [2]).

Lemma 2.4 ([2, 22]). *Let $f : \{0,1\}^n \to \{-1,1\}$ be a function such that $\mathrm{Cor}(f,P_d) \le \epsilon$. Then $U_{d+1}(f) \le 1 - \Omega\left(\min\left\{2^d(1-\epsilon), 1/(d\cdot 2^d)\right\}\right)$.*

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 [34] improved this dependence in the special case of quadratic polynomials (i.e., $d=2$).

Lemma 2.5 ([34]). *Let $f : \{0,1\}^n \to \{-1,1\}$ be a function such that $\mathrm{Cor}(f,P_2) \le \epsilon$. Then $U_3(f) \le \epsilon'$, where $\epsilon' \le \log^{-\Omega(1)}(1/\epsilon)$.*

Another key property for us of the norm is the fact, formally stated next, that the norm of the product of two functions defined on disjoint input bits is the product of the norms of the two functions. The proof of this fact is immediate from the definition.

Fact 2.6. *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 $GF(2)$ polynomials

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

[6] The degree-d norm is indeed a norm when raised to the power of $1/2^d$; see, e.g., [13].

Theorem 1.1 (XOR lemma for $GF(2)$ polynomials)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$ be a function such that $\mathrm{Cor}(f, P_d) \leq 1 - 1/2^d$. Then $\mathrm{Cor}(f^{\oplus m}, P_d) \leq \exp\left(-\Omega\left(m/\left(4^d \cdot d\right)\right)\right)$.*

Proof. We have

$$\begin{aligned} E_x\left[f^{\oplus m}(x) \cdot p(x)\right] &\leq U_{d+1}\left(f^{\oplus m}\right)^{1/2^{d+1}} \\ &= U_{d+1}\left(f\right)^{m/2^{d+1}} \leq \left(1 - 1/\left(2^d \cdot d\right)\right)^{m/2^{d+1}} \\ &\leq 2^{-\Omega\left(m/\left(4^d \cdot d\right)\right)}, \end{aligned}$$

where the first inequality holds by Lemma 2.3, the next equality by Fact 2.6, and the next inequality by Lemma 2.4. □

Note that if the initial correlation is $\epsilon \geq 1 - 1/\left(d \cdot 4^d\right)$, then in fact we can obtain an XOR lemma with the 'correct' dependence on ϵ, namely $\exp\left(-\Omega\left(m \cdot (1-\epsilon)\right)\right) \approx \epsilon^m$ (we did not state this in the theorem for simplicity). However, if the initial correlation is $\epsilon \leq 1 - 1/\left(d \cdot 4^d\right)$, 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 2.4 and reasoning as in the proof Theorem 1.1, we obtain the following XOR lemma for quadratic $GF(2)$ polynomials.

Theorem 2.7 (XOR lemma for quadratic $GF(2)$ polynomials)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$ be a function such that $\mathrm{Cor}(f, P_2) \leq \epsilon$. Then $\mathrm{Cor}(f^{\oplus m}, P_2) \leq (\epsilon')^m$, where $\epsilon' \leq \log^{-\Omega(1)}(1/\epsilon)$.*

As discussed in Section 1.2.3, combining Theorem 1.1 with Proposition 1.1 we immediately obtain a direct product lemma for low-degree polynomials over $GF(2)$.

Corollary 1.1 (Direct product lemma for $GF(2)$ polynomials)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$ be a function such that $\mathrm{Cor}\left(f, P_d\right) \leq 1 - 1/2^d$. Then $\mathrm{Suc}\left(f^{(m)}, P_d\right) \leq \exp\left(-\Omega\left(m/\left(4^d \cdot d\right)\right)\right)$.*

Similarly, one can obtain a direct product lemma for quadratic $GF(2)$ polynomials by combining Proposition 1.1 with Theorem 2.7.

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

In this section we study the correlation of low-degree $GF(2)$ polynomials with the function $Mod_m : \{0,1\}^n \to \{-1,1\}$, for odd $m \geq 3$, where $Mod_m(x_1, x_2, \ldots, x_n)$ 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 defines the correlation between f and p as

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

In this section correlation always means correlation in the above sense.

