COMPLEXITY OF SOME ARITHMETIC PROBLEMS FOR BINARY POLYNOMIALS

Eric Allender, Anna Bernasconi, Carsten Damm, Joachim von zur Gathen, Michael Saks, and Igor Shparlinski

Abstract. We study various combinatorial complexity measures of Boolean functions related to some natural arithmetic problems about binary polynomials, that is, polynomials over $\mathbb{F}_2$. In particular, we consider the Boolean function deciding whether a given polynomial over $\mathbb{F}_2$ is squarefree. We obtain an exponential lower bound on the size of a decision tree for this function, and derive an asymptotic formula, having a linear main term, for its average sensitivity. This allows us to estimate other complexity characteristics such as the formula size, the average decision tree depth and the degrees of exact and approximative polynomial representations of this function. Finally, using a different method, we show that testing squarefreeness or irreducibility of polynomials over $\mathbb{F}_2$ are not in $\mathrm{AC}^0[p]$ for any odd prime p. Similar results are obtained for deciding coprimality of two polynomials over $\mathbb{F}_2$ as well.

1. Introduction

In light of their many applications in modern cryptography, Boolean functions related to number-theoretic problems are a natural object to study from the complexity viewpoint. Recently, lower bounds for several such functions have been obtained, for computational models such as unbounded fan-in Boolean circuits, decision trees, and real polynomials (see Allender *et al.* (2001), Bernasconi *et al.* (1999, 2000, 2001), Bernasconi & Shparlinski (1999), Coppersmith & Shparlinski (1998), Plaku & Shparlinski (2001), and Shparlinski (1999a)). The two main ingredients of these papers are harmonic analysis and estimates based on number-theoretic considerations.

In this paper we extend some results of the aforementioned papers to problems concerning arithmetic properties of polynomials over $\mathbb{F}_2$. Our primary motivation is to extend the class of natural number-theoretic problems for which lower

bounds can be rigorously proved. As one might expect, some of the techniques that have proved useful in establishing lower bounds for number-theoretic problems over the integers are also helpful for polynomials over $\mathbb{F}_2$, and indeed our techniques are similar to those of Bernasconi *et al.* (1999), (2000), (2001), and Bernasconi & Shparlinski (1999). Nevertheless, some new difficulties and effects arise when working over $\mathbb{F}_2[x]$. For example, some of our results are more precise than those known for analogous problems over the integers. On the other hand, we have not been able to extend some of the results of Allender *et al.* (2001) to the case of polynomials.

There is a well-known analogy between integers and polynomials, in particular when we take the binary representation and polynomials in $\mathbb{F}_2[x]$, respectively. Basic arithmetic can be done with analogous algorithms, for example multiplication, division with remainder, or computing the gcd. The recent result of Agrawal *et al.* (2002) also puts primality and irreducibility testing, and finding primes or irreducibles, on roughly equal footing.

However, some problems seem more difficult for integers than for polynomials, at the current state of knowledge. The most dramatic example is factorization; squarefreeness behaves similarly. For parallel computation, the gcd is an example. Bach & Shallit (1996), Cohen (1997), von zur Gathen & Gerhard (1999), and Shparlinski (1999b) present overviews on arithmetic. The lower bounds on complexity we obtain are approximately of the same strength in both cases; this is presumably just a further indication that currently available methods do not reach the computational "heart" of the difficult problems like factoring integers.

In this paper we consider Boolean functions defined by of the following properties of polynomials in $\mathbb{F}_2[x]$: A polynomial u in $\mathbb{F}_2[x]$ of degree greater than 0 is *irreducible* if $u = vw$ implies $v = 1$ or $w = 1$, and it is *squarefree* if $u = v^2 w$ implies $v = 1$ (in particular, the constant polynomial $u = 1$ is squarefree). Two polynomials are *coprime* if there is no nonconstant polynomial dividing both, that is, if their gcd is 1.

Throughout the paper we identify polynomials of degree k over $\mathbb{F}_2$ with constant coefficient 1 and the corresponding k-bit vectors of coefficients. Writing

$$u = u_n x^n + \cdots + u_1 x + 1, v = v_\ell x^\ell + \cdots + v_1 x + 1, w = w_\ell x^\ell + \cdots + w_1 x + 1,$$

we consider the following functions:

- the *irreducibility function* $f\colon \{0,1\}^n \to \{0,1\}$ defined by

$$f(u_1,\dots,u_n) = \begin{cases} 1, & \text{if } u \text{ is irreducible,} \\ 0, & \text{otherwise.} \end{cases}$$

- the *squarefreeness function* $g\colon \{0,1\}^n \to \{0,1\}$ defined by

$$g(u_1,\dots,u_n) = \begin{cases} 1, & \text{if } u \text{ is squarefree,} \\ 0, & \text{otherwise.} \end{cases}$$

- the *coprimality function* $h\colon \{0,1\}^\ell \times \{0,1\}^\ell \to \{0,1\}$ defined by

$$h(v_1,\dots,v_\ell;w_1,\dots,w_\ell) = \begin{cases} 1, & \text{if } v \text{ and } w \text{ are coprime,} \\ 0, & \text{otherwise.} \end{cases}$$

We provide estimates for the *average sensitivity* and the size of the *Fourier coefficient* of highest order for these functions. These measures are important indicators for the computational complexity of functions, and therefore they have often received study, see Bernasconi *et al.* (1996), Boppana (1997), Linial *et al.* (1993), and Nisan & Szegedy (1994). Then, using our estimates, we derive lower bounds on the decision tree size, on the average decision tree depth, on the formula size and on the degree of certain polynomial representations for g and h.

Although, as we mentioned, our techniques are similar to those used for the analogues of the functions above over the integers, here the functions exhibit a somewhat different behavior which has allowed us to obtain more precise results. For example, our results which are based on the properties of the highest order Fourier coefficient have no analogs for the integers.

Finally, we show that the technique of Allender *et al.* (2001) can be modified to provide circuit lower bounds for problems about polynomials over $\mathbb{F}_2$.

2. Basic definitions

Throughout the paper $\log z$ denotes the binary logarithm. The implied constants in symbols 'O' are absolute and can be explicitly evaluated.

