Correlation bounds against polynomials[*]

Emanuele Viola[†]

October 14, 2022

This survey is about one of the most basic computational models: low-degree polynomials over the field $\{0,1\} = \mathrm{GF}(2)$. For example, the following is a polynomial of degree 2 in 3 variables

$$p(x_1, x_2, x_3) := x_1 \cdot x_2 + x_2 + x_3 + 1,$$

given by the sum of the 4 monomials x_1x_2, x_2, x_3, and 1, of degree $2, 1, 1$, and 0, respectively. This polynomial computes a function from $\{0,1\}^3$ to $\{0,1\}$, which we also denote p, by performing the arithmetic over $\{0,1\}$. Thus the sum "+" is modulo 2 and is the same as "xor," while the product "$\cdot$" is the same as "and." For instance, $p(1,1,0) = 1$. Being complexity theorists rather than algebraists, we are only interested in the function computed by a polynomial, not in the polynomial itself; therefore we need not bother with variables raised to powers bigger than 1, since for $x \in \{0,1\}$ one has $x = x^2 = x^3$ and so on. In general, a polynomial p of degree d in n Boolean variables $x_1, \ldots, x_n \in \{0,1\}$ is a sum of monomials of degree at most d:

$$p(x_1, \ldots, x_n) = \sum_{M \subseteq \{1,\ldots,n\}, |M| \le d} c_M \prod_{i \in M} x_i,$$

where $c_M \in \{0,1\}$ and we let $\prod_{i \in \emptyset} x_i := 1$; such a polynomial p computes a function $p : \{0,1\}^n \to \{0,1\}$, interpreting again the sum modulo 2. We naturally measure the complexity of a polynomial by its degree d: the maximum number of variables appearing in any monomial. Since every function $f : \{0,1\}^n \to \{0,1\}$ can be computed by a polynomial of degree n, specifically $f(x_1, \ldots, x_n) = \sum_{a_1,\ldots,a_2} f(a_1, \ldots, a_n) \prod_{1 \le i \le n} (1 + a_i + x_i)$, we are interested in polynomials of low degree $d \ll n$.

Low-degree polynomials constitute a fundamental model of computation that arises in a variety of contexts, ranging from error-correcting codes to circuit lower bounds. As for any computational model, a first natural challenge is to exhibit explicit functions that cannot be computed in the model. This challenge is easily won: the monomial $\prod_{i=1}^{d} x_i$ requires degree d. A second, natural challenge has baffled researchers, and is our focus. One now asks for functions that not only cannot be computed by low-degree polynomials, but do not even *correlate* with them.

[*] This survey is a revision of Chapter 1 in [Vio09b] (cf. [Vio09a]), but does not completely subsume the older text.

[†] Supported by NSF grant CCF-2114116. Email: `viola@ccs.neu.edu`

1 Correlation bounds

We start by defining the correlation between a function f and polynomials of degree d. This quantity captures how well we can approximate f by polynomials of degree d, and is also known as the average-case hardness of f against polynomials of degree d.

Definition 1 (Correlation)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a function, n an integer, and D a distribution on $\{0,1\}^n$. We define the* correlation between f and a polynomial $p : \{0,1\}^n \to \{0,1\}$ with respect to D *as*

$$\mathrm{Cor}_D(f,p) := \left|\Pr_{x\sim D}[f(x) = p(x)] - \Pr_{x\sim D}[f(x) \neq p(x)]\right| = 2\left|1/2 - \Pr_{x\sim D}[f(x) \neq p(x)]\right| \in [0,1].$$

We define the correlation between f and polynomials of degree d with respect to D *as*

$$\mathrm{Cor}_D(f,d) := \max_p \mathrm{Cor}_D(f,p) \in [0,1],$$

where the maximum is over all polynomials $p : \{0,1\}^n \to \{0,1\}$ of degree d.

We also define the correlation between f and polynomials of degree d as the minimum of the above over all distributions:

$$\mathrm{Cor}(f,d) := \min_D \mathrm{Cor}_D(f,d).$$

By Yao's duality principle [Yao77] (a.k.a. min/max, linear programming duality, Hahn–Banach, etc.), $\mathrm{Cor}(f,d) \geq \epsilon$ if and only if there is a distribution P on degree-d polynomials such that for every input x we have $\Pr[P(x) = f(x)] \geq 1/2 + \epsilon/2$. If $\mathrm{Cor}(f,d) \geq \epsilon$ we also say that f has $(1/2 - \epsilon/2)$*-error probabilistic degree d,* or that f has a degree-d *probabilistic polynomial* with error $(1/2 - \epsilon/2)$.

Since the 80's, researchers have sought to exhibit explicit functions that have small correlation with high-degree polynomials. We refer to this enterprise as obtaining, or proving, "correlation bounds." A dream setting of parameters would be to exhibit a function $f \in \mathrm{P}$ such that for every n, and for D the uniform distribution over $\{0,1\}^n$, $\mathrm{Cor}_D(f, \epsilon \cdot n) \leq \exp(-\epsilon \cdot n)$, where $\epsilon > 0$ is an absolute constant, and $\exp(x) := 2^x$. For context, we mention that a random function satisfies such strong correlation bounds, with high probability.

The original motivation for seeking correlation bounds comes from circuit complexity, because functions with small correlation with polynomials require large constant-depth circuits of certain types, see e.g. [Raz87, Smo87, HMP$^+$93, Bei93]. An additional motivation comes from pseudorandomness: as we will see, sufficiently strong correlation bounds can be used to construct pseudorandom generators [Nis91, NW94], which in turn have myriad applications. Moreover, as we shall shortly see, progress on correlation bounds is *necessary* for progress on many other open problems in complexity theory. But as this survey also aims to put forth, today the challenge of proving correlation bounds is interesting *per se*, and its status is a fundamental benchmark for our understanding of complexity theory: it is not known how to achieve the dream setting of parameters mentioned above, and in fact nobody can even achieve the following seemingly much weaker parameters.

Open question 1. *Is there a function $f \in \mathrm{NP}$ such that for arbitrarily large n there is a distribution D on $\{0,1\}^n$ with respect to which $\mathrm{Cor}_D(f, \log_2 n) \leq 1/\sqrt{n}$?*

Before discussing known results in the next sections, we add to the above concise motivation for tackling correlation bounds the following discussion of their relationship with other open problems.

Correlation bounds are necessary for progress on many central problems. We begin by pointing out that a negative answer to Question 1 implies that NP has circuits of quasipolynomial size $s = n^{O(\log n)}$. This relatively standard fact can be proved via boosting [Fre95, Section 2.2] or min-max/linear-programming duality [GHR92, Section 5]. Thus, an affirmative answer to Question 1 is necessary to prove that NP does not have circuits of quasipolynomial size, a leading goal of theoretical computer science. Of course, this connection can be strengthened in various ways, for example noting that the circuits for NP given by a negative answer to Question 1 can be written on inputs of length n as a majority of $n^{O(1)}$ polynomials of degree $\log_2 n$; thus, an affirmative answer to Question 1 is necessary even to prove that NP does not have circuits of the latter type. On the other hand, Question 1 cannot easily be related to polynomial-size lower bounds such as NP $\not\subseteq$ P/poly, because a polynomial of degree $\log n$ may have a quasipolynomial number of monomials.

Progress on Question 1 is also necessary for progress on *number-on-forehead communication complexity lower bounds* [CFL83]. Specifically, a long-standing open question in communication complexity is to exhibit an explicit function $f : (\{0,1\}^n)^k \to \{0,1\}$ that cannot be computed by number-on-forehead k-party protocols exchanging $O(k)$ bits, for some $k \geq \log_2 n$ [KN97, Problem 6.21]. As pointed out in [Vio17, Vio], a $\omega(\log^3 n)$ communication lower bound for $k = O(\log^2 n)$ would also answer Question 1; but the converse is not known, so correlation bounds, if true, might be easier to prove.

Moreover, Progress on Question 1 is also necessary for progress on Valiant's *matrix rigidity* problem [Val77]. We refer the reader to [Vio] for a proof of this connection, cf. .[SV12]. For more on matrix rigidity see the surveys [Lok09, Ram20].

Finally, the recent "polarizing random walks" paradigm [CHHL18, CHLT19, CGL$^+$20] constructs new pseudorandom generators against classes of functions with "bounded Fourier tails." Pseudorandom generators are discussed later in Section 2. In an effort to use this framework to improve the state of pseudorandom generators against low-degree polynomials, several conjectures have been put forth about the Fourier spectrum of polynomials. We refer the reader to [Vio21] for discussion, but, to give a quick example, one of the conjectures was that the sum of the size-2 Fourier coefficients of a degree-d polynomial is at most $O(d^2)$. The conjectures have been verified for special classes of polynomials [BIJ$^+$21]. One interesting feature of the polarizing random walks approach is that, unlike the influential approach by Nisan [Nis91], it is not directly based on correlation bounds. In particular, it was conceivable that one could prove the conjectures and obtain better pseudorandom generators without proving new correlation bounds. However, it was shown in [Vio21] that in fact correlation bounds are also *necessary:* Any of the proposed conjectures, or even weaker ones, if true would imply an answer to Question 1.

The above connections arguably set apart the goal of proving correlation bounds from other long-standing goals in computational complexity: We are not aware of another problem which is unrelated to correlation bounds and has a comparable number of connections with other long-standing problems.