Theorem 2.8. *For any odd m, $\mathrm{Cor}(Mod_m, P_d) \leq \exp\left(-\alpha \cdot n/4^d\right)$, where $\alpha > 0$ is a constant that depends on m only.*

Proof. To model the Mod_m function, define $f : \{0,1\}^n \to \mathbb{C}$ as $f(x_1, \ldots, x_n) := e_m\left(\sum_i x_i\right) = \prod_i e_m\left(x_i\right)$, where $e_m\left(y\right) := e^{2\pi \cdot i \cdot y/m}$. As shown in [4], the correlation between a $GF(2)$ polynomial $p(x) : \{0,1\}^n \to \{-1,1\}$ of degree d and the Mod_m function can be bound from above by the maximum over $a \in \{1, \ldots, m-1\}$ of

$$\left|E_{x \in \{0,1\}^n}\left[f(x)^a \cdot p(x)\right]\right|, \tag{1}$$

up to a factor $O(m)$ and a term $2^{-\epsilon \cdot n}$ for a constant $\epsilon > 0$ which depends only on m. To bound the above quantity (1), we use Lemma 2.3 to relate it 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 2.6). Formally, letting $k := d + 1$, we obtain:

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

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

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

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

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

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 $y_1 = y_2 = \ldots = y_k = 1$ with probability 2^{-k}, we have, letting $\delta := \Re\left(e_m\left(a \cdot 2^{k-1}\right)\right)$:

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

Putting everything together, we obtain

$$\left|E_{x\in\{0,1\}^n}\left[f(x)^a \cdot p(x)\right]\right| \leq \left(1 - \frac{1-\delta}{2^k}\right)^{\frac{n}{2^k}} < e^{-\frac{(1-\delta)n}{2^{2k}}},$$

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

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

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

Theorem 2.9. *There is a polynomial-time computable function $f : \{0,1\}^n \to \{-1,1\}$ such that $\mathrm{Cor}(f, P_d) \leq \exp\left(-\alpha \cdot n/2^d\right)$ for every $d < n/2$, where $\alpha > 0$ is a universal constant.*

In [3] they prove a $\exp\left(-\alpha \cdot n/(d \cdot 2^d)\right)$ bound in the stronger computational model of $(d+1)$-party protocols. Their proof and ours are very similar: Both exploit a property of the target function which is captured in Lemma 2.10 below. Our main contribution is to show that using the degree-norm one obtains a 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 x-th output bit of a small-bias generator on seed σ. The following lemma summarizes the definition and the existence of small-bias generators.

Lemma 2.10 ([26, 1][7])**.** *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:*

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

Let f be the function in Lemma 2.10 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.

[7] 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 x-th bit of the sample that corresponds to α.

Claim 2.11. *$E_\sigma\left[U_k\left(f_\sigma\right)\right] \leq 2^{-\alpha\cdot n}$, for every $k \leq n/2$, where $\alpha > 0$ is a universal constant.*

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:

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

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 2.10. 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

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

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 $GF(2)$ polynomial of degree d, and notice that

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

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 2.11. □

Remark 2.12 (On the tightness of Theorem 2.9)**.** *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 2.9, which recall computes the x-th 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 $GF(2)$ polynomial 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 [16, 19][8] and the well-known fact that any such function has correlation at least $1 - o(1)$ with some $GF(2)$ polynomial of degree $\log^{O(1)} n$ [33, 38].

3. Multiparty protocols

In this section we discuss our results on multiparty protocols. As we will see, some of the results were already obtained in the literature using the notion of discrepancy, which is an important tool in proving multiparty lower bounds (cf. [23]). Our approach avoids discrepancy and seems to us more direct.

3.1. k-party norm

In this section we discuss the k-party norm (recall Notation 2.1).

Definition 3.1 (k-party norm)**.** *Let $f : D^k \to \mathbb{C}$ be a function. The* k-party norm *of f is defined as $R_k(f) :=$*

$$E_{\substack{x_1^0,\ldots,x_k^0 \in D \\ x_1^1,\ldots,x_k^1 \in D}} \left[\prod_{\epsilon_1,\ldots,\epsilon_k \in \{0,1\}} f\left(x_1^{\epsilon_1},\ldots,x_k^{\epsilon_k}\right)^{\frac{\sum_{i\le k}\epsilon_i}{}} \right].$$

We now recall the model of multiparty protocols, and then discuss its relationship to the above norm. 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 in k blocks, and the i-th party knows all the input bits except those corresponding to the i-th 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, and at the end output a value in $\{-1, 1\}$. The measure of interest is the number of bits exchanged by the parties (for background, see the book by Kushilevitz and Nisan [23]).