Let $\mathbb{B} = \{0,1\} \subset \mathbb{R}$, so that $\mathbb{B}^n$ is the n-dimensional Boolean cube. The *Hamming weight* $|u|$ of $u \in \mathbb{B}^n$ is the number of 1's in u. If $\varphi\colon \mathbb{B}^n \longrightarrow \mathbb{B}$ is a

Boolean function and $w \in \mathbb{B}^n$, then we define

$$\hat{\varphi}(w) = \frac{1}{2^n} \sum_{u \in \mathbb{B}^n} (-1)^{\varphi(u) + \sum_i u_i w_i}.$$

The quantities $(\hat{\varphi}(w))_{w \in \mathbb{B}^n}$ are the **Fourier coefficients** of φ. We define $c(\varphi) = \hat{\varphi}(1^n)$ as the **highest order Fourier coefficient**, and $E(\varphi) = \hat{\varphi}(0^n)$ as the **lowest order Fourier coefficient**. One can easily verify that

$$\begin{aligned} c(\varphi) &= \frac{1}{2^n} \sum_{u \in \mathbb{B}^n} (-1)^{\varphi(u)+|u|} = \frac{1}{2^n} \Big(\sum_{\substack{u \in \mathbb{B}^n \\ \varphi(u)=0}} (-1)^{|u|} - \sum_{\substack{u \in \mathbb{B}^n \\ \varphi(u)=1}} (-1)^{|u|} \Big) \\ &= \frac{1}{2^n} \Big(- \sum_{\substack{u \in \mathbb{B}^n \\ \varphi(u)=1}} (-1)^{|u|} - \sum_{\substack{u \in \mathbb{B}^n \\ \varphi(u)=1}} (-1)^{|u|} \Big) = \frac{-1}{2^{n-1}} \sum_{\substack{u \in \mathbb{B}^n \\ \varphi(u)=1}} (-1)^{|u|}, \end{aligned} \tag{2.1}$$

since $\sum_{u \in \mathbb{B}^n} (-1)^{|u|} = 0$. Thus $c(\varphi)$ is the number of inputs accepted by φ with odd Hamming weight minus the number of accepted inputs with even Hamming weight, divided by 2^{n-1}. In other words, $c(\varphi)$ is the correlation of φ with the parity function on the same n input bits. Also,

$$E(\varphi) = \frac{1}{2^n} \sum_{w \in \mathbb{B}^n} (-1)^{\varphi(w)}, \tag{2.2}$$

so that $E(\varphi)$ is the expectation of the function $(-1)^{\varphi(u)}$ with regard to the uniform distribution on its domain. Combining (2.1) and (2.2) gives

$$|E(\varphi)| + |c(\varphi)| \leq 1. \tag{2.3}$$

For a bit vector $u \in \mathbb{B}^n$, we denote by $u^{(i)}$ the vector obtained from u by flipping its ith coordinate. The **sensitivity of φ at input** $u \in \mathbb{B}^n$ is the number

$$\sigma_u(\varphi) = \sum_{1 \leq i \leq n} \big| \varphi(u) - \varphi(u^{(i)}) \big|$$

of inputs at Hamming distance 1 from u where φ takes a different value. The **sensitivity** of φ is defined as

$$\sigma(\varphi) = \max_{u \in \mathbb{B}^n} \sigma_u(\varphi),$$

and the **average sensitivity** of φ is

$$s(\varphi) = 2^{-n} \sum_{u \in \mathbb{B}^n} \sigma_u(\varphi).$$

Clearly, $s(\varphi) \le \sigma(\varphi) \le n$ for any φ. The sensitivity provides lower bounds for the CREW PRAM complexity of Boolean functions; see Nisan (1989), Dietzfelbinger *et al.* (1996), Fich (1990), Parberry & Yan (1991), and Wegener (1987).

The average sensitivity of a function φ equals the sum of the **influences** of all variables on φ, where the influence of u_i on φ, denoted as $I_i(\varphi)$, is the probability that flipping the ith bit of a random Boolean input will flip the output. In other words, $I_i(\varphi)$ is a measure of how influential the variable u_i is in determining the outcome of φ. Thus we have

$$I_i(\varphi) = 2^{-n} \sum_{u \in \mathbb{B}^n} \left| \varphi(u) - \varphi(u^{(i)}) \right|,$$

which immediately implies

$$s(\varphi) = \sum_{1 \le i \le n} I_i(\varphi).$$

A binary **decision tree** T is a binary tree with inner nodes labeled by Boolean variables $U_1, \ldots, U_n$ and leaves labeled by 0 or 1. Further, edges leaving the same node are labeled 0 and 1, respectively.

Every input assignment $u \in \mathbb{B}^n$ to the variables in the tree determines a computation path from the root to one of the leaves: at each visited inner node that is labeled by variable U_i the path follows the edge labeled $u_i \in \{0, 1\}$. The tree computes the function that maps every assignment to the label of the leaf reached by its computation path.

For an input assignment u, let $D_u(T)$ be length of its computation path in T. **Depth** and **average depth** of the tree are defined by

$$D(T) = \max\{D_u(T) \colon u \in \mathbb{B}^n\}, \quad \overline{D}(T) = 2^{-n} \sum_{u \in \mathbb{B}^n} D_u(T).$$

The number of leaves is called the **size** of the decision tree.

For a Boolean function φ, let $M(\varphi)$, $D(\varphi)$, and $\overline{D}(\varphi)$, respectively, denote the **minimal size**, **minimal depth** and **minimal average depth**, respectively, of the decision trees that compute φ.

Clearly, $\overline{D}(\varphi) \leq D(\varphi) \leq n$ for any φ. Moreover, minimal average depth and minimal size of any decision tree for a function φ are related as follows:

$$\overline{D}(\varphi) \geq \log M(\varphi). \tag{2.4}$$

This follows since the sequence of labels along a root-to-leaf path gives a binary encoding for that leaf and the average length of any encoding for the leaves is at least $\log M$ bits.

Further we mention the following definitions from Nisan & Szegedy (1994): for a Boolean function $\varphi : \mathbb{B}^n \to \{0,1\}$, let the **real degree** $\Delta(\varphi)$ of φ be the degree of the unique multilinear real polynomial $P \in \mathbb{R}[U_1, \ldots, U_n]$ for which $\varphi(u) = P(u)$ holds for every $u \in \mathbb{B}^n$. Multilinearity means that each variable appears with degree at most 1.