Barriers. Regarding the limitation of current techniques, we point out that oracle results such as [BGS75, AW08] are not relevant to non-uniform models such as the one considered here. It is also not clear if natural proofs [RR97] present an obstacle, since we do not have candidate pseudorandom functions that correlate with low-degree polynomials. Bhowmick and Lovett [BL15] present a "barrier" result that is specific to low-degree polynomials. They show that certain long-standing correlation bounds, such as strong bounds for the mod_3 function which is defined in the next subsection, are false for a generalization of polynomials known as non-classical polynomials. This means that, should those strong bounds be true, the proof cannot apply to non-classical polynomials. They also discuss which of the available proof techniques apply to non-classical polynomials, and we shall discuss another approach which circumvents this obstacle later in §1.4.

After this discussion, we now move to presenting the known correlation bounds. It is a remarkable state of affairs that, while we are currently unable to make the correlation small and the degree large *simultaneously*, as asked in Question 1, we can make the correlation small and the degree large *separately*. And in fact we can even achieve this for the same explicit function $f = \text{mod}_3$. We examine these two types of results in turn.

1.1 Large degree $d \gg \log n$ but noticeable correlation $\epsilon \gg 1/n$

Razborov [Raz87] (also in [CK02, Section 2.7.1]) proves the existence of a symmetric function $f : \{0,1\}^n \to \{0,1\}$ that has correlation at most $1 - 1/n^{O(1)}$ with polynomials of degree $\Omega(\sqrt{n})$ (a function is symmetric when its value only depends on the Hamming weight of the input).

In [Smo93] Smolensky proves the following sharper result. For a more recent exposition see e.g. [Kop11]. We denote by U the uniform distribution over $\{0,1\}^n$.

Theorem 1. *[Smo93]* $\text{Cor}_U(Majority, d) \le O(d/\sqrt{n})$.

This was shown to be tight in [Vio21]. Perhaps surprisingly, tight bounds on $\text{Cor}(Majority, d)$ are not known.

Theorem 2. *[Vio21] We have:*
(1) $\text{Cor}_U(Majority, d) \ge O(d/\sqrt{n})$,
(2) $\text{Cor}(Majority, d) \ge O(d^2/n)$, *and*
(3) $\text{Cor}(Majority, 1) \le O(1/n)$.

Item (1) in Theorem 2 matches the bound in Theorem 1. It is conjectured in [Vio21] that (2) is tight. Note that (3) shows that the bound in Theorem 1 is not tight for non-uniform distributions.

We now move to functions which, unlike majority, are candidate for having very small correlation. We consider the function $\text{mod}_3 : \{0,1\}^n \to \{0,1\}$ which evaluates to 1 if and only if the number of input bits that are 1 is of the form $3k + 1$ for some integer k, i.e., it is congruent to 1 modulo 3:

$$\text{mod}_3(x_1, \ldots, x_n) = 1 \Leftrightarrow \sum_i x_i = 1(\text{mod } 3).$$

For example, $\text{mod}_3(1,0,0) = \text{mod}_3(0,1,0) = 1 \neq \text{mod}_3(1,0,1)$.

Smolensky proved the same correlation bounds for this function as well.

Theorem 2. *[Smo87]* $\text{Cor}(\text{mod}_3, d) \leq O(d/\sqrt{n})$.

While the proof of Smolensky's result has appeared several times, e.g. [Smo87, BS90, Bei93, AB09], we are unaware of a source that directly proves Theorem 2, and thus we include next a proof for completeness. We break up the proof in two parts, also to illustrate a powerful "amplification" methodology. In the first part we show that for any n that, for simplicity, is divisible by 3, and for U the uniform distribution over $\{0,1\}^n$, $\text{Cor}_U(\text{mod}_3, \epsilon\sqrt{n}) \leq 0.9$, where $\epsilon > 0$ is an absolute constant.

Proof. The idea is to consider the set of inputs $X \subseteq \{0,1\}^n$ where the polynomial computes the mod_3 function correctly, and use the polynomial to represent any function defined on X by a polynomial of degree $n/2 + d$. This means that the number of functions defined on X should be smaller than the number of polynomials of degree $n/2 + d$, which leads to the desired tradeoff between $|X|$ and d. To carry through this argument, one works over a field F that extends $\{0,1\}$.

We start by noting that, since n is divisible by 3, one has

$$\sum_i x_i = 2(\text{mod } 3) \Leftrightarrow \sum_i 1 - x_i = 1(\text{mod } 3) \Leftrightarrow \text{mod}_3(1 + x_1, \ldots, 1 + x_n) = 1, \quad (1)$$

where the sums $1 + x_i$ in the input to mod_3 are modulo 2. Let F be the field of size 4 that extends $\{0,1\}$, which we can think of as $F = \{0,1\}[t]/(t^2 + t + 1)$: the set of polynomials over $\{0,1\}$ modulo the irreducible polynomial $t^2 + t + 1$. Note that $t \in F$ has order 3, since $t^2 = t + 1 \neq 1$, while $t^3 = t^2 + t = 1$. Let $h : \{1, t\} \to \{0,1\}$ be the "change of domain" linear map $h(\alpha) := (\alpha + 1)/(t + 1)$; this satisfies $h(1) = 0$ and $h(t) = 1$.

Observe that for every $y \in \{1,t\}^n$ we have, using Equation (1):

$$y_1 \cdots y_n = 1 + (t+1) \cdot \text{mod}_3(h(y_1), \ldots, h(y_n)) + (t^2 + 1) \cdot \text{mod}_3(1 + h(y_1), \ldots, 1 + h(y_n)). \quad (2)$$

Now fix any polynomial $p : \{0,1\}^n \to \{0,1\}$ and let

$$\Pr_{x \in \{0,1\}^n}[p(x) \neq \text{mod}_3(x)] =: \delta,$$

which we aim to bound from below. Let $p' : \{1,t\}^n \to F$ be the polynomial

$$p'(y_1, \ldots, y_n) := 1 + (t+1) \cdot p(h(y_1), \ldots, h(y_n)) + (t^2 + 1) \cdot p(1 + h(y_1), \ldots, 1 + h(y_n));$$

note p' has the same degree d of p. By the definition of p' and δ, a union bound, and Equation (2) we see that

$$\Pr_{y \in \{1,t\}^n}[y_1 \cdots y_n = p'(y_1, \ldots, y_n)] \geq 1 - 2\delta. \quad (3)$$

Now let $S \subseteq \{1,t\}^n$ be the set of $y \in \{1,t\}^n$ such that $y_1 \cdots y_n = p'(y_1, \ldots, y_n)$; we have just shown that $|S| \geq 2^n(1 - 2\delta)$. Any function $f : S \to F$ can be written as a polynomial over F where no variable is raised to powers bigger than 1: $f(y_1, \ldots, y_n) = \sum_{a_1,\ldots,a_n} f(a_1, \ldots, a_n) \prod_{1 \leq i \leq n}(1 + h(y_i) + h(a_i))$. In any such polynomial we can replace any monomial M of degree $|M| > n/2$ by a polynomial of degree at most $n/2 + d$ as follows, without affecting the value on any input $y \in S$:

$$\prod_{i \in M} y_i = y_1 \cdots y_n \prod_{i \notin M} (y_i(t+1) + t) = p'(y_1, \ldots, y_n) \prod_{i \notin M} (y_i(t+1) + t),$$

where the first equality is not hard to verify. Doing this for every monomial we can write $f : S \to F$ as a polynomial over F of degree $\lfloor n/2 + d \rfloor$.

The number of functions from S to F is $|F|^{|S|}$, while the number of polynomials over F of degree $\lfloor n/2 + d \rfloor$ is $|F|^{\sum_{i=0}^{\lfloor n/2+d \rfloor} \binom{n}{i}}$. Thus

$$\log_{|F|} \#\text{functions} = |S| = 2^n(1 - 2\delta) \leq \sum_{i=0}^{\lfloor n/2+d \rfloor} \binom{n}{i} = \log_{|F|} \#\text{polynomials}.$$

Since $d = \epsilon\sqrt{n}$, we have

$$\sum_{i=0}^{\lfloor n/2+d \rfloor} \binom{n}{i} \leq 2^n/2 + d \cdot \binom{n}{\lfloor n/2 \rfloor} \leq 2^n/2 + \epsilon\sqrt{n} \cdot \Theta\left(\frac{2^n}{\sqrt{n}}\right) = (1/2 + \Theta(\epsilon))2^n,$$

where the second inequality follows from standard estimates on binomial coefficients. The standard estimate for even n is for example in [CT06, Lemma 17.5.1]; for odd $n = 2k + 1$ one can first note $\binom{n}{\lfloor n/2 \rfloor} = \binom{2k+1}{k} < \binom{2k+2}{k+1} = \binom{n+1}{(n+1)/2}$ and then again apply [CT06, Lemma 17.5.1]. Therefore $1 - 2\delta \leq 1/2 + \Theta(\epsilon)$ and the theorem is proved. □

In the second part, we would like to build on the first part to prove the theorem. We cannot use the uniform distribution, since the constant 0 polynomial has correlation $\Omega(1)$. The natural "hard" distribution is the one that makes mod_3 balanced, and is used in the next section. Instead, we rely on a general fact pointed out in [Vio] (Lemma 18).

Lemma 3. *[Vio] If* $\text{Cor}(f, d) \geq \epsilon$ *then* $\text{Cor}(f, O(d/\epsilon)) \geq 0.99$.