A *c-bit k-party protocol* is a protocol between k parties that exchanges at most c bits. For a domain D, we denote by $\Pi_{k,c}$ the class of all such protocols $\pi : D^k \to \{-1, 1\}$. The correlation $\mathrm{Cor}(f, \pi)$ between a complex-valued function and a protocol π is again defined as the (complex) norm of the complex value $E_{x\in D^k}[f(x) \cdot \pi(x)]$.

The next Lemma 3.2 shows that $R_k(f)$ upper bounds the correlation of f with low-communication k-party protocols. The same lemma (for real-valued functions) appears in [7, 32] and is proved using the notion of *discrepancy* of a function. Building on [32], we give a direct proof of the lemma which avoids discrepancy, and also generalizes to complex-valued functions.

Lemma 3.2. *For every function $f : D^k \to \mathbb{C}$, $\mathrm{Cor}(f, \Pi_{k,c}) \le 2^c \cdot R_k(f)^{1/2^k}$.*

To prove Lemma 3.2 we need the following lemma, which is a generalization of Lemma 4.1 in [32] to complex-valued functions.

Lemma 3.1. *For any function $f : D^k \to \mathbb{C}$, $|E_{x\in D^k}[f(x)]| \le R(f)^{1/2^k}$.*

The proof of the above lemma amounts to writing $R_k(f)$ in an appropriate form, then using the inequality $E\left[|Z|^2\right] \ge |E[Z]|^2$, and repeating this argument k times. Due to space restrictions we omit the standard derivation of this inequality which can be found in the full version of this paper.

Proof of Lemma 3.2 assuming Lemma 3.1. Let π be a c-bit protocol. It is well-known that π partitions D^k in at most 2^c monochromatic cylinder intersections $CI_i = C_{i,1} \cap \ldots \cap C_{i,k}$ where $C_{i,j} \subseteq D^k$ is a *cylinder* in the j-th coordinate, that is, the membership of $(x_1, \ldots, x_k)$ in $C_{i,j}$ does not depend on the j-th coordinate x_j (see, e.g., [23, Lemma 6.10]).

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 a protocol is naturally written in terms of $0/1$ functions, but our norms require $-1/1$ functions. For any i, j, consider the random function $g_{i,j} : D^k \to \{-1, 1\}$ defined as $g_{i,j}(x) := 1$ with probability 1 if $x \in C_{i,j}$, and $g_{i,j}(x) := 1$ with probability $1/2$ if $x \notin C_{i,j}$ (and consequently $g_{i,j}(x) := -1$ also with probability $1/2$ if $x \notin C_{i,j}$). Now observe that for every $i \le 2^c$ and every $x \in (\{0,1\}^n)^k$, the expectation $E_{g_{i,1},\ldots,g_{i,k}}[g_{i,1}(x) \cdot g_{i,2}(x) \cdots g_{i,k}(x)] = \prod_{j\le k} E_{g_{i,j}}[g_{i,j}(x)]$ equals 1 if $x \in CI_i = C_{i,1} \cap \ldots \cap C_{i,k}$, and 0 otherwise. Therefore, denoting by $v(i)$ the value of the protocol π on inputs in the cylinder intersection CI_i, the protocol can be written as

$$\pi(x) = \sum_{i\le 2^c} v(i) \cdot E_{g_{i,1},\ldots,g_{i,k}} \left[\prod_{j\le k} g_{i,j}(x) \right].$$

We now have, by linearity of expectation,

$$E_x[f(x)\pi(x)] = E_{g's} \left[\sum_{i\le 2^c} E_x \left[f(x)v(i) \prod_{j\le k} g_{i,j}(x) \right] \right].$$

[8] These works give *uniform* circuits, while for the point made here non-uniform circuits would suffice. However, we do not know how to substantially simplify the construction in [16, 19] if one allows for non-uniformity.

By fixing the randomness for the g's so as to maximize the outermost expectation, we have

$$E_x\left[f(x)\pi(x)\right] \le 2^c \max_i E_x\left[f(x)v(i)\prod_{j\le k} g_{i,j}(x)\right]$$
$$\le 2^c \max_i R_k\left(f\cdot v(i)\prod_{j\le k} g_{i,j}\right)^{1/2^k} = 2^c R_k\left(f\right)^{1/2^k},$$

where the second inequality follows from Lemma 3.1. The last equality (for real-valued functions) is Claim 5.2 in [32]. This equality intuitively holds because each $g_{i,j}$ is cylindrical, i.e. its value does not depend on the j-th coordinate of its input. More specifically, it is not hard to see that for every function h and every $-1/1$ function g that does not depend on some j-th coordinate, $R_k\left(h\cdot g\right) = R_k\left(h\right)$. To see this, let $j = k$ without loss of generality, and note that for every fixed $x_1^0, x_2^0, \ldots, x_k^0, x_1^1, x_2^1, \ldots, x_k^1$, we have