More generally, for $\varepsilon \in [0, 1/3]$ we say that a real polynomial P in n variables ε-approximates φ if

$$|\varphi(u) - P(u)| \leq \varepsilon$$

for all $u \in \mathbb{B}^n$, and define the **real ε-approximate degree** $\delta_\varepsilon(\varphi)$ of φ as the minimum degree of a multilinear real polynomial that ε-approximates φ. We simply write $\delta(\varphi)$ for $\delta_{1/3}(\varphi)$. This notion was introduced by Nisan and Szegedy with $\varepsilon = 1/3$, but it will be convenient to extend it to smaller ε. Clearly $n \geq \delta_\varepsilon(\varphi) \geq \delta(\varphi)$ for any $\varepsilon \in (0, 1/3)$. The following shows that decreasing ε can increase the real approximate degree only by a constant factor.

LEMMA 2.5. *For any Boolean function φ and for $\varepsilon \in (0, 1/3)$ we have*

$$\delta_\varepsilon(\varphi) \leq C \cdot \delta(\varphi) \log \varepsilon^{-1},$$

where C is a constant independent of ε and φ.

PROOF. Let $P \in \mathbb{R}[U_1, \ldots, U_n]$ be a polynomial that 1/3-approximates φ. Defining $Q_0 = \frac{1}{5}(3P + 1)$, we have $Q_0(u) \in [0, \frac{2}{5}]$ for $u \in \varphi^{-1}(0)$ and $Q_0(u) \in [\frac{3}{5}, 1]$ for $u \in \varphi^{-1}(1)$.

Let $w(z) = 3z^2 - 2z^3$. This is the unique cubic polynomial with vanishing constant and linear terms that satisfies $w(1-z) = 1 - w(z)$. It is routine to show that for any $z \in [0, \frac{2}{5}]$ we have $w(z) \in [0, z^{1.1})$ and $w(1-z) \in (1-z^{1.1}, 1]$.

For $i \geq 0$, define the polynomial $Q_{i+1} = w(Q_i)$. Then it follows by induction that $Q_i(u) \in [0,1]$ for all $u \in \mathbb{B}^n$, that Q_i γ_i-approximates φ with $\gamma_i = (2/5)^{1.1^i}$, and that $\deg(Q_i) = 3^i \deg(P)$. We can then choose m of size $O(\log\log \varepsilon^{-1})$, so that $\gamma_m \leq \varepsilon$, and $\deg(Q_m) \leq C \deg(P) \log \frac{1}{\varepsilon}$, for some absolute constant C. □

A Boolean **circuit** C on n variables is a directed acyclic graph with Boolean inputs $0, 1, x_1, \ldots, x_n$ and some number of output gates $y_1, \ldots, y_r$. The gates of C (except for the input gates) are labelled by $\neg$, $\wedge$, or $\vee$ and have the corresponding in-degree; their number is the **size** $s(C)$ of C. The **depth** $d(C)$ is the length of a longest path from an input to an output in C. The circuit computes a Boolean function from $\mathbb{B}^n$ to $\mathbb{B}^r$ in the natural way.

Formulae are defined in the following recursive way: 0, 1, the variables $x_1, \ldots, x_n$ and their negations $\neg x_1, \ldots, \neg x_n$ are formulae; if F_1 and F_2 are formulae, then so are $F_1 \wedge F_2$ and $F_1 \vee F_2$. The **size** of a formula is the number of occurrences of variables in it. Formulae are equivalent to Boolean circuits where the fan-out of each gate is bounded by one. Let $L(\varphi)$ denote the minimal size of formulas that compute φ.

Usually, we are interested in Boolean functions $\varphi\colon \mathbb{B}^* \to \mathbb{B}$. In order to discuss the (non-uniform) circuit complexity of such functions φ, it is necessary to consider *families* of circuits $(C_n)_{n\in\mathbb{N}}$, where C_n has n variables. Then the family $(C_n)_{n\in\mathbb{N}}$ computes φ if C_n outputs $\varphi(u)$ for all n and $u \in \mathbb{B}^n$. A circuit family has size and depth bounded by $s(n)$ and $d(n)$, respectively, if $s(C_n) \leq s(n)$ and $d(C_n) \leq d(n)$.

A function φ is in AC^0 if there is a circuit family $(C_n)_{n\in\mathbb{N}}$ of size $n^{O(1)}$ and depth $O(1)$ consisting of inputs, negated inputs, and unbounded fan-in AND and OR gates and computing φ.

For integers $d, n \geq 1$ we define the Boolean function $\mathrm{MOD}_d\colon \mathbb{B}^n \longrightarrow \mathbb{B}$ as

$$\mathrm{MOD}_d(u) = \begin{cases} 1 & \text{if } \sum_{1\leq i\leq n} u_i \equiv 0 \bmod d, \\ 0 & \text{otherwise,} \end{cases}$$

for $u \in \mathbb{B}^n$. The function MOD_2 is known as the **parity** function. It has been known since Ajtai (1983) and Furst *et al.* (1984) that the parity function is not in AC^0. This has led researchers to consider the power of AC^0 circuits that are augmented with parity gates, and more generally with MOD_d gates.

Let $d > 1$ be an integer. A function φ is in $\mathrm{AC}^0[d]$ if there is a circuit family $(C_n)_{n\in\mathbb{N}}$ of size $n^{O(1)}$ and depth $O(1)$ computing φ and consisting of inputs, negated inputs, and unbounded fan-in AND, OR, and MOD_d gates.

For circuits with MOD_p gates with prime p one can prove exponential lower size bounds for explicitly defined functions. Circuits with MOD_d gates for composite d are of interest as well but unfortunately almost no nontrivial results are known about such circuits, even in the simplest case $d = 6$.

3. Fourier coefficients, average sensitivity, and computational complexity

We mention here some known relations among various complexity measures.

The following simple and fundamental fact, which we heard from Nati Linial, seems not to have appeared before.

LEMMA 3.1. *For any Boolean function* $\varphi\colon \mathbb{B}^n \longrightarrow \mathbb{B}$, $s(\varphi) \geq |c(\varphi)| \cdot n$.