This is proved by taking the majority of $t := O(1/\epsilon)^2$ independent copies of a probabilistic polynomial for f, and then using the probabilistic polynomial of degree $O(\sqrt{t})$ for majority constructed by Alman and Williams [AW15], which will be discussed in §8 (see Theorem 4). (An earlier, slightly weaker result by Srinivasan [Sri13] suffices for the main point here.)

Combining Lemma 3 with the first part proves Theorem 2.

The above theorems give non-trivial bounds for degree up to $\Omega(\sqrt{n})$, and the argument appears incapable of handling larger degrees. In fact, lower bounds for higher degrees remain unknown for any explicit function (say in NP). It is known [Vio] that some function in E^{NP} requires probabilistic degree $\Omega(n/\log^2 n)$ for constant error (which is optimal up to the $\log^2 n$ factor, since every function has polynomials of degree n).

Xor lemma. The above results ([Raz87] and Theorem 2) prove non-trivial correlation bounds for polynomials of very high degree $d = n^{\Omega(1)}$. In this sense they address the computational model which is the subject of Question 1, they "just" fail to provide a strong enough bound on the correlation. For other important computational models this would not be a problem: the extensive study of *hardness amplification* has developed many techniques to improve correlation bounds in the following sense: given an explicit function $f : \{0,1\}^n \to \{0,1\}$ that has correlation ϵ with some class $\mathcal{C}_n$ of functions on n bits, construct another explicit function $f' : \{0,1\}^{n'} \to \{0,1\}$, where $n' \approx n$, that has correlation $\epsilon' \ll \epsilon$ with a closely related class $\mathcal{C}_{n'}$ of functions on n' bits (see [GSV18] for a comprehensive list of references on hardness amplification). While the following discussion holds for any hardness amplification, for concreteness we focus on the foremost: Yao's xor lemma. Here $f' : (\{0,1\}^n)^k \to \{0,1\}$ is defined as the xor (or parity, or sum modulo 2) of k independent outputs of f:

$$f'(x^1, \ldots, x^k) := f(x^1) + \cdots + f(x^k) \in \{0,1\}, \qquad x^i \in \{0,1\}^n.$$

The compelling intuition is that, since functions from $\mathcal{C}_n$ have correlation at most ϵ with f, and f' is the xor of k independent evaluations of f, the correlation should decay exponentially with k: $\epsilon' \approx \epsilon^k$. This is indeed the case if one tries to compute $f'(x^1, \ldots, x^k)$ as $g_1(x^1) + \cdots + g_k(x^k)$ where $g_i : \{0,1\}^n \to \{0,1\}, g_i \in \mathcal{C}_n, 1 \leq i \leq k$, but in general a function $g : (\{0,1\}^n)^k \to \{0,1\}, g \in \mathcal{C}_{n'}$, need not have this structure, making proofs of Yao's xor lemma more subtle. If we could prove this intuition true for low-degree polynomials, we could combine this with Theorem 2 to answer affirmatively Question 1 via the function

$$f(x^1, \ldots, x^k) := \text{mod}_3(x^1) + \cdots + \text{mod}_3(x^k) \tag{4}$$

for $k = n$.

Of course the obstacle is that nobody knows whether Yao's xor lemma holds for polynomials. In general it was remarked that "*none of the known hardness amplification results can be applied to the computational models for which we actually can establish the existence of hard functions (i.e. prove lower bounds)" [Vio06a, Page 7]* .

Open question 2. *Does Yao's xor lemma hold for polynomials of degree $d \geq \log_2 n$? For example, let $f : \{0,1\}^n \to \{0,1\}$ satisfy $\text{Cor}(f, n^{1/3}) \leq 1/3$, and for $n' := n^2$ define $f' : \{0,1\}^{n'} \to \{0,1\}$ as $f'(x^1, \ldots, x^n) := f(x^1) + \cdots + f(x^n)$. Is $\text{Cor}(f', \log_2 n') \leq 1/n'$?*

We now discuss why, despite the many alternative proofs of Yao's xor lemma that are available (e.g., [GNW95]), we cannot apply any of them to the computational model of low-degree polynomials. To prove that f' has correlation at most ϵ' with some class of functions, all known proofs of the lemma need (a slight modification of) the functions in the class to compute the majority function on $> (1/\epsilon')^2$ bits. However, the majority function on this many bits requires polynomials of degree $\Omega(1/\epsilon')$ by Theorem 1. This means that known proofs can only establish correlation bounds $\epsilon' > 1/\sqrt{n}$, failing to answer Question 2. More generally, the work [GSV18] completes a line of research initiated in [Vio06b, SV10] and shows that computing the majority function on $\Omega(1/\epsilon')$ bits is necessary for a central class of hardness amplification proofs known as *black-box*. (Improving the bound to $\Omega(1/\epsilon')^2$ is open.)

Xor lemmas are however known for polynomials of small degree $d \ll \log n$ [Vio06c] (cf. [VW08]). More recently Chattopadhyay, Hatami, Hosseini, Lovett, and Zuckerman [CHH$^+$20] introduced a novel technique with which they established an xor lemma for low-degree polynomials and consequently new correlation bounds. In particular, they proved that the correlation of the xor of two majority functions with *constant-degree* polynomials is $(\log^{O(1)} n)/n$, a result for which previous xor lemmas do not seem sufficient. Note that the bound is indeed about the square of the bound in Theorem 1.

The key ingredient in the approach in [CHH$^+$20] is a structural result about the Fourier spectrum of low-degree polynomials. They show that for any n-variate polynomial p of degree $\leq d$, there is a set S of variables such that almost all of the Fourier mass of p lies on Fourier coefficients that intersect with S, and the size of S is exponential in d. This limits their results to degree $d \leq \log n$ (roughly the same setting as the next section). Further, they conjecture that the size of S needs to be just polynomial in d. If true, this would expand their results to degrees polynomial in n, and yield exciting new correlation bounds. However, a counterexample to their conjecture is given in [IPV22]. In fact, [IPV22] rules out even weaker parameters and shows that what is proved in [CHH$^+$20] is essentially the best possible.

1.2 Negligible correlation $\epsilon \ll 1/\sqrt{n}$ but small degree $d \ll \log n$

It is easy to prove exponentially small correlation bounds for polynomials of degree 1, for example the *inner product* function $\mathrm{IP} : \{0,1\}^n \to \{0,1\}$, defined for even n as

$$\mathrm{IP}(x_1, \ldots, x_n) := x_1 \cdot x_2 + x_3 \cdot x_4 + \cdots + x_{n-1} \cdot x_n,$$

satisfies $\mathrm{Cor}(\mathrm{IP}, 1) = 2^{-n/2}$. Already obtaining exponentially small bounds for polynomials of constant degree appears to be a challenge. The first such bounds come from the famed work by Babai, Nisan, and Szegedy [BNS92] proving exponentially small correlation bounds between polynomials of degree $d := \epsilon \log n$ and, for $k := d + 1$, the *generalized inner product* function $\mathrm{GIP}_k : \{0,1\}^n \to \{0,1\}$,

$$\mathrm{GIP}_k(x_1, \ldots, x_n) := \prod_{i=1}^{k} x_i + \prod_{i=k+1}^{2k} x_i + \cdots + \prod_{i=n-k+1}^{n} x_i,$$

assuming for simplicity that n is a multiple of k. The intuition for this correlation bound is precisely that behind Yao's xor lemma (cf. §1.1): (i) any polynomial of degree d has correlation that is bounded away from 1 with any monomial of degree $k = d + 1$ in the definition of GIP, and (ii) since the monomials in the definition of GIP are on disjoint sets of variables, the correlation decays exponentially. (i) is not hard to establish formally. With some work, (ii) can also be established to obtain the following theorem.

Theorem 3. *[BNS92] For every n, d,* $\mathrm{Cor}(\mathrm{GIP}_{d+1}, d) \leq \exp\left(-\Omega(n/4^d \cdot d)\right)$.

When $k \gg \log n$, GIP is almost always 0 on a uniform input, and thus GIP is not a candidate for having small correlation with respect to the uniform distribution with polynomials of degree $d \gg \log n$.

Our exposition of the results in [BNS92] differs in multiple ways from the original. First, [BNS92] does not discuss polynomials but rather number-on-forehead multiparty protocols. The results for polynomials are obtained via the observation of Håstad and Goldmann [HG91, Proof of Lemma 4] that k-party protocols can efficiently simulate polynomials of degree

$k-1$, for any input partition. Second, [BNS92] presents the proof with a different language. Alternative languages have been put forth in a series of papers [CT93, Raz00, VW08], with the last one stressing the above intuition and the connections between multiparty protocols and polynomials.

Bourgain [Bou05] later proves bounds similar to those in Theorem 3 but for the mod_3 function. A minor mistake in his proof is corrected by F. Green, Roy, and Straubing [GRS05].

Theorem 4. *[Bou05, GRS05] For every n, d there is a distribution D on $\{0,1\}^n$ such that $\text{Cor}_D(\text{mod}_3, d) \leq \exp\left(-n/c^d\right)$, where c is an absolute constant.*

A random sample from the distribution D in Theorem 4 is obtained as follows: toss a fair coin, if "heads" then output a uniform $x \in \{0,1\}^n$ such that $\text{mod}_3(x) = 1$, if "tails" then output a uniform $x \in \{0,1\}^n$ such that $\text{mod}_3(x) = 0$. The value $c = 8$ in [Bou05, GRS05] is later improved to $c = 4$ in [Vio06c, Cha07] (cf. [VW08]). [Vio06c] also presents the proof in a different language.