$$\prod_{\epsilon_1,\ldots,\epsilon_k\in\{0,1\}} (h\cdot g)\left(x_1^{\epsilon_1}, x_2^{\epsilon_2}, \ldots, x_k^{\epsilon_k}\right)^{\sum_{i\le k}\epsilon_i}$$
$$= \prod_{\epsilon_1,\ldots,\epsilon_{k-1}} \left(g\left(x_1^{\epsilon_1}, x_2^{\epsilon_2}, \ldots, x_{k-1}^{\epsilon_{k-1}}, 0\right)^2 \cdot \prod_{\epsilon_k} h\left(x_1^{\epsilon_1}, x_2^{\epsilon_2}, \ldots, x_k^{\epsilon_k}\right)^{\sum_{i\le k}\epsilon_i}\right)$$
$$= \prod_{\epsilon_1,\ldots,\epsilon_k} h\left(x_1^{\epsilon_1}, x_2^{\epsilon_2}, \ldots, x_k^{\epsilon_k}\right)^{\sum_{i\le k}\epsilon_i},$$

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

We now state and prove a new lemma that shows that the k-party norm lower bounds the correlation.

Lemma 3.3. *For every function $f : D^k \to \{-1,1\}$, $\mathrm{Cor}(f, \Pi_{k,k}) \ge R_k\left(f\right)$.*

Proof. For $x_1^1, x_2^1, \ldots, x_k^1 \in D$, consider the function $g_{x_1^1,\ldots,x_k^1} : D^k \to \{-1,1\}$, where $g_{x_1^1,\ldots,x_k^1}(x_1^0, \ldots, x_k^0)$ is defined as

$$\prod_{\epsilon_1,\ldots,\epsilon_k\in\{0,1\}:\text{not all zero}} f\left(x_1^{\epsilon_1}, \ldots, x_k^{\epsilon_k}\right).$$

Now observe that

$$E_{x_1^1,\ldots,x_k^1}\Big[E_{x_1^0,\ldots,x_k^0}\Big[f(x_1^0,\ldots,x_k^0) \cdot g_{x_1^1,\ldots,x_k^1}(x_1^0,\ldots,x_k^0)\Big]\Big] = R_k\left(f\right).$$

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

To conclude the proof, note that $g(x_1^0, \ldots, x_k^0)$ is computable by a k-bit k-party protocol because it is the product of terms none of which depends on all x^0's. More specifically, we can partition the $2^k - 1$ terms appearing in the definition of g in k sets in such a way that party i can evaluate all the terms in set i; now a protocol for g can be obtained by having each party i communicate the product of the terms in set i, which takes 1 bit. □

Again, a key property for us of the norm is the fact, formally stated next, that the norm of the product of two functions defined on disjoint input bits is the product of the norms of the two functions (cf., [7]). The proof of this fact is immediate from the definition.

Fact 3.4. *For functions $f : D^k \to \mathbb{C}$ and $f' : \left(D'\right)^k \to \mathbb{C}$, define the function $\left(f\cdot f'\right) : \left(D\times D'\right)^k \to \mathbb{C}$ by $\left(f\cdot f'\right)\left((x_1,x_1'),\ldots,(x_k,x_k')\right) := f(x_1,\ldots,x_k)\cdot f'(x_1',\ldots,x_k')$. Then $R_k\left(f\cdot f'\right) = R_k\left(f\right)\cdot R_k\left(f'\right)$.*

3.2. XOR and direct product lemmas for multiparty protocols

In this section we show how to obtain the following XOR lemma for multiparty protocols.

Theorem 1.2 (XOR lemma for multiparty protocols)**.** *Let $f : D^k \to \{-1,1\}$ be a function such that $\mathrm{Cor}\left(f, \Pi_{k,k}\right) \le \epsilon$. Then $\mathrm{Cor}\left(f^{\oplus m}, \Pi_{k,c}\right) \le 2^c\cdot\epsilon^{m/2^k}$.*

Proof. We have

$$E_x\left[f^{\oplus m}(x)\cdot\pi(x)\right] \le 2^c\cdot R_k\left(f^{\oplus m}\right)^{1/2^k} = 2^c\cdot R_k\left(f\right)^{m/2^k} \le 2^c\cdot\epsilon^{m/2^k},$$

where the first inequality holds by Lemma 3.2, the next equality by Fact 3.4, and the next inequality by Lemma 3.3. □

Combining the above XOR lemma with Proposition 1.1 we immediately obtain a direct product lemma for k-party protocols.