PROOF. The hypercube H_n is the undirected graph on vertex set $\mathbb{B}^n$, whose edge set E_n consists of pairs of points that differ in a single bit position. We partition $\mathbb{B}^n$ into four sets $V_i(j)$ for $i, j \in \{0,1\}$, where $V_i(j)$ is the set of $x \in \mathbb{B}^n$ such that $|x| \equiv i \bmod 2$ and $\varphi(x) = j$. Let $N_i(j) = |V_i(j)|$, so that $N_i(0) + N_i(1) = 2^{n-1}$ for $i \in \{0,1\}$. Let $V(j) = V_0(j) \cup V_1(j) = \varphi^{-1}(j)$ and let $V_i = \{x \in \mathbb{B}^n : |x| \equiv i \bmod 2\} = V_i(0) \cup V_i(1)$. We split E_n into three sets E^0, E^1, and $E^{\neq}$, where $(x,y) \in E^0$ if $\varphi(x) = \varphi(y) = 0$, $(x,y) \in E^1$ if $\varphi(x) = \varphi(y) = 1$ and $(x,y) \in E^{\neq}$ if $\varphi(x) \neq \varphi(y)$. We write $d^{\neq}(x)$ for the number of neighbors y of x with $\varphi(y) \neq \varphi(x)$. Then $|c(\varphi)| = 2^{1-n}|N_0(1) - N_1(1)|$ and $s(\varphi) = 2^{-n}\sum_{x\in\mathbb{B}^n} d^{\neq}(x) = 2^{1-n}|E^{\neq}|$, so it suffices to show that $|E^{\neq}| \geq n(|N_0(1)| - |N_1(1)|)$. Now, $|E^0| \leq n\min\{N_0(0), N_1(0)\}$ and

$$\begin{aligned} |E^1| \leq n\min\{N_0(1), N_1(1)\} &= n\min\{2^{n-1} - N_0(0), 2^{n-1} - N_1(0)\} \\ &= n(2^{n-1} - \max\{N_0(0), N_1(0)\}). \end{aligned}$$

Thus

$$\begin{aligned} |E^{\neq}| &= n2^{n-1} - |E^0| - |E^1| \\ &\geq n(\max\{N_0(0), N_1(0)\} - \min\{N_0(0), N_1(0)\}) \\ &= n|N_0(0) - N_1(0)| = n|N_0(1) - N_1(1)|. \end{aligned}$$

□

The following lower bound on the minimal decision tree size in terms of the Fourier coefficients was proved in Jukna *et al.* (1999). It combines results from Kushilevitz & Mansour (1993) and Linial *et al.* (1993).

LEMMA 3.2. *For an* n*-variate Boolean function* φ *and* $w \in \mathbb{B}^n$, *we have*

$$M(\varphi) \geq 2^{|w|} \sum_{u \geq w} |\hat{\varphi}(u)|,$$

where the sum is taken over all $u \in \mathbb{B}^n$ *such that* $u_i \geq w_i$ *for all* i.

The following well known fact says that if $c(\varphi) \neq 0$, then the decision tree depth and real degree of φ are determined.

LEMMA 3.3. *Let φ be an n-variate Boolean function. If $c(\varphi) \neq 0$, then $D(\varphi) = \Delta(\varphi) = n$.*

PROOF. First, $D(\varphi) \geq \Delta(\varphi)$ since any decision tree T for φ of depth d gives a polynomial P of degree d that equals φ: take P to be the sum over accepting paths (with leaf label 1) in the decision tree of the product of u_i if $u_i = 1$ on the path and $1 - u_i$ if $u_i = 0$ on the path.

Let P be the unique multilinear polynomial representing φ. Then $(-1)^{\varphi(u)} = 1 - 2P(u)$ for any $u \in \mathbb{B}^n$ and therefore from (2.1) we have $c(\varphi) = \sum_{u \in \mathbb{B}^n} (1 - 2P(u))(-1)^{|u|}$. Expanding $1 - 2P$ as a linear combination of multilinear monomials, we see that any monomial of degree less than n has a net contribution of 0 to the sum $\sum_{u \in \mathbb{B}^n} (1 - 2P(u))(-1)^{|u|}$. Thus if $c(\varphi) \neq 0$, then P has degree at least n, and the trivial bound $u \geq D(\varphi)$ implies the claim. □

The next lemma and its corollary relate the real approximate degree of φ to $c(\varphi)$. We recall a few basic facts about the set F_n of functions mapping $\{-1, 1\}^n$ to $\mathbb{R}$. This is a 2^n dimensional real vector space. Let $[n]$ denote the set $\{1, \ldots, n\}$. For $J \subseteq [n]$, define χ_J to be the n-variate polynomial $\chi_J(X_1, \ldots, X_n) = \prod_{j \in J} X_j$. Then the functions $\frac{1}{2^{n/2}} \chi_J$ when restricted to $\{-1, 1\}^n$ form an orthonormal basis of F_n (with the usual inner product $\langle \cdot, \cdot \rangle$). The representation of a function $f \in F_n$ in terms of this basis gives a real multilinear polynomial $\Psi_f = \sum_{J \in [n]} a_J \chi_J$ whose restriction to $\{-1, 1\}^n$ agrees with f. By the orthonormality of the basis $\frac{1}{2^{n/2}} \chi_J$, we have that the high order coefficient $a_{[n]}$ satisfies

$$a_{[n]} = \left\langle f, \frac{\chi_{[n]}}{2^{n/2}} \right\rangle = \frac{1}{2^n} \sum_{v \in \{-1,1\}^n} \Psi_f(v) \chi_{[n]}(v). \tag{3.4}$$

Also, since the coefficients $a_J 2^{n/2}$ for $J \subseteq [n]$ are obtained by applying an orthonormal change of basis from f, we have Parseval's identity:

$$\sum_{v \in \{-1,1\}^n} (f(v))^2 = 2^n \sum_{J \subseteq [n]} {a_J}^2. \tag{3.5}$$

LEMMA 3.6. *Let $\varphi \colon \mathbb{B}^n \longrightarrow \mathbb{B}$ be an n-variate Boolean function, and P an n-variate real multilinear polynomial of degree $d < n$. Then*

$$\max_{u \in \mathbb{B}^n} |\varphi(u) - P(u)| \geq |c(\varphi)|/2.$$

PROOF. Let f be the function in F_n obtained by changing the Boolean value $b \in \{0,1\}$ to $(-1)^b$ in both the range and domain of φ. Thus for $(v_1, \ldots, v_n) \in \{-1,1\}^n$, we have

$$f(v_1, \ldots, v_n) = 1 - 2\varphi\left(\frac{1-v_1}{2}, \ldots, \frac{1-v_n}{2}\right).$$

Similarly, define the real polynomial Q by

$$Q(X_1, \ldots, X_n) = 1 - 2P\left(\frac{1-X_1}{2}, \ldots, \frac{1-X_n}{2}\right).$$