Theorem 4 appears more than a decade after Theorem 3. However, Noam Nisan (personal communication) observes that in fact the first easily follows from the latter.

Sketch of Nisan's proof of Theorem 4. Grolmusz's [Gro95] extends the results in [BNS92] and shows that there is a distribution D' on $\{0,1\}^n$ such that for $k := d+1$ the function

$$\text{mod}_3 \wedge_k (x_1, \ldots, x_n) := \text{mod}_3 \left(\prod_{i=1}^{k} x_i, \prod_{i=k+1}^{2k} x_i, \ldots, \prod_{i=n-k+1}^{n} x_i \right)$$

has correlation $\exp(-n/c^d)$ with polynomials of degree d, for an absolute constant c. A proof of this can also be found in [VW08, §3.3]. An inspection of the proof reveals that, with respect to another distribution D'' on $\{0,1\}^n$, the same bound applies to the function

$$\text{mod}_3\text{mod}_2(x_1, \ldots, x_n) := \text{mod}_3(x_1 + \cdots + x_k, x_{k+1} + \cdots + x_{2k}, \ldots, x_{n-k+1} + \cdots + x_n)$$

where we replace "and" with "parity" (the sums in the input to mod_3 are modulo 2).

Now consider the distribution D on $\{0,1\}^{n/k}$ that D'' induces on the input to mod_3 of length n/k. (To sample from D one can sample from D'', perform the n/k sums modulo 2, and return the string of length n/k.) Suppose that a polynomial $p(y_1, \ldots, y_{n/k})$ of degree d has correlation ϵ with the mod_3 function with respect to D. Then the polynomial

$$p'(x_1, \ldots, x_n) := p(x_1 + \cdots + x_k, x_{k+1} + \cdots + x_{2k}, \ldots, x_{n-k+1} + \cdots + x_n)$$

has degree d and correlation ϵ with the mod_3mod_2 function with respect to the distribution D'' on $\{0,1\}^n$. Therefore $\epsilon \leq \exp(-n/c^d)$. □

The "squaring trick." Most or all proofs of very small correlation bounds (including theorems 3, 4, and 8) use a common technique which we now discuss also to highlight its limitation. The idea is to reduce the challenge of proving a correlation bound for a polynomial of degree d to that of proving related correlation bounds for polynomials of degree $d-1$, by *squaring*. To illustrate, let $f : \{0,1\}^n \to \{0,1\}$ be any function and $p : \{0,1\}^n \to \{0,1\}$ a polynomial of degree d. Using the extremely convenient notation $e[z] := (-1)^z$, we write the correlation between f and p with respect to the uniform distribution as

$$\mathrm{Cor}(f,p) = \left| \Pr_{x\in\{0,1\}^n}[f(x) = p(x)] - \Pr_{x\in\{0,1\}^n}[f(x) \neq p(x)] \right| = \left| E_{x\in\{0,1\}^n} e[f(x) + p(x)] \right|.$$

If we square this quantity, and use that $E_Z[g(Z)]^2 = E_{Z,Z'}[g(Z)\cdot g(Z')]$, we obtain

$$\mathrm{Cor}(f,p)^2 = E_{x,y\in\{0,1\}^n} e[f(x) + f(y) + p(x) + p(y)].$$

Letting now $y = x + h$ we can rewrite this as

$$\mathrm{Cor}(f,p)^2 = E_{x,h\in\{0,1\}^n} e[f(x) + f(x+h) + p(x) + p(x+h)].$$

The crucial observation is now that, for every fixed h, the polynomial $p(x) + p(x+h)$ has degree $d-1$ in x, even though $p(x)$ has degree d. For example, if $d = 2$ and $p(x) = x_1x_2+x_3$, we have $p(x)+p(x+h) = x_1x_2+x_3+(x_1+h_1)(x_2+h_2)+(x_3+h_3) = x_1h_2+h_1x_2+h_1h_2+h_3$,which indeed has degree $d-1 = 1$ in x. Thus we managed to reduce our task of bounding from above $\mathrm{Cor}(f,p)$ to that of bounding from above a related quantity which involves polynomials of degree $d-1$, specifically the average over h of the correlation between the function $f(x) + f(x+h)$ and polynomials of degree $d-1$. To iterate, we apply the same trick, this time coupled with the Cauchy-Schwarz inequality $E[Z]^2 \leq E[Z^2]$:

$$\begin{aligned} \mathrm{Cor}(f,p)^4 &= E_{x,h} e[f(x) + f(x+h) + p(x) + p(x+h)]^2 \\ &\leq E_h \left[E_x e[f(x) + f(x+h) + p(x) + p(x+h)]^2 \right]. \end{aligned}$$

We can now repeat the argument in the inner expectation, further reducing the degree of the polynomial. After d repetitions, the polynomial p becomes a constant. After one more, a total of $d+1$ repetitions, the polynomial p "disappears" and we are left with a certain expectation involving f, known as the "Gowers norm" of f and introduced independently in [Gow98, Gow01] and in [AKK$^+$03]:

$$\mathrm{Cor}(f,p)^{2^{d+1}} \leq E_{x,h_1,h_2,\ldots,h_{d+1}} e\left[\sum_{S\subseteq[d+1]} f\left(x + \sum_{i\in S} h_i \right) \right]. \qquad (5)$$

For interesting functions f, the expectation in the right-hand side of (5) can be easily shown to be small, sometimes exponentially in n, yielding correlation bounds. For example, applying this method to the generalized inner product function gives Theorem 3, while a complex-valued generalization of the method can be applied to the mod_3 function to obtain Theorem 4. This concludes the exposition of this technique; see, e.g., [Vio06c, VW08] for more details.

This "squaring trick" for reducing the analysis of a polynomial of degree d to that of an expression involving polynomials of lower degree $d-1$ dates back at least to the work by Weyl at the beginning of the 20th century; for an exposition of the relevant proof by Weyl, as well as pointers to his work, the reader may consult [GR07]. This method was apparently introduced in computer science by Babai, Nisan, and Szegedy [BNS92], and employed later by various researchers in different contexts, see [VW08] for discussion.

The obvious limitation of this technique is that, to bound the correlation with polynomials of degree d, it squares the correlation d times; this means that the bound on the correlation will be $\exp(-n/2^d)$ at best: nothing for degree $d = \log_2 n$. This bound is almost achieved by [BNS92] which gives an explicit function f such that $\mathrm{Cor}(f, d) \leq \exp(-\Omega(n/2^d \cdot d))$. The extra factor of d in the exponent arises because of the different context of multiparty protocols in [BNS92], but a similar argument, given in [Vio06c] and also appearing in [VW08], establishes the following stronger bound.

Theorem 5. *[Vio06c] There is an explicit $f \in P$ such that for every n and d, and U the uniform distribution over $\{0,1\}^n$, $\mathrm{Cor}_U(f, d) \leq \exp(-\Omega(n/2^d))$.*

The function $f : \{0,1\}^n \to \{0,1\}$ in Theorem 5 takes as input an index $i \in \{0,1\}^{\epsilon n}$ and a seed $s \in \{0,1\}^{(1-\epsilon)n}$, and outputs the i-th output bit of a certain pseudorandom generator on seed s [NN93] (Theorem 7 in §2). The natural question of whether these functions have small correlation with polynomials of degree $d \gg \log_2 n$ is answered negatively also in [Vio06c]: for a specific implementation of the generator [GV04, HV06, Hea08], the associated function $f : \{0,1\}^n \to \{0,1\}$ satisfies $\mathrm{Cor}(f, \log^{O(1)} n) \geq 1 - o(1)$.

1.3 Correlating with symmetric functions

Most of the correlation bounds discussed in §1.1 and §1.2 are for functions that are symmetric: their value depends only on the number of input bits that are 1. In this section we prove that such functions somewhat correlate with low-degree polynomials. We begin with the result that any symmetric function has large correlation with polynomials of degree $O(\sqrt{n})$, thus excluding symmetric functions from the candidates to the dream setting of parameters $\mathrm{Cor}(f, \epsilon \cdot n) \leq \exp(-\epsilon \cdot n)$.

Under the uniform distribution this result was proved by Viola. A slight simplification of his proof, by Wigderson, is presented in [Vio09a, Vio09b]. In those works it was also suggested to investigate whether the result extends to other distributions. Such an extension is obtained by Srinivasan [Sri13] and with slightly better parameters by Alman and Williams [AW15] (Theorem 1.2).

Theorem 4. *[AW15] (Theorem 1.2) For any d, ϵ and any symmetric $f : \{0,1\}^n \to \{0,1\}$ we have $\mathrm{Cor}(f, O(\sqrt{n \log 1/\epsilon})) \geq 1 - \epsilon$.*

This result however does not give information for degree less than $\sqrt{n}$, a regime of great interest as we have seen. We obtain the following tradeoff, which also improves on Theorem 9 in [CP16].

Theorem 5. *For any d and any symmetric $f : \{0,1\}^n \to \{0,1\}$ we have $\mathrm{Cor}(f, d) \geq 2^{-n\Omega(\log^2 n)/d^2}$.*

Except for the $\log^2 n$ factor, this tradeoff includes the bound in Theorem 4. We suspect that the factor can be removed, but it is not clear how to do that at this moment. In the remainder of this section we present and discuss the proof of Theorem 5.