Corollary 1.2 (Direct product lemma for multiparty protocols)**.** *Let $f : D \to \{-1,1\}$ be a function such that $\mathrm{Cor}\left(f, \Pi_{k,k}\right) \le \epsilon \le 2^{-(c+1)\cdot 2^k}$. Then $\mathrm{Suc}\left(f^{(m)}, \Pi_{k,c}\right) \le 2^{-\Omega(m)}$.*

Proof. Proposition 1.1 implies that $\mathrm{Suc}\left(f^{(m)}, \Pi_{k,c}\right)$ can be upper bounded by $\mathrm{Cor}\left(f^{\oplus m'}, C'\right) + 2^{-\Omega(m)}$, 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 $\mathrm{Suc}\left(f^{(m)}, \Pi_{k,c}\right) \le \mathrm{Cor}\left(f^{\oplus m'}, \Pi_{k,m'\cdot c}\right) + 2^{-\Omega(m)}$.

By Theorem 1.2, we have that $\mathrm{Cor}\left(f^{\oplus m'}, \Pi_{k,m'\cdot c}\right) \leq 2^{m'\cdot c}\cdot \epsilon^{m'/2^k} \leq 2^{-m'}$, which gives the result. □

Remark 3.5 ($R_k(f)$ captures correlation with $SimXOR_k$)**.** *The results in Section 3.1 show that the norm $R_k(f)$ captures the correlation of f with c-bit k-party protocols, with some loss in the parameters, in particular a factor 2^c. We remark that this norm more accurately captures the correlation of f with another computational model, which is more restricted than k-party protocols but more general than $GF(2)$-polynomials. This model, which we denote $SimXOR_k$, is a k-party simultaneous communication model, where all k players communicate exactly 1 bit with no communication, and the output is the XOR of these k-bits (i.e., the referee computes XOR). The observation of Håstad and Goldmann [18, Proof of Lemma 4] shows that $SimXOR_k$ is more powerful than degree-$(k-1)$ $GF(2)$ polynomials, while obviously $SimXOR_k$ is a special case of k-bit k-party protocols. We note that the proofs of Lemmas 3.2 and 3.3 establish that $R_k(f) \leq \mathrm{Cor}(f, SimXOR_k) \leq R_k(f)^{1/2^k}$. Using this and Fact 3.4 one can prove an XOR lemma for $SimXOR_k$ which is analogous to Theorem 1.2 with $c = 0$.*

3.2.1. The case of two parties

In this section we further discuss XOR lemmas for the special interesting case of $k = 2$ parties. We start by comparing our results with an XOR lemma by Shaltiel [36], and then we present a counterexample to the "ideal" setting of parameteres of the XOR lemma, i.e. going from correlation ϵ to correlation ϵ^m.

Remark 3.6 (Comparison with the XOR lemma by Shaltiel [36])**.** *For $k = 2$ parties, Shaltiel proves an XOR lemma which (up to different constants) has the same conclusion as ours (Theorem 1.2) but starts from the assumption that the original function f has bounded* discrepancy *(as opposed to bounded correlation with 2-bit protocols in our result), where the discrepancy of a function $f : D \times D \to \{-1,1\}$ is defined as the maximum, over all rectangles R, of $|E_{x,y}[f(x,y)|(x,y) \in R]| \cdot \Pr[(x,y) \in R]$ (cf. [23]). 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, this is perhaps misleading, as the discrepancy of f in fact equals the maximum correlation of f with 2-bit protocols (up to constant factors). This fact is very similar to what is shown in Lemmas 3.3 and 3.2, and we now discuss it in the language of discrepancy. 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 fact, i.e. that the discrepancy upper bounds the correlation with 2-bit protocols, is standard and can be found, e.g., in [23]. Thus, for $k = 2$, our XOR lemma (Theorem 1.2) 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 1.2) 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 3.7. *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. *$Corr(f, \Pi_{2,2}) \leq 5/9$.*
2. *$Corr(f^{\oplus 2}, \Pi_{2,2}) \geq 33/81 > (5/9)^2$.*

Remark 3.8 (Comparison with the counterexample by Shaltiel [36])**.** *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 3.7) 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 3.7. Although the proof involves a certain amount of calculation, it is perhaps instructive to observe how a 2-bit protocol can correlate with $f^{\oplus 2}$ 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, $\Pr[\mathcal{E} | x \neq x' \wedge y \neq y'] = (1/3)(1/2) + (2/3)(3/4) = 1/6 + 1/2 = 2/3$.