Let $\Psi = \sum_{J\subseteq[n]} a_J \chi_J$ be the polynomial representation of f, let $H = \Psi - Q$, and write $H = \sum_{J\subseteq[n]} h_J \chi_J$. Then

$$\begin{aligned}
\max_{u\in\mathbb{B}^n} |\varphi(u) - P(u)| &= \frac{1}{2} \max_{v\in\{-1,1\}^n} |H(v)| \\
&\geq \frac{1}{2}\left(2^{-n} \sum_{v\in\{-1,1\}^n} (H(v))^2\right)^{1/2} \\
&= \frac{1}{2}\left(\sum_{J\subseteq[n]} (h_J)^2\right)^{1/2} \geq \frac{1}{2} h_{[n]},
\end{aligned}$$

where the last equality comes from (3.5). Now $h_{[n]} = a_{[n]}$, since $\deg(Q) = \deg(P) < n$. By (3.4), and since $f((-1)^{u_1}, \ldots, (-1)^{u_n}) = (-1)^{\varphi(u)}$ for $u \in \mathbb{B}^n$, we have

$$h_{[n]} = a_{[n]} = \frac{1}{2^n} \sum_{v\in\{-1,1\}^n} f(v)\chi_{[n]}(v) = \frac{1}{2^n} \sum_{u\in\mathbb{B}^n} (-1)^{\varphi(u)+|u|} = c(\varphi). \qquad \square$$

Lemmas 2.5 and 3.6 imply the following.

COROLLARY 3.7. *For each $\gamma > 0$ there is an (effectively computable) constant $K(\gamma) > 0$ such that for any n-variate Boolean function φ, $c(\varphi) > \gamma$ implies $\delta(\varphi) \geq K(\gamma)n$.*

The following bound on the formula size in terms of average sensitivity was derived in Bernasconi *et al.* (1999, 2000).

LEMMA 3.8. *Let φ be a Boolean function depending on n variables. Then*

$$L(\varphi) \geq \frac{s(\varphi)^2}{1 - E(\varphi)^2}.$$

This bound has essentially been mentioned also in Bernasconi *et al.* (1996) and Boppana (1997). Finally we have the following slightly relaxed version of the result of Smolensky (1987).

LEMMA 3.9. *Let p be a prime, and let $d \geq 2$ not be a power of p. Then the Boolean function* MOD_d *is not in* $AC^0[p]$.

4. Squarefree and coprime polynomials

4.1. A preliminary identity. We denote by $\mathcal{I}$ the set of all irreducible polynomials $w \in \mathbb{F}_2[x]$, and let

$$\mathcal{I}_0 = \mathcal{I} \backslash \{x\} \qquad \text{and} \qquad \mathcal{I}_1 = \mathcal{I} \backslash \{x+1\}.$$

LEMMA 4.1. *We have*

$$\prod_{w \in \mathcal{I}} \left(1 - 2^{-2\deg w}\right) = \frac{1}{2},$$
$$\prod_{w \in \mathcal{I}_0} \left(1 - 2^{-2\deg w}\right) = \prod_{w \in \mathcal{I}_1} \left(1 - 2^{-2\deg w}\right) = \frac{2}{3}.$$

PROOF. By picking $z = 1/4$, the first equality follows from the identity

$$\prod_{w \in \mathcal{I}} \left(1 - z^{\deg w}\right)^{-1} = \frac{1}{1 - 2z}$$

which is a special case (with $q = 2$) of Theorem 3.32 of Berlekamp (1968). The other two products equal the first one times $4/3$. □

We will make frequent use of an equivalent formulation of the products in the preceding lemma. For polynomials over $\mathbb{F}_2$ we have an analog of the Euler product formula (which is better known over the integers)

$$\prod_{w \in \mathcal{I}} \left(1 - 2^{-2\deg w}\right) = \sum_{w} \mu(w) 2^{-2\deg w}, \tag{4.2}$$

where the sum is taken over all nonzero polynomials w in $\mathbb{F}_2[x]$, and $\mu(w)$ is the Möbius function for these polynomials. Recall that $\mu(1) = 1$ and if $\deg w \geq 1$, then $\mu(w) = 0$ if w is not squarefree and $\mu(w) = (-1)^{\nu(w)}$ otherwise, where $\nu(w)$ is the number of distinct irreducible divisors of $w \in \mathbb{F}_2[x]$.

4.2. Estimating the highest order Fourier coefficient. The next two lemmas yield estimates for the highest order Fourier coefficient of the squarefreeness function g and of the coprimality function h. As our standard notation, we use the set $\mathcal{M}_n = \{u \in \mathbb{F}_2[x] \colon \deg u \leq n,\ u \equiv 1 \bmod x\}$ of polynomials of the form

$$u = u_n x^n + \ldots + u_1 x + 1 \in \mathbb{F}_2[x].$$

The congruence

$$u \equiv 1 \bmod x \tag{4.3}$$

is equivalent to $u(0) = 1$. We can identify a polynomial $u \in \mathcal{M}_n$ with the bit string $u = (u_1, \ldots, u_n) \in \mathbb{B}^n$. Then the string has odd Hamming weight if and only if $u(1) = 0$.

LEMMA 4.4. *For the squarefreeness function g we have*

$$c(g) = -\frac{1}{3} + o(1).$$

PROOF. Let D_n denote the number of squarefree polynomials $u \in \mathcal{M}_n$ with $u(1) = 0$ minus the number of squarefree polynomials $u \in \mathcal{M}_n$ with $u(1) = 1$. Then

$$c(g) = \frac{D_n}{2^{n-1}}.$$

For a nonzero polynomial $m \in \mathbb{F}_2[x]$, let $\mathcal{R}_n(m)$ be the set of polynomials $u \in \mathcal{M}_n$ with $u \equiv 0 \bmod m^2$, and let $R_n(m)$ denote the cardinality of that set. We also denote by $T_n(m)$ the number of $u \in \mathcal{R}_n(m)$ with $u(1) = 0$ minus the number of $u \in \mathcal{R}_n(m)$ with $u(1) = 1$.

The inclusion-exclusion principle implies that

$$D_n = \sum_{\deg m \leq n/2} \mu(m) T_n(m).$$

The constant polynomial $u = 1$ contributes -1 both to $T_n(1)$ and to D_n. The idea of the principle here, and in our later applications, is that we start

with the main term $T_n(1) = \mathcal{M}_n$. Then for irreducible m the polynomials in $T_n(m)$ are subtracted, since $\mu(m) = -1$. Those in $T_n(m_1 m_2)$ for irreducible $m_1 \neq m_2$ have been subtracted twice, so now they are added in once again, since $\mu(m_1 m_2) = 1$. And so on.