A toy case: Non-zero correlation with any function. Let $f : \{0,1\}^n \to \{0,1\}$ be a function, not necessarily symmetric. We note that $\mathrm{Cor}(f, 1) \geq \Omega(2^{-n}) > 0$. To verify this, consider the following way to sample a polynomial witnessing the correlation. Pick a uniform $y \in \{0,1\}^n$, and uniform $v_1, v_2, \ldots, v_n \in \{0,1\}$. Output the polynomial $\sum_i (y_i - x_i) \cdot v_i + f(y)$. Note that if $y \neq x$ then some $y_i - x_i = 1$, and then the polynomial computes a uniform bit thanks to v_i, giving correlation 0. On the other hand, if $y = x$ then all the v_i cancel out and we get $f(y)$. The probability that $y = x$ is obviously 2^{-n} and this gives the claimed result.

This simple result highlights the "power of a random bit" which is available to polynomials modulo 2 but not to polynomials over other domains such as the reals, indeed, as we shall see in § 1.5, the correlation between the parity function and real polynomials of low degree is zero.

Warm-up: Proof of the weaker bound $2^{-n\Omega(\log d)/d}$. This warm-up can be seen as an extension of the toy case to blocks, exploiting the symmetry of the function. Divide the input x of n bits into b blocks of n/b bits, and let y_i denote the n/b variables in block i. For $w \in \{0, 1, \ldots, n/b\}$ let $p_{\neq w}$ be the polynomial in n/b variables that is equal to 1 if the input Hamming weight is not equal to w, and 0 otherwise.

To sample a polynomial witnessing the correlation proceed as follows. Pick uniformly at random $w_1, w_2, \ldots, w_b \in \{0, 1, \ldots, n/b\}$. Pick uniformly random $v_1, v_2, \ldots, v_b \in \{0,1\}$. Output

$$f(\sum_{i=1}^{b} w_i) + \sum_{i=1}^{b} p_{\neq w_i}(y_i) \cdot v_i.$$

Here for an integer m we write $f(m)$ for the value of f on any input of Hamming weight m. Each $p_{\neq w_i}$ is on n/b bits and hence can be computed by a polynomial of the same degree; so each sample has degree n/b.

To analyze, fix an input x. Note that if at least one of the guesses w_i is wrong (i.e., it does not equal the Hamming weight of y_i) then as in the toy case we obtain correlation 0 because of the random choice v_i. But if all the guesses are right then all the v_i are multiplied by 0 and we get $f(x)$. The probability of this is $(n/b + 1)^{-b}$.

Given d, set $b := n/d$ to obtain a distribution on degree-d polynomials with correlation $(d + 1)^{-n/d}$.

Proof of the claimed bound (Theorem 5). The idea is to use the probabilistic polynomials from Theorem 4 to reduce the degree of the $p_{\neq w_i}$ at the cost of some error. However, we can't afford to set the error so small to take a simple union bound. We will condition on the number t of wrong guesses, and show that the probability of having t wrong guesses compares favorably to the probability that each of the t corresponding polynomials makes a mistake. Details follow.

The distribution on polynomials is defined as before, except that for each $p_{\neq w}$ we now pick the probabilistic polynomial from Theorem 4 with degree $O(\sqrt{(n/b) \log n})$ such that on any input, the probability that this polynomial outputs the wrong value is α/n for a small enough constant $\alpha > 0$ to be set later . Fix x and let W be the random number of wrong guesses w_i. Conditioned on $W = 0$, the probability that the final polynomial computes the function correctly is by a union bound $\geq 1 - b \cdot \alpha/n \geq 1 - \alpha \geq 0.9$ for small enough α. Conditioned on $W = t$ for $t > 0$, the probability that the polynomial does not compute a random bit is at most $(\alpha/n)^t$, because we need each of the t corresponding polynomials to be wrong to cancel the associated v_i. Hence the polynomial computes incorrectly with probability at most $1/2 + (\alpha/n)^t$

Hence the probability that we compute f incorrectly is at most

$$\Pr[W = 0] \cdot (1/2 - 0.4) + \sum_{t=1}^{b} \Pr[W = t] \cdot (1/2 + (\alpha/n)^t).$$

Collecting the $1/2$ terms we obtain a bound of

$$1/2 - 0.4 \Pr[W = 0] + \sum_{t=1}^{b} \Pr[W = t] \cdot (\alpha/n)^t.$$

Now we use the bound

$$\Pr[W = t] = \frac{\binom{b}{t}(n/b)^t}{(1 + n/b)^b} \leq \frac{(3n)^t}{(1 + n/b)^b}.$$

Hence the summation above is

$$\leq \frac{1}{(1 + n/b)^b} \sum_{t=1}^{b} (3\alpha)^t \leq \frac{0.01}{(1 + n/b)^b} = 0.01 \Pr[W = 0],$$

for a small enough α. Consequently the error probability is at most

$$1/2 - \Omega(1/(1 + n/b)^b).$$

Finally, given d set b so that $O(\sqrt{(n/b)\log n}) \leq d$. This is possible unless $d = O(\sqrt{\log n})$. In the latter case we can, say, use the polynomials from the proof of the warm-up case. This setting of b makes the error probability at most

$$1/2 - ((\log n)/d^2)^{\Omega(n \log n)/d^2},$$

as desired.

1.4 More on mod_3

For the mod_3 function, we can summarize the bounds from the previous sections as follows:

$$2^{-n\Omega(\log n)^2/d^2} \leq \mathrm{Cor}(\mathrm{mod}_3, d) \leq \min\{2^{-\Omega(n/2^d)}, O(d/\sqrt{n})\}. \tag{6}$$

Improving any of these bounds substantially would be a major advance, but seems to require new ideas.

One possible approach is aiming for *exact results.* Low-degree polynomials appear simple enough that instead of merely proving bounds one might be able to pinpoint the best polynomials. In this spirit, it is conjectured in [IPV22] that the polynomials that maximize correlation with mod_3 are *symmetric* (under the same natural distribution as in Section 1.2, and for all n congruent to 3 or 9 modulo 12; a cleaner version of their conjecture is precisely stated below). If the conjecture is true then very strong correlation bounds would follow, since it is known that degree-d polynomials have correlation at most $d \cdot 2^{-\Omega(n/d^2)}$ with mod_3 (see [IPV22] for a proof). Note that this bound matches the leftmost expression in Equation (6) up to lower order terms, and thus is essentially tight.

Moreover, this conjecture does not require distinguishing classical from non-classical polynomials, thus circumventing the "barrier" raised in [BL15] and discussed in § 1. Further, [IPV22] prove their conjecture for polynomials of degree $d = 2$.

To state the conjecture and results in [IPV22] more precisely, and also for context, we mention that for functions like mod_3 it is sometimes convenient to work with the following complex-valued version of correlation (sometimes referred to as an exponential sum):

$$C(p) := |E_{x\in\{0,1\}^n}[(-1)^{p(x)}\omega^{\sum_i x_i}]|,$$

where ω is the third root of unity $e^{\sqrt{-1}2\pi/3}$ and $|.|$ is the complex norm. This expression has the advantage that if p is the sum of polynomials over disjoint variables then it multiplies. This for example leads to a simple proof of the correlation with linear polynomials, see e.g. [IPV22]. In general, it appears to be a slightly more convenient quantity to work with.

On the other hand, $C(p)$ is closely related to $\mathrm{Cor}(\mathrm{mod}_3, p)$. An upper bound on the former implies an upper bound on the latter, a fact that is sketched in several works and fully proved in [VW08, IPV22]. A weak form of the converse holds as well, where the distribution depends on the polynomials in a limited way.

The conjecture in [IPV22] is that C is maximized by symmetric polynomials, and the conjecture is proved in [IPV22] for $d = 2$.

Open question 3. *Is the maximum of C over degree-d polynomials attained by symmetric polynomials for every n and $d \geq 3$?*

The case $d = 3$ is already open but seems within reach; partial progress is reported in [IPV22].

1.5 Real advantage

Instead of polynomials modulo 2 one can consider polynomials over the integers or the reals where a non-boolean output is always counted as a mistake. Perhaps surprisingly, most or all of the open questions discussed until now are also open for such real polynomials. The challenge of proving stronger correlation bounds for real polynomials was raised in [RV13]. They specifically asked what is the correlation with the *parity function.* Also, they showed that when the degree d of the polynomial is very small, at most $0.5 \log\log n$, then in fact this correlation is not just small, but *zero.* That is, the polynomial cannot correlate better than a constant function. This is unlike the case of polynomials over finite fields, which have non-zero correlation with *any* function, thanks to the "power of a random bit," which we saw in § 1.3.

The follow-up work [MNV16] improved these bounds and showed that the correlation is zero up to degree $\log n/(15 \log\log n)$. On the other hand the construction mentioned in \S 1.3 of polynomials of degree $O(\sqrt{n})$ that have constant correlation with symmetric functions such as parity under the uniform distribution (proved in [Vio09a, Vio09b]) in fact gives real polynomials.

This leads to the following question:

Open question 4. *What is the minimum d such that there exists a degree-d real polynomial with non-zero correlation with the parity function?*

Summarizing the discussion above, it is known that $\log n/(15 \log\log n) \le d \le O(\sqrt{n})$.