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

By symmetry, $\Pr[\mathcal{E} | x \neq x' \wedge y = y'] = 2/3$ as well.

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

We now briefly explain how to extend the counterexample in Claim 3.7 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 3.7 for the function $f : D^2 \to \{-1, 1\}$ defined as $f(x,y) := 1$ if and only if $(x \mod 3) = (y \mod 3)$. 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. For a game G we define the game $G^{\oplus m}$ in the obvious way: The verifier asks m independent questions and expects m answers. Then it checks how many of the m games are accepting, and accepts according to the parity of this number.

Consider the following game G between a verifier and *one* prover A. (The example to follow does not exploit the fact that there is only one prover and in fact can be trivially transformed into an example for any number of provers.) The verifier sends two bits (p,t) to A. A then sends one bit a 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 on 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 a game with $p = 0$ in $G^{\oplus m}$, A can force the XOR of the games to be accepted.

Specifically, we claim that any strategy A achieves correlation at most $1/2$ with G. This is because when $p = 1$, no matter what A says, the verifier accepts with probability $1/2$ according to t, and therefore we have correlation 0.

Now consider the game $G^{\oplus m}$ and the following strategy 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 verifier accepts $G^{\oplus m}$ whenever there is an i such that $p_i = 0$, which happens with probability $1 - 2^{-m}$, and therefore this strategy has correlation $1 - 2^{-m}$ with $G^{\oplus m}$.

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. This simple proof already appears in [7], but we now present it using a different language.[9] In what follows we denote by $\wedge_k : \{0,1\}^k \to \{0,1\}$ 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 $\left((-1)^{\wedge_k}\right)^{\oplus n}$, i.e. $GIP(x_1, \ldots, x_k) := \prod_{i \leq n} (-1)^{\wedge_{j \leq k}(x_j)_i}$.

Theorem 3.9 ([3, 7]). $\mathrm{Cor}(GIP, \Pi_{k,c}) \leq 2^{c - \Omega\left(n/4^k\right)}$.

Proof.

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

where the first inequality is Lemma 3.2, the next inequality is Fact 3.4, and $R_k\left((-1)^{\wedge_k}\right) = 1 - 2^{-k+1}$ is a straightforward calculation. □

[9] While in [7, Theorem 5] they claim a $\exp\left(-\Omega\left(n/2^k\right)\right)$ lower bound, their proof only gives $\exp\left(-\Omega\left(n/4^k\right)\right)$, which we also obtain here. To the best of our knowledge it is not known whether a $\exp\left(-\Omega\left(n/2^k\right)\right)$ lower bound holds.

Using our extension of the k-party norm to complex-valued functions (as opposed to real-valued), 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 [15] obtained the corresponding lower bound using the techniques in [3]. In the following theorem we again use the definition of correlation in Section 2.3. Let $GIP_m : (\{0,1\}^n)^k \to \{-1,1\}$ be the function that equals 1 iff $\sum_{i \le n} \wedge_{j \le k} (x_j)_i$ is divisible by m.

Theorem 3.10. $\mathrm{Cor}(GIP_m, \Pi_{k,c}) \le 2^{c - \alpha \cdot n/4^k}$, *where* $\alpha > 0$ *depends on* m *only.*

Proof. Following the proof of Theorem 2.8, we consider the function $f : (\{0,1\}^n)^k \to \mathbb{C}$ defined as $f(x) := e_m\left(\sum_{i \le n} \wedge_{j \le k} (x_j)_i\right)$, where $e_m(y) := e^{2\pi \cdot i \cdot y/m}$. As in the proof of Theorem 2.8, to obtain the claimed bound on the correlation it is enough to bound from above the maximum over $a \in \{1, \ldots, m-1\}$ of $\left| E_{x \in (\{0,1\}^n)^k} \left[f(x)^a \cdot \pi(x) \right] \right|$, where $\pi \in \Pi_{k,c}$. The proof of this latter claim is a relatively standard algebraic manipulation. For the case $k = 1$ of the Mod_m function, this manipulation appears for example in [4, Equation (4)], while for general k one can follow the same approach in [4, Equation (4)], this time making use of the fact that $|\Pr_x[GIP_m(x) = 1] - 1/m| \le \exp(-\alpha \cdot n/2^k)$ for a constant $\alpha > 0$ that depends on m only, which in turn is not too hard to verify. To bound the above quantity, we use Lemma 3.2 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 3.4). Thus we obtain $\left| E_{x \in (\{0,1\}^n)^k} \left[f(x)^a \cdot \pi(x) \right] \right| \le R_k\left(e_m\left(a \cdot \wedge_k\right)\right)^{n/2^k}$ and we are left with the task of bounding