Now if $m(1) = 0$, then every polynomial $u \in \mathcal{R}_n(m)$ is such that $u(1) = 0$; thus $T_n(m) = R_n(m)$. On the other hand, if $m(1) = 1$ and $\deg m < n/2$, then exactly half of the polynomials $u \in \mathcal{R}_n(m)$ satisfy $u(1) = 1$; thus $T_n(m) = 0$. Therefore we have

$$D_n = \sum_{\substack{m(1)=0 \\ \deg m \le n/2}} \mu(m) R_n(m) + \sum_{\substack{m(1)=1 \\ \deg m = n/2}} \mu(m) T_n(m).$$

Since $|T_n(m)| \le R_n(m) = 2^{n-2\deg m}$ for $\deg m \le n/2$, the second sum is of order $O(2^{n/2})$. Since $R_n(m) = 2^{n-2\deg m}$, for the first sum we have

$$\begin{aligned} \sum_{\substack{m(1)=0 \\ \deg m \le n/2}} \mu(m) R_n(m) &= \sum_{\substack{m(1)=0 \\ \deg m \le n/2}} \mu(m) 2^{n-2\deg m} \\ &= \sum_{m(1)=0} \mu(m) 2^{n-2\deg m} + O\Bigg(\sum_{\substack{m(1)=0 \\ \deg m > n/2}} 2^{n-2\deg m} \Bigg). \end{aligned}$$

For the error term we obtain

$$\sum_{\substack{m(1)=0 \\ \deg m > n/2}} 2^{n-2\deg m} = \sum_{d > n/2} \sum_{\substack{m(1)=0 \\ \deg m = d}} 2^{n-2d} \le \sum_{d > n/2} 2^{n-d} = O\left(2^{n/2}\right).$$

Thus we have

$$D_n = 2^n \sum_{m(1)=0} \mu(m) 2^{-2\deg m} + O\left(2^{n/2}\right).$$

Taking into account that the only irreducible polynomial w with $w(1) = 0$ is $w = x + 1$, using (4.2) and Lemma 4.1 we obtain

$$\begin{aligned} \sum_{m(1)=0} \mu(m) 2^{-2\deg m} &= \sum_{m} \mu(m) 2^{-2\deg m} - \sum_{m(1)=1} \mu(m) 2^{-2\deg m} \\ &= \prod_{w \in \mathcal{I}} \left(1 - 2^{-2\deg w}\right) - \prod_{w \in \mathcal{I}_1} \left(1 - 2^{-2\deg w}\right) \\ &= \frac{1}{2} - \frac{2}{3} = \frac{-1}{6}. \end{aligned}$$

□

LEMMA 4.5. *For the coprimality function h, the highest order Fourier coefficient is*

$$c(h) = -\frac{1}{3} + o(1).$$

PROOF. Let $n = 2\ell$ be an even integer and let $\mathcal{N}_n$ denote the set of pairs $(u, v) \in \mathcal{M}_\ell^2$ of coprime polynomials, with $\gcd(u, v) = 1$. We denote by G_n the number of pairs $(u, v) \in \mathcal{N}_n$ such that $u(1) \neq v(1)$ minus the number of pairs $(u, v) \in \mathcal{N}_n$ such that $u(1) = v(1)$. That is, G_n is the number of pairs $(u, v) \in \mathcal{N}_n$ with odd $|u| + |v|$ minus those with even $|u| + |v|$. Then

$$c(h) = \frac{G_n}{2^{n-1}}.$$

For a nonzero polynomial $m \in \mathbb{F}_2[x]$, we let $\mathcal{S}_n(m)$ be the set of $(u, v) \in \mathcal{M}_\ell^2$ with $u \equiv v \equiv 0 \bmod m$, and let $S_n(m)$ denote its cardinality. We also denote by $Q_n(m)$ the number of pairs $(u, v) \in \mathcal{S}_n(m)$ with $u(1) \neq v(1)$ minus the number of $(u, v) \in \mathcal{S}_n$ with $u(1) = v(1)$.

From the inclusion-exclusion principle it follows that

$$G_n = \sum_{\deg m \leq \ell} \mu(m) Q_n(m).$$

Now if $m(1) = 0$, then every pair $(u, v) \in \mathcal{S}_n(m)$ is such that $u(1) = v(1) = 0$; thus $Q_n(m) = S_n(m)$. On the other hand, if $m(1) = 1$ and $\deg m < n/2$, then exactly half of the pairs $(u, v) \in \mathcal{S}_n(m)$ satisfy $u(1) = v(1)$; thus $Q_n(m) = 0$. When $\deg m = \ell = n/2$, then

$$S_n(m) = \begin{cases} \{(m, m)\} & \text{if } m \equiv 1 \bmod x, \\ \emptyset & \text{otherwise.} \end{cases}$$

We then have

$$G_n - \sum_{\substack{m(1)=0 \\ \deg m \leq \ell}} \mu(m) \mathcal{S}_n(m) = \sum_{\substack{\deg m = \ell \\ m(1)=1}} \mu(m) Q_n(m),$$

and this is absolutely bounded by

$$|\{m \in \mathbb{F}_2[x] \colon \deg m = \ell, m(0) = m(1) = 1\}| = 2^{\ell-1}.$$

Now the desired result follows from the calculation in Lemma 4.4. □

4.3. Influences and average sensitivity. In this section we derive asymptotically optimal formulas for $s(g)$ and $s(h)$.

Using Lemma 3.1, we could immediately derive linear lower bounds for the average sensitivities $s(g)$ and $s(h)$ from the bounds on the highest order Fourier coefficient given in Lemmas 4.4 and 4.5. However, with a different, and longer, calculation it is possible to improve these bounds and give asymptotically tight upper and lower bounds on $s(g)$ and $s(h)$.

We define the constant

$$\gamma = \frac{2}{3} - 2 \prod_{w \in \mathcal{I}} \left(1 - \frac{2}{2^{2 \deg w}}\right). \tag{4.6}$$

Numerical calculations yield $\gamma \approx 0.27358$.

THEOREM 4.7. *Let g be the squarefreeness function and $1 \le i \le n$. Then $I_i(g) = 2\gamma + o(1)$, and $s(g) = 2\gamma n + o(n)$.*

PROOF. Let M_i denote the number of polynomials $u \in \mathcal{M}_n$ which are not squarefree and for which $u^{(i)} = u + x^i$ is squarefree. Thus $I_i(g) = 2M_i/2^n$. We now show that $M_i = \gamma 2^n + O(2^{7n/8})$, which suffices to prove the theorem, since $s(g) = \sum_{1 \le i \le n} I_i(g)$.