1.6 Other works

There are many papers on correlation bounds we have not discussed. In general, we have chosen to focus on polynomials modulo 2 for simplicity, whereas many previous works focus on polynomials modulo $m \neq 2$. Indeed, many of the results we discussed, such as Theorems 2, 4, and the results in \S 1 can be extended to polynomials modulo $m \neq 2$. For example, Theorem 4 extends to polynomials modulo m vs. the mod_q function for any relatively prime m and q, while Theorem 2 extends to polynomials modulo m vs. the mod_q function for any prime m and q not a power of m.

F. Green [Gre04, Theorem 3.10] computes exactly the correlation between the parity function and quadratic ($d = 2$) polynomials over $\{0, 1, 2\}$. [Gre04] further discusses the difficulties currently preventing an extension of the result to degree $d > 2$ or to polynomials over fields different from $\{0, 1, 2\}$, while [GR10] studies the structure of quadratic polynomials over $\{0, 1, 2\}$ that correlate with the parity function best.

A natural question, also asked in [AB01], is whether the symmetric polynomials mod m, for odd m, that correlate best with parity are symmetric. [Gre04] answers this negatively for degree 2. The work [GKV17] gives a negative answer for more degrees, including degrees that are relevant to Question 1. The same work also reports results of a computer search to determine the polynomials modulo 2 that correlate best with mod_3. Interestingly, the polynomials that can be computed are symmetric. Indeed, this was a starting point for the work [IPV22] and the conjecture therein which we discussed in \S 1.4.

Other special classes of polynomials are studied in [CGT96, GT06, BEHL08]. Finally, we mention that several papers [Vio07, Han06, LS11, ST18] prove correlation bounds against *sparse* polynomial.

2 Pseudorandom generators

In this section we discuss pseudorandom generators for polynomials and their connections to correlation bounds. Pseudorandom generators are fascinating algorithms that stretch short input seeds into much longer sequences that "look random;" naturally, here we interpret "look random" as "look random to polynomials," made formal in the next definition.

Definition 6 (Generator)**.** *A* generator *G* that fools polynomials of degree $d = d(n)$ with error $\epsilon = \epsilon(n)$ and seed length $s = s(n)$ *is a an algorithm running in time polynomial in its* *output length such that for every n: (i) G maps strings of length $s(n)$ to strings of length n, and (ii) for any polynomial $p : \{0,1\}^n \to \{0,1\}$ of degree d we have*

$$\left| \Pr_{S \in \{0,1\}^s}[p(G(S)) = 1] - \Pr_{U \in \{0,1\}^n}[p(U) = 1] \right| \leq \epsilon. \tag{7}$$

For brevity, we write $G : \{0,1\}^s \to \{0,1\}^n$ for a generator with seed length $s(n)$.

Ideally, we would like generators that fool polynomials of large degree d with small error ϵ and small seed length s. We discuss below various connections between obtaining such generators and correlation bounds, but first we point out a notable difference: while for correlation bounds we do have results for large degree $d \gg \log n$ (e.g., Theorem 2), we know of no generator that fools polynomials of degree $d \geq \log_2 n$, even with constant error ϵ.

Open question 5. *Is there a generator $G : \{0,1\}^{n/2} \to \{0,1\}^n$ that fools polynomials of degree $\log_2 n$ with error $1/3$?*

While the smaller the error ϵ the better, generators for constant error are already of great interest; for example, a constant-error generator that fools small circuits is enough to derandomize BPP. However, we currently seem to be no better at constructing generators that fool polynomials with constant error than generators with shrinking error, such as $1/n$.

We now discuss the relationship between generators and correlation bounds, and then present the known generators.

From pseudorandomness to correlation. It is easy to see and well-known that a generator implies a worst-case lower bound, i.e., an explicit function that cannot be computed by (essentially) the class of functions fooled by the generator. The following simple observation, which does not seem to have appeared before [Vio09c, §3], shows that in fact a generator implies even a correlation bound. We will use it later to obtain new candidates for answering Question 1.

Observation 1. *Suppose that there is a generator $G : \{0,1\}^{n-\log n-1} \to \{0,1\}^n$ that fools polynomials of degree $\log_2 n$ with error $0.5/n$. Then the answer to Question 1 is affirmative.*

Proof sketch. Let D be the distribution on $\{0,1\}^n$ where a random $x \in D$ is obtained as follows: toss a fair coin, if "heads" then let x be uniformly distributed over $\{0,1\}^n$, if "tails" then let $x := G(S)$ for a uniformly chosen $S \in \{0,1\}^{n-\log n-1}$. Define the function $f : \{0,1\}^n \to \{0,1\}$ as $f(x) = 1$ if and only if there is $s \in \{0,1\}^{n-\log n-1}$ such that $G(s) = x$; f is easily seen to be in NP. It is now a routine calculation to verify that any function $t : \{0,1\}^n \to \{0,1\}$ that satisfies $\mathrm{Cor}_D(f,t) \geq 1/n$ has advantage at least $0.5/n$ in distinguishing the output of the generator from random. Letting t range over polynomials of degree $\log_2 n$ concludes the proof. □

From correlation to pseudorandomness. The celebrated construction by Nisan [Nis91] (cf. [NW94]) shows that a sufficiently strong correlation bound with respect to the uniform distribution can be used to obtain a generator that fools polynomials. However, to obtain a generator $G : \{0,1\}^s \to \{0,1\}^n$ against polynomials of degree d, [Nis91] in particular needs a function f on $m \leq n$ input bits that has correlation at most $1/n$ with polynomials of degree d. Thus, the current correlation bounds are not strong enough to obtain generators for polynomials of degree $d \geq \log_2 n$. It is a pressing open problem to determine whether alternative constructions of generators are possible, ideally based on constant correlation bounds such as Theorem 2. Here, an uncharted direction is to understand which distributions D enable one to construct generators starting from correlation bounds with respect to D.

In [Vio19] it is shown that computing majority on many bits is required for black-box proofs of pseudorandom generators, even if the latter only have *constant error,* but only for NW-style [NW94]constructions that are "seed revealing." This suggests that there may not be a way around correlation bounds even for constructing generators with constant error. However, extending the results in [Vio19] to other constructions, such as NW-style that do not reveal their seed remains open.

Nisan's construction is however sharp enough to give non-trivial generators based on the current correlation bounds such as Theorem 3. Specifically, Luby, Veličković, and Wigderson [LVW93, Theorem 2] obtain generators for polynomials that have arbitrary degree but at most $n^{\alpha \cdot \log n}$ terms for a small absolute constant $\alpha > 0$; a different proof of this result appears in the paper [Vio07] which we already mentioned in §1.6. Albeit falling short of answering Question 5 (cf. §1.6), this generator [LVW93, Theorem 2] does fool polynomials of constant degree. However, its seed length, satisfying $n = s^{O(\log s)}$, has been superseded in this case by recent developments, which we now discuss.

Generators for degree $d \ll \log n$. Naor and Naor [NN93] construct a generator that fools polynomials of degree 1 (i.e., linear) with a seed length that is optimal up to constant factors – a result with a surprising range of applications (cf. references in [BSVW03]).

Theorem 7. *[NN93] There is a generator $G : \{0,1\}^{O(\log n)} \to \{0,1\}^n$ that fools polynomials of degree 1 with error $1/n$.*

Later, Alon et al. [AGHP92] give three alternative constructions. A nice one is $G(a,b)_i := \langle a^i, b\rangle$ where $\langle \cdot, \cdot \rangle$ denotes inner product modulo 2, $a, b \in \{0,1\}^\ell$ for $\ell = O(\log n)$, and a^i denotes the result of considering a as an element of the field with 2^ℓ elements, and raising it to the power i.

Bogdanov and Viola introduced [BV10] a simple paradigm to fool polynomials of higher degree d: just sum together d generators for polynomials of degree 1. This paradigm has been analyzed in [BV10, Lov08, Vio09c], with the last work giving the following result.

Theorem 8. *[Vio09c] Let $G : \{0,1\}^s \to \{0,1\}^n$ be a generator that fools polynomials of degree 1 with error ϵ. Then $G_d : (\{0,1\}^s)^d \to \{0,1\}^n$ defined as*

$$G_d(x^1, x^2, \ldots, x^d) := G(x^1) + G(x^2) + \cdots + G(x^d)$$

fools polynomials of degree d with error $\epsilon_d := 16 \cdot \epsilon^{1/2^{d-1}}$, where "+" denotes bit-wise xor.

In particular, the combination of Theorems 7 and 8 yields generators $G : \{0,1\}^s \to \{0,1\}^n$ that fool polynomials of degree d with error $\epsilon_d = 1/n$ and seed length $s = O(d \cdot 2^d \cdot \log(n))$.

Proof sketch of Theorem 8. This proof uses the notation $e[z] := (-1)^z$ and some of the techniques presented at the end of §1.2. First, let us rewrite Inequality (7) in the Definition 6 of a generator as

$$\left|E_{S\in\{0,1\}^s} e[p(G_d(S))] - E_{U\in\{0,1\}^n} e[p(U)]\right| \leq \epsilon_d/2. \tag{8}$$

To prove Inequality (8), we proceed by induction on the degree d of the polynomial $p : \{0,1\}^n \to \{0,1\}$ to be fooled. The inductive step is structured as a case analysis based on the value $\tau := \mathrm{Cor}_U(p,0) = |E_{U\in\{0,1\}^n} e[p(U)]|$.

If τ is large then p correlates with a constant, which is a polynomial of degree lower than d, and by induction one can prove the intuitive fact that G_{d-1} fools p. This concludes the overview of the proof in this case.