$$R_k\left(e_m\left(a \cdot \wedge_k\right)\right) = E_{\substack{x_1^0, \ldots, x_k^0 \in \{0,1\} \\ x_1^1, \ldots, x_k^1 \in \{0,1\}}} \left[e_m\left(a \cdot \sum_{\epsilon_1, \ldots, \epsilon_k \in \{0,1\}} (-1)^{\sum_i \epsilon_i} \wedge_k \left(x_1^{\epsilon_1}, \ldots, x_k^{\epsilon_k}\right) \right) \right].$$

Consider now the event $V :=$ "$x_i^0 \neq x_i^1$ for every i." When V happens, there is exactly one choice for the ϵ_i's that gives $\wedge_k(x_1^{\epsilon_1}, x_2^{\epsilon_2}, \ldots, x_k^{\epsilon_k}) = 1$, and that choice is $\epsilon_i := x_i^1$ (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

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

where $\Re(\cdot)$ denotes the real part. Above, the first equality uses the fact that $\sum_i x_i^1$ 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 CCC referees for helpful comments. The first author would like to thank Salil Vadhan for his helpful reading of a preliminary version of this work [40].

References

[1] N. Alon, O. Goldreich, J. Håstad, and R. Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures & Algorithms*, 3(3):289–304, 1992.

[2] N. Alon, T. Kaufman, M. Krivelevich, S. Litsyn, and D. Ron. Testing low-degree polynomials over $\mathrm{GF}(2)$. In *Approximation, randomization, and combinatorial optimization*, volume 2764 of *Lecture Notes in Comput. Sci.*, pages 188–199. Springer, Berlin, 2003.

[3] L. Babai, N. Nisan, and M. Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. Comput. System Sci.*, 45(2):204–232, 1992. Twenty-first Symposium on the Theory of Computing (Seattle, WA, 1989).

[4] J. Bourgain. Estimation of certain exponential sums arising in complexity theory. *C. R. Math. Acad. Sci. Paris*, 340(9):627–631, 2005.

[5] A. K. Chandra, M. L. Furst, and R. J. Lipton. Multi-party protocols. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*, Boston, Massachusetts, 25–27 Apr. 1983.

[6] A. Chattopadhyay. An improved bound on correlation between polynomials over z_m and mod_q. *Electronic Colloquium on Computational Complexity*, Technical Report TR06-107, 2006. `http://www.eccc.uni-trier.de/eccc`.

[7] F. R. K. Chung and P. Tetali. Communication complexity and quasi randomness. *SIAM J. Discrete Math.*, 6(1):110–123, 1993.

[8] U. Feige. Error reduction by parallel repetition-the state of the art. Technical report, Jerusalem, Israel, Israel, 1995.

[9] O. Goldreich and L. A. Levin. A hard-core predicate for all one-way functions. In *Proceedings of the Twenty First Annual ACM Symposium on Theory of Computing*, pages 25–32, Seattle, Washington, 15–17 May 1989.

[10] O. Goldreich, N. Nisan, and A. Wigderson. On Yao's XOR lemma. Technical Report TR95–050, Electronic Colloquium on Computational Complexity, March 1995. `http://www.eccc.uni-trier.de/eccc`.

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

[12] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.

[13] B. Green and T. Tao. An inverse theorem for the gowers u^3 norm, 2005. arXiv.org:math/0503014.

[14] F. Green, A. Roy, and H. Straubing. Bounds on an exponential sum arising in Boolean circuit complexity. *C. R. Math. Acad. Sci. Paris*, 341(5):279–282, 2005.

[15] V. Grolmusz. Separating the communication complexities of $\mathrm{MOD}\, m$ and $\mathrm{MOD}\, p$ circuits. *J. Comput. System Sci.*, 51(2):307–313, 1995.

[16] D. Gutfreund and E. Viola. Fooling parity tests with parity gates. In *Proceedings of the Eight International Workshop on Randomization and Computation (RANDOM)*, Lecture Notes in Computer Science, Volume 3122, pages 381–392. Springer-Verlag, 2004.