For $m \in \mathbb{F}_2[x]$, let $W_{i,n}(m)$ be the number of polynomials $u \in \mathbb{F}_2[x]$ which are not squarefree and for which

$$u + T^i \equiv 0 \bmod m^2.$$

Let $\mathcal{S}_n$ denote the set of squarefree polynomials in $\mathcal{M}_n$. The inclusion-exclusion principle says that

$$M_i = \sum_{\substack{\deg m \le n/2 \\ m \in \mathcal{S}_n}} \mu(m)\, W_{i,n}(m).$$

The constant polynomial $u = 1$ is squarefree, and thus contributes neither to M_i nor to $W_{i,n}(m)$ for any i, m.

Given a further polynomial $k \in \mathbb{F}_2[x]$, let $R_{i,n}(k,m)$ be the number of polynomials $u \in \mathcal{M}_n$ such that

$$u \equiv 0 \bmod k^2 \qquad \text{and} \qquad u + T^i \equiv 0 \bmod m^2. \tag{4.8}$$

Again, the inclusion-exclusion principle says that

$$W_{i,n}(m) = - \sum_{\substack{0 < \deg k \le n/2 \\ k \in \mathcal{S}_n}} \mu(k)\, R_{i,n}(k,m).$$

Since u and $u + x^i$ are coprime, $R_{i,n}(k,m) = 0$ unless k and m are coprime. If x divides either k or m, then (4.3) is inconsistent with (4.8) since $i \geq 1$, and hence $R_{i,n}(k,m) = 0$. Otherwise, x, k, and m are pairwise coprime, and by the Chinese Remainder Theorem, u is uniquely determined modulo xk^2m^2, so that $R_{i,n}(k,m) = 2^{n-2\deg km}$ if $n \geq 2\deg km$, and $R_{i,n}(k,m) = 0$ otherwise. Together, we obtain

$$R_{i,n}(k,m) = 2^{n-2\deg km} + O\left(2^{n-t}\right) = 2^{n-2\deg km} + O\left(2^{n/2}\right) \tag{4.9}$$

for any polynomial $k \in \mathcal{S}_n$. It is also clear that

$$R_{i,n}(k,m) \leq 2^{n-2\deg k} \tag{4.10}$$

for any $k \in \mathcal{S}_n$ of degree at most $n/2$.

In our estimates below we will use several times that a sum of the form

$$S(D) = \sum_{\substack{D < \deg k \leq n/2 \\ k \in \mathcal{S}_n}} 2^{n-2\deg k}$$

can be bounded as

$$S(D) = \sum_{n/2 \geq d > D} \sum_{\substack{\deg k = d \\ k \in \mathcal{S}_n}} 2^{n-2d} \leq \sum_{d > D} 2^{n-d} \leq 2^{n-D}.$$

Now fix some integer $K \geq 1$. Using (4.9) for $\deg k \leq K$ and (4.10) for $\deg k > K$, we obtain

$$\begin{aligned} W_{i,n}(m) &= -\sum_{\substack{0 < \deg k \leq n/2 \\ \gcd(k,m)=1 \\ k \in \mathcal{S}_n}} \mu(k)\, R_{i,n}(k,m) \\ &= -\sum_{\substack{0 < \deg k \leq K \\ \gcd(k,m)=1 \\ k \in \mathcal{S}_n}} \mu(k) 2^{n-2\deg m - 2\deg k} + O\left(\sum_{\substack{0 < \deg k \leq K \\ k \in \mathcal{S}_n}} 2^{n/2}\right) + O\left(S(K)\right) \\ &= -\sum_{\substack{0 < \deg k \leq K \\ \gcd(k,m)=1 \\ k \in \mathcal{S}_n}} \mu(k) 2^{n-2\deg m - 2\deg k} + O\left(2^{K+n/2} + 2^{n-K}\right). \end{aligned}$$

For the first summand, we have

$$\begin{aligned}
&\sum_{\substack{0<\deg k\le K\\ \gcd(k,m)=1\\ k\in\mathcal{S}_n}} \mu(k)2^{n-2\deg m-2\deg k}\\
&= \sum_{\substack{\deg k>0\\ \gcd(k,m)=1\\ k\in\mathcal{S}_n}} \mu(k)2^{n-2\deg m-2\deg k} + O\left(\sum_{\substack{\deg k> K\\ k\in\mathcal{S}_n}} 2^{n-2\deg k-2\deg m}\right)\\
&= 2^{n-2\deg m}\left(-1+\prod_{\substack{\gcd(w,m)=1\\ w\in\mathcal{I}_0}}(1-2^{-2\deg w})\right)+O\left(2^{n-2\deg m-K}\right).
\end{aligned}$$

Selecting $K=\lceil n/4\rceil$, we obtain

$$W_{i,n}(m)=2^{n-2\deg m}\left(1-\prod_{\substack{\gcd(w,m)=1\\ w\in\mathcal{I}_0}}(1-2^{-2\deg w})\right)+O\left(2^{3n/4}\right). \tag{4.11}$$

It is also clear from the definition that

$$W_{i,n}(m)\le 2^{n-2\deg m} \tag{4.12}$$

for any $m\in\mathcal{S}_n$. We may use (4.11) for $\deg m\le n/8$, since then $(n-2\deg m)/2\ge K$, and (4.12) for $\deg m>n/8$, and obtain

$$\begin{aligned}
M_i &= \sum_{\substack{\deg m\le n/8\\ m\in\mathcal{S}_n}} \mu(m)\,2^{n-2\deg m}\left(1-\prod_{\substack{\gcd(w,m)=1\\ w\in\mathcal{I}_0}}(1-2^{-2\deg w})\right)\\
&\qquad +O\left(2^{3n/4}\sum_{\substack{\deg m\le n/8\\ m\in\mathcal{S}_n}}\mu(m)+S(n/8)\right).
\end{aligned}$$

As before we obtain for the error term

$$\begin{aligned}
&|2^{3n/4}\sum_{\substack{\deg m\le n/8\\ m\in\mathcal{S}_n}}\mu(m)+S(n/8)|\\
&\qquad\le 2^{3n/4}|\{m\in\mathcal{S}_n\colon\ \deg m\le n/8\}|+n^{7n/8}=O\left(2^{7n/8}\right).
\end{aligned}$$

Therefore

$$M_i=\sum_{\substack{\deg m\le n/8\\ m\in\mathcal{S}_n}}\mu(m)\,2^{n-2\deg m}\left(1-\prod_{\substack{\gcd(w,m)=1\\ w\in\mathcal{I}_0}}(1-2^{-2\deg w})\right)+O\left(2^{7n/8}\right). \tag{4.13}$$