If τ is small we reason as follows. Let us denote by W the output of G_{d-1} and by Y an independent output of G, so that the output of G_d is $W+Y$. We start by an application of the Cauchy-Schwarz inequality:

$$\mathrm{E}_{W,Y}\, e\left[p(W+Y)\right]^2 \;\leq\; E_W\left[\mathrm{E}_Y\, e\left[p(W+Y)\right]^2\right] \;=\; \mathrm{E}_{W,Y,Y'}\, e\left[p(W+Y)+p(W+Y')\right], \tag{9}$$

where Y' is independent from and identically distributed to Y. Now we observe that for every fixed Y and Y', the polynomial $p(U+Y)+p(U+Y')$ has degree $d-1$ in U, though p has degree d. Since by induction W fools polynomials of degree $d-1$ with error ϵ_{d-1}, we can replace W with the uniform distribution $U \in \{0,1\}^n$:

$$\mathrm{E}_{W,Y,Y'}\, e\left[p(W+Y)+p(W+Y')\right] \quad\leq\quad \mathrm{E}_{U,Y,Y'}\, e\left[p(U+Y)+p(U+Y')\right] \;+\; \epsilon_{d-1}. \tag{10}$$

At this point, a standard argument shows that

$$\mathrm{E}_{U,Y,Y'}\, e\left[p(U+Y)+p(U+Y')\right] \quad\leq\quad \mathrm{E}_{U,U'}\, e\left[p(U)+p(U')\right] \;+\; \epsilon^2 \quad=\quad \tau^2 \;+\; \epsilon^2. \tag{11}$$

Therefore, chaining Equations (9), (10), and (11), we have that

$$\left|\mathrm{E}_{W,Y}\, e\left[p(W+Y)\right] - \mathrm{E}_U\, e\left[p(U)\right]\right| \;\leq\; \left|\mathrm{E}_{W,Y}\, e\left[p(W+Y)\right]\right| + \tau \;\leq\; \sqrt{\tau^2+\epsilon^2+\epsilon_{d-1}} + \tau.$$

This proves Inequality (8) for a suitable choice of ϵ_d, concluding the proof in this remaining case. □

Observe that Theorem 8 gives nothing for polynomials of degree $d = \log_2 n$. The reason is that its proof again relies on the "squaring trick" discussed in §1.2. But it is still not known whether the construction in Theorem 8 fools polynomials of degree $d \geq \log_2 n$.

Open question 6. *Does the sum of $d \gg \log n$ copies of a generator $G : \{0,1\}^s \to \{0,1\}^n$ that fools polynomials of degree 1 with error $1/n$ fools polynomials of degree d with error $1/3$?*

The recent work [DV22] gives a positive answer *over large fields,* when G is "algebraic." This leads to generators with optimal seed length, improving on a line of works starting with a seminal paper by Bogdanov [Bog05]. But the case of small fields remains wide open. An interesting special case is that of $d = 2$.

Open question 7. *What is the minimum ϵ' such that the sum of two generators for degree 1 polynomials with error ϵ fools the Inner Product function on n bits with error ϵ'?*

Theorem 8 implies $\epsilon' \leq O(\sqrt{\epsilon})$. Whether the square-root loss can be improved seems another basic question.

Acknowledgments. We are grateful to Noam Nisan for his permission to include his proof of Theorem 4. We also thank Frederic Green, Peter Ivanov, Shachar Lovett, Rocco Servedio, and Avi Wigderson for helpful comments and discussions.

References

[AB01] Noga Alon and Richard Beigel. Lower bounds for approximations by low degree polynomials over Z_m. In *IEEE Conf. on Computational Complexity (CCC)*, pages 184–187, 2001.

[AB09] Sanjeev Arora and Boaz Barak. *Computational Complexity.* Cambridge University Press, 2009. A modern approach.

[AGHP92] 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.

[AKK$^+$03] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing low-degree polynomials over GF(2). In *7th Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM)*, volume 2764 of *Lecture Notes in Computer Science*, pages 188–199. Springer, 2003.

[AW08] Scott Aaronson and Avi Wigderson. Algebrization: a new barrier in complexity theory. In *40th ACM Symp. on the Theory of Computing (STOC)*, pages 731–740, 2008.

[AW15] Josh Alman and Ryan Williams. Probabilistic polynomials and hamming nearest neighbors. In *IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 136–150, 2015.

[BEHL08] Ido Ben-Eliezer, Rani Hod, and Shachar Lovett. Random low degree polynomials are hard to approximate. Manuscript, 2008.

[Bei93] Richard Beigel. The polynomial method in circuit complexity. In *8th Structure in Complexity Theory Conference*, pages 82–95. IEEE, 1993.

[BGS75] Theodore Baker, John Gill, and Robert Solovay. Relativizations of the *P=?NP* question. *SIAM J. on Computing*, 4(4):431–442, 1975.

[BIJ$^+$21] Jaroslaw Blasiok, Peter Ivanov, Yaonan Jin, Chin Ho Lee, Rocco A. Servedio, and Emanuele Viola. Fourier growth of structured f2-polynomials and applications. In *Workshop on Randomization and Computation (RANDOM)*, 2021.

[BL15] Abhishek Bhowmick and Shachar Lovett. Nonclassical polynomials as a barrier to polynomial lower bounds. In *IEEE Conf. on Computational Complexity (CCC)*, pages 72–87, 2015.

[BNS92] László Babai, Noam Nisan, and Márió Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. of Computer and System Sciences*, 45(2):204–232, 1992.

[Bog05] Andrej Bogdanov. Pseudorandom generators for low degree polynomials. In *ACM Symp. on the Theory of Computing (STOC)*, pages 21–30, 2005.

[Bou05] Jean Bourgain. Estimation of certain exponential sums arising in complexity theory. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 340(9):627–631, 2005.

[BS90] Ravi B. Boppana and Michael Sipser. The complexity of finite functions. In *Handbook of theoretical computer science, Vol. A*, pages 757–804. Elsevier, Amsterdam, 1990.

[BSVW03] Eli Ben-Sasson, Madhu Sudan, Salil Vadhan, and Avi Wigderson. Randomness-efficient low degree tests and short PCPs via epsilon-biased sets. In *ACM Symp. on the Theory of Computing (STOC)*, pages 612–621, 2003.

[BV10] Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. *SIAM J. on Computing*, 39(6):2464–2486, 2010.

[CFL83] Ashok K. Chandra, Merrick L. Furst, and Richard J. Lipton. Multi-party protocols. In *15th ACM Symp. on the Theory of Computing (STOC)*, pages 94–99, 1983.

[CGL$^+$20] Eshan Chattopadhyay, Jason Gaitonde, Chin Ho Lee, Shachar Lovett, and Abhishek Shetty. Fractional pseudorandom generators from any fourier level. *CoRR*, abs/2008.01316, 2020.

[CGT96] Jin-Yi Cai, Frederic Green, and Thomas Thierauf. On the correlation of symmetric functions. *Mathematical Systems Theory*, 29(3):245–258, 1996.

[Cha07] Arkadev Chattopadhyay. Discrepancy and the power of bottom fan-in in depth-three circuits. In *48th IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 449–458. IEEE, 2007.

[CHH$^+$20] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, Shachar Lovett, and David Zuckerman. XOR lemmas for resilient functions against polynomials. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, *ACM Symp. on the Theory of Computing (STOC)*, pages 234–246. ACM, 2020.

[CHHL18] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, and Shachar Lovett. Pseudorandom generators from polarizing random walks. In *CCC*, volume 102 of *LIPIcs*, pages 1:1–1:21. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018.

[CHLT19] Eshan Chattopadhyay, Pooya Hatami, Shachar Lovett, and Avishay Tal. Pseudorandom generators from the second fourier level and applications to AC0 with parity gates. In *ACM Innovations in Theoretical Computer Science conf. (ITCS)*, pages 22:1–22:15, 2019.

[CK02] Peter Clote and Evangelos Kranakis. *Boolean Functions and Computation Models*. Springer, 2002.

[CP16] Shiteng Chen and Periklis A. Papakonstantinou. Correlation lower bounds from correlation upper bounds. *Inf. Process. Lett.*, 116(8):537–540, 2016.

[CT93] Fan R. K. Chung and Prasad Tetali. Communication complexity and quasi randomness. *SIAM Journal on Discrete Mathematics*, 6(1):110–123, 1993.

[CT06] Thomas Cover and Joy Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

[DV22] Harm Derksen and Emanuele Viola. Fooling polynomials by preserving indecomposability. In *IEEE Symp. on Foundations of Computer Science (FOCS)*, 2022.

[Fre95] Yoav Freund. Boosting a weak learning algorithm by majority. *Information and Computation*, 121(2):256–285, 1995.

[GHR92] Mikael Goldmann, Johan Håstad, and Alexander A. Razborov. Majority gates vs. general weighted threshold gates. *Computational Complexity*, 2:277–300, 1992.

[GKV17] Frederic Green, Daniel Kreymer, and Emanuele Viola. Block-symmetric polynomials correlate with parity better than symmetric. *Computational Complexity*, 26(2):323–364, 2017. Available at http://www.ccs.neu.edu/home/viola/.

[GNW95] Oded Goldreich, Noam Nisan, and Avi Wigderson. On Yao's XOR lemma. Technical Report TR95–050, *Electronic Colloquium on Computational Complexity*, March 1995. www.eccc.uni-trier.de/.

[Gow98] Timothy Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geometric and Functional Analysis*, 8(3):529–551, 1998.