[17] A. Hajnal, W. Maass, P. Pudlák, M. Szegedy, and G. Turán. Threshold circuits of bounded depth. *J. Comput. System Sci.*, 46(2):129–154, 1993.

[18] J. Håstad and M. Goldmann. On the power of small-depth threshold circuits. *Comput. Complexity*, 1(2):113–129, 1991.

[19] A. Healy. Randomness-efficient sampling within nc^1. In *Proceedings of the 10th International Workshop on Randomization and Computation (RANDOM)*, Lecture Notes in Computer Science., 2006.

[20] R. Impagliazzo. Hard-core distributions for somewhat hard problems. In *36th Annual Symposium on Foundations of Computer Science*, pages 538–545, Milwaukee, Wisconsin, 23–25 Oct. 1995. IEEE.

[21] R. Impagliazzo and A. Wigderson. $P = BPP$ if E requires exponential circuits: Derandomizing the XOR lemma. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 220–229, El Paso, Texas, 4–6 May 1997.

[22] C. S. Jutla, A. C. Patthak, A. Rudra, and D. Zuckerman. Testing low-degree polynomials over prime fields. *focs*, 00:423–432, 2004.

[23] E. Kushilevitz and N. Nisan. *Communication complexity*. Cambridge University Press, Cambridge, 1997.

[24] L. A. Levin. One way functions and pseudorandom generators. *Combinatorica*, 7(4):357–363, 1987.

[25] M. Luby, B. Velickovic, and A. Wigderson. Deterministic approximate counting of depth-2 circuits. In *Proceedings of the 2nd Israeli Symposium on Theoretical Computer Science (ISTCS)*, pages 18–24, 1993.

[26] J. Naor and M. Naor. Small-bias probability spaces: efficient constructions and applications. In *Proceedings of the 22nd Annual ACM Symposium on the Theory of Computing*, pages 213–223, 1990.

[27] N. Nisan. Pseudorandom bits for constant depth circuits. *Combinatorica*, 11(1):63–70, 1991.

[28] N. Nisan, S. Rudich, and M. Saks. Products and help bits in decision trees. *SIAM J. Comput.*, 28(3):1035–1050, 1999.

[29] N. Nisan and A. Wigderson. Hardness vs randomness. *J. Computer & Systems Sciences*, 49(2):149–167, Oct. 1994.

[30] I. Parnafes, R. Raz, and A. Wigderson. Direct product results and the GCD problem, in old and new communication models. In *STOC '97 (El Paso, TX)*, pages 363–372 (electronic). ACM, New York, 1999.

[31] R. Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803 (electronic), 1998.

[32] R. Raz. The BNS-Chung criterion for multi-party communication complexity. *Comput. Complexity*, 9(2):113–122, 2000.

[33] A. 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.

[34] A. Samorodnitsky. Low-degree tests at large distances. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing 2007, San Diego, CA USA, June 2007*, 2007.

[35] A. Samorodnitsky and L. Trevisan. Gowers uniformity, influence of variables, and PCPs. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing 2006, Seattle, WA, USA, May 21–23, 2006*, pages 11–20, 2006.

[36] R. Shaltiel. Towards proving strong direct product theorems. *Comput. Complexity*, 12(1-2):1–22, 2003.

[37] R. Shaltiel and C. Umans. Simple extractors for all min-entropies and a new pseudorandom generator. *J. ACM*, 52(2):172–216, 2005.

[38] R. Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, pages 77–82, New York City, 25–27 May 1987.

[39] E. Viola. *The Complexity of Hardness Amplification and Derandomization*. PhD thesis, Harvard University, 2006. `http://www.eccc.uni-trier.de/eccc`.

[40] E. Viola. New correlation bounds for gf(2) polynomials using gowers uniformity. *Electronic Colloquium on Computational Complexity*, Technical Report TR06-097, 2006. `http://www.eccc.uni-trier.de/eccc`.

[41] E. Viola. Pseudorandom bits for constant-depth circuits with few arbitrary symmetric gates. *SIAM Journal on Computing*, 36(5):1387–1403, 2007.

[42] A. C. Yao. Theory and applications of trapdoor functions (extended abstract). In *23rd Annual Symposium on Foundations of Computer Science*, pages 80–91, Chicago, Illinois, 3–5 Nov. 1982. IEEE.

[43] A. C.-C. Yao. Some complexity questions related to distributive computing. In *STOC '79: Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 209–213, New York, NY, USA, 1979. ACM Press.