[Gow01] Timothy Gowers. A new proof of Szemerédi's theorem. *Geometric and Functional Analysis*, 11(3):465–588, 2001.

[GR07] Andrew Granville and Zeév Rudnick. Uniform distribution. In *Equidistribution in Number Theory, An Introduction*, volume 237 of *NATO Science Series II: Mathematics, Physics and Chemistry*, pages 1–13. Springer, 2007.

[GR10] Frederic Green and Amitabha Roy. Uniqueness of optimal mod 3 circuits for parity. *Journal of Number Theory*, 130:961 – 975, 2010.

[Gre04] Frederic Green. The correlation between parity and quadratic polynomials mod 3. *J. of Computer and System Sciences*, 69(1):28–44, 2004.

[Gro95] Vince Grolmusz. Separating the communication complexities of MOD m and MOD p circuits. *J. of Computer and System Sciences*, 51(2):307–313, 1995.

[GRS05] Frederic Green, Amitabha Roy, and Howard Straubing. Bounds on an exponential sum arising in Boolean circuit complexity. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 341(5):279–282, 2005.

[GSV18] Aryeh Grinberg, Ronen Shaltiel, and Emanuele Viola. Indistinguishability by adaptive procedures with advice, and lower bounds on hardness amplification proofs. In *IEEE Symp. on Foundations of Computer Science (FOCS)*, 2018. Available at http://www.ccs.neu.edu/home/viola/.

[GT06] Anna Gál and Vladimir Trifonov. On the correlation between parity and modular polynomials. In *31st Symposium on Mathematical Foundations of Computer Science (MFCS)*, volume 4162 of *Lecture Notes in Computer Science*, pages 387–398. Springer, 2006.

[GV04] Dan Gutfreund and Emanuele Viola. Fooling parity tests with parity gates. In *8thWorkshop on Randomization and Computation (RANDOM)*, pages 381–392. Springer, 2004.

[Han06] Kristoffer Arnsfelt Hansen. Lower bounds for circuits with few modular gates using exponential sums. *Electronic Colloquium on Computational Complexity*, Technical Report TR06-079, 2006. www.eccc.uni-trier.de/.

[Hea08] Alexander Healy. Randomness-efficient sampling within NC^1. *Computational Complexity*, 17(1):3–37, 2008.

[HG91] Johan Håstad and Mikael Goldmann. On the power of small-depth threshold circuits. *Computational Complexity*, 1(2):113–129, 1991.

[HMP$^+$93] András Hajnal, Wolfgang Maass, Pavel Pudlák, Márió Szegedy, and György Turán. Threshold circuits of bounded depth. *J. of Computer and System Sciences*, 46(2):129–154, 1993.

[HV06] Alexander Healy and Emanuele Viola. Constant-depth circuits for arithmetic in finite fields of characteristic two. In *23rd Symp. on Theoretical Aspects of Computer Science (STACS)*, pages 672–683. Springer, 2006.

[IPV22] Peter Ivanov, Liam Pavlovic, and Emanuele Viola. On correlation bounds against polynomials. 2022.

[KN97] Eyal Kushilevitz and Noam Nisan. *Communication complexity.* Cambridge University Press, 1997.

[Kop11] Swastik Kopparty. On the complexity of powering in finite fields. In *ACM Symp. on the Theory of Computing (STOC)*, 2011.

[Lok09] Satyanarayana V. Lokam. Complexity lower bounds using linear algebra. *Foundations and Trends in Theoretical Computer Science*, 4(1-2):1–155, 2009.

[Lov08] Shachar Lovett. Unconditional pseudorandom generators for low degree polynomials. In *40th ACM Symp. on the Theory of Computing (STOC)*, pages 557–562, 2008.

[LS11] Shachar Lovett and Srikanth Srinivasan. Correlation bounds for poly-size $\mbox{\rm AC}ˆ0$ circuits with n 1 - o(1) symmetric gates. In Leslie Ann Goldberg, Klaus Jansen, R. Ravi, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 14th International Workshop, APPROX 2011, and 15th International Workshop, RANDOM 2011, Princeton, NJ, USA, August 17-19, 2011. Proceedings*, volume 6845 of *Lecture Notes in Computer Science*, pages 640–651. Springer, 2011.

[LVW93] Michael Luby, Boban Veličković, and Avi Wigderson. Deterministic approximate counting of depth-2 circuits. In *2nd Israeli Symposium on Theoretical Computer Science (ISTCS)*, pages 18–24, 1993.

[MNV16] Raghu Meka, Oanh Nguyen, and Van Vu. Anti-concentration for polynomials of independent random variables. *Theory Comput.*, 12(1):1–17, 2016.

[Nis91] Noam Nisan. Pseudorandom bits for constant depth circuits. *Combinatorica. An Journal on Combinatorics and the Theory of Computing*, 11(1):63–70, 1991.

[NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: efficient constructions and applications. *SIAM J. on Computing*, 22(4):838–856, 1993.

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *J. of Computer and System Sciences*, 49(2):149–167, 1994.

[Ram20] C. Ramya. Recent progress on matrix rigidity - A survey. *CoRR*, abs/2009.09460, 2020.

[Raz87] Alexander Razborov. Lower bounds on the dimension of schemes of bounded depth in a complete basis containing the logical addition function. *Akademiya Nauk SSSR. Matematicheskie Zametki*, 41(4):598–607, 1987. English translation in Mathematical Notes of the Academy of Sci. of the USSR, 41(4):333-338, 1987.

[Raz00] Ran Raz. The BNS-Chung criterion for multi-party communication complexity. *Computational Complexity*, 9(2):113–122, 2000.

[RR97] Alexander Razborov and Steven Rudich. Natural proofs. *J. of Computer and System Sciences*, 55(1):24–35, August 1997.

[RV13] Alexander Razborov and Emanuele Viola. Real advantage. *ACM Trans. Computation Theory*, 5(4):17, 2013.

[Smo87] Roman Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *19th ACM Symp. on the Theory of Computing (STOC)*, pages 77–82. ACM, 1987.

[Smo93] Roman Smolensky. On representations by low-degree polynomials. In *34th IEEE IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 130–138, 1993.

[Sri13] Srikanth Srinivasan. On improved degree lower bounds for polynomial approximation. In *FSTTCS*, volume 24 of *LIPIcs*, pages 201–212. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2013.

[ST18] Rocco A. Servedio and Li-Yang Tan. Luby-Velickovic-Wigderson revisited: Improved correlation bounds and pseudorandom generators for depth-two circuits. In *Workshop on Randomization and Computation (RANDOM)*, volume 116 of *LIPIcs*, pages 56:1–56:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

[SV10] Ronen Shaltiel and Emanuele Viola. Hardness amplification proofs require majority. *SIAM J. on Computing*, 39(7):3122–3154, 2010.

[SV12] Rocco A. Servedio and Emanuele Viola. On a special case of rigidity. Available at http://www.ccs.neu.edu/home/viola/, 2012.

[Val77] Leslie G. Valiant. Graph-theoretic arguments in low-level complexity. In *6th Symposium on Mathematical Foundations of Computer Science*, volume 53 of *Lecture Notes in Computer Science*, pages 162–176. Springer, 1977.

[Vio] Emanuele Viola. New lower bounds for probabilistic degree and AC0 with parity gates. *Theory of Computing*. Available at http://www.ccs.neu.edu/home/viola/.

[Vio06a] Emanuele Viola. The complexity of hardness amplification and derandomization. *Ph.D. thesis, Harvard University*, 2006.

[Vio06b] Emanuele Viola. *The Complexity of Hardness Amplification and Derandomization.* PhD thesis, Harvard University, 2006.

[Vio06c] Emanuele Viola. New correlation bounds for GF(2) polynomials using Gowers uniformity. *Electronic Colloquium on Computational Complexity*, Technical Report TR06-097, 2006. www.eccc.uni-trier.de/.

[Vio07] Emanuele Viola. Pseudorandom bits for constant-depth circuits with few arbitrary symmetric gates. *SIAM J. on Computing*, 36(5):1387–1403, 2007.

[Vio09a] Emanuele Viola. Correlation bounds for polynomials over $\{0,1\}$. *SIGACT News, Complexity Theory Column*, 40(1), 2009.

[Vio09b] Emanuele Viola. On the power of small-depth computation. *Foundations and Trends in Theoretical Computer Science*, 5(1):1–72, 2009.

[Vio09c] Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d. *Computational Complexity*, 18(2):209–217, 2009.

[Vio17] Emanuele Viola. Challenges in computational lower bounds. *SIGACT News, Open Problems Column*, 48(1), 2017.

[Vio19] Emanuele Viola. Constant-error pseudorandomness proofs from hardness require majority. *ACM Trans. Computation Theory*, 11(4):19:1–19:11, 2019. Available at http://www.ccs.neu.edu/home/viola/.

[Vio21] Emanuele Viola. Fourier conjectures, correlation bounds, and majority. In *Coll. on Automata, Languages and Programming (ICALP)*, 2021. Available at http://www.ccs.neu.edu/home/viola/.

[VW08] Emanuele Viola and Avi Wigderson. Norms, XOR lemmas, and lower bounds for polynomials and protocols. *Theory of Computing*, 4:137–168, 2008.

[Yao77] Andrew Chi-Chih Yao. Probabilistic computations: Toward a unified measure of complexity (extended abstract). In *IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 222–227. IEEE Computer Society, 1977.