Extending the summation range in (4.13) to all polynomials $m \in \mathcal{S}_n$ gives an additional error term not exceeding $S(n/8) \leq 2^{7n/8}$. Thus we have

$$M_i = \sum_{m \in \mathcal{S}_n} \mu(m) 2^{n-2\deg m} - \sum_{m \in \mathcal{S}_n} \mu(m) 2^{n-2\deg m} \prod_{\substack{\gcd(w,m)=1 \\ w \in \mathcal{I}_0}} (1 - 2^{-2\deg w}) + O\left(2^{7n/8}\right).$$

The first sum equals $\frac{2}{3} \cdot 2^n$ by Lemma 4.1 and

$$\sum_{m \in \mathcal{S}_n} \mu(m) 2^{-2\deg m} = \prod_{w \in \mathcal{I}_0} \left(1 - 2^{-2\deg w}\right).$$

Finally, we calculate the second sum as follows, using the squarefreeness of the m in the sum for the third equation.

$$\begin{aligned}
&\sum_{m \in \mathcal{S}_n} \mu(m) 2^{-2\deg m} \prod_{\substack{\gcd(w,m)=1 \\ w \in \mathcal{I}_0}} (1 - 2^{-2\deg w}) \\
&\quad = \prod_{w \in \mathcal{I}_0} \left(1 - 2^{-2\deg w}\right) \sum_{m \in \mathcal{S}_n} \mu(m) 2^{-2\deg m} \prod_{\substack{w \in \mathcal{I}_0 \\ w \mid m}} \left(1 - 2^{-2\deg w}\right)^{-1} \\
&\quad = \prod_{w \in \mathcal{I}_0} \left(1 - 2^{-2\deg w}\right) \sum_{m \in \mathcal{S}_n} \mu(m) 2^{-2\deg m} \prod_{\substack{w \in \mathcal{I}_0 \\ w \mid m}} \left(\frac{2^{2\deg w}}{2^{2\deg w} - 1}\right) \\
&\quad = \prod_{w \in \mathcal{I}_0} \left(1 - 2^{-2\deg w}\right) \sum_{m \in \mathcal{S}_n} \mu(m) \prod_{\substack{w \in \mathcal{I}_0 \\ w \mid m}} \left(\frac{1}{2^{2\deg w} - 1}\right) \\
&\quad = \prod_{w \in \mathcal{I}_0} \left(1 - 2^{-2\deg w}\right) \prod_{w \in \mathcal{I}_0} \left(1 - \frac{1}{2^{2\deg w} - 1}\right) \\
&\quad = \prod_{w \in \mathcal{I}_0} \left(1 - \frac{2}{2^{2\deg w}}\right) = 2 \prod_{w \in \mathcal{I}} \left(1 - \frac{2}{2^{2\deg w}}\right).
\end{aligned}$$

Adding up, we find $M_i = \gamma 2^n + O(2^{7n/8})$ and $s(g) = \sum_{1 \leq i \leq n} I_i(g) = 2\gamma n + O(n 2^{-n/8})$, as desired. □

THEOREM 4.14. *Let h be the coprimality function and $1 \leq i \leq n = 2\ell$. Then $I_i(h) = 2\gamma + o(1)$ and $s(h) = 2\gamma n + o(n)$.*

PROOF. We denote by L_i the number of pairs $(u, v) \in \mathcal{M}_\ell^2$ of polynomials such that u and v are not coprime and

- if $1 \le i \le \ell$, then $u + x^i$ and v are coprime,
- if $\ell < i \le 2\ell$, then u and $v + x^{i-\ell}$ are coprime.

Clearly, $I_i(h) = 2L_i/2^n$. We now show that $L_i = \gamma 2^n + O(2^{7n/8})$ which suffices to prove the theorem.

For $m \in \mathbb{F}_2[x]$, let $U_{i,n}(m)$ be the number of pairs of polynomials $(u, v) \in \mathcal{M}_\ell^2$ which are not coprime and

$$\begin{array}{l} \bullet \text{ if } 1 \le i \le \ell, \text{ then } u + x^i \equiv v \equiv 0 \bmod m, \\ \bullet \text{ if } \ell < i \le 2\ell, \text{ then } u \equiv v + x^{i-\ell} \equiv 0 \bmod m. \end{array} \tag{4.15}$$

As before, let $\mathcal{S}_n$ denote the set of squarefree polynomials in $\mathcal{M}_n$. The inclusion-exclusion principle yields

$$L_i = \sum_{\substack{\deg m \le \ell \\ m \in \mathcal{S}_n}} \mu(m)\, U_{i,n}(m).$$

Given a further nonconstant $k \in \mathbb{F}_2[x]$, let $V_{i,n}(k, m)$ be the number of pairs of polynomials $(u, v) \in U_{i,n}(m)$ satisfying

$$u \equiv v \equiv 0 \bmod k. \tag{4.16}$$

Again, applying the inclusion-exclusion principle we derive that

$$U_{i,n}(m) = -\sum_{\substack{0 < \deg k \le \ell \\ k \in \mathcal{S}_n}} \mu(k)\, V_{i,n}(k, m).$$

The pairs (u, v) contributing to $V_{i,n}(k, m)$ are characterized by $\deg u$, $\deg v \le \ell$, $\gcd(u, v) \ne 1$, and the three sets of congruences: $u \equiv v \equiv 1 \bmod x$, (4.15), and (4.16). The first congruence implies that $\gcd(u, u+x^j) = \gcd(v, v+x^j) = 1$ for any $j \ge 0$. It follows that $V_{i,n}(k, m) = 0$ unless the three moduli x, k, and m are pairwise coprime. In that case, $R_{i,n}(k, m) = 2^{2(\ell - \deg km)}$ if $\ell \ge \deg km$, and $R_{i,n}(k, m) \le 1$ otherwise. We obtain

$$V_{i,n}(k, m) = 2^{2\ell - 2\deg km} + O\left(2^{2\ell - 2t}\right) = 2^{2\ell - 2\deg km} + O\left(2^{n/2}\right) \tag{4.17}$$

for any polynomial $k \in \mathcal{S}_n$. It is also clear that

$$V_{i,n}(k, m) \le 2^{n - 2\deg k} \tag{4.18}$$

for any $k \in \mathcal{S}_n$. The rest of the proof is identical to the proof of Theorem 4.7, using (4.17) and (4.18) instead of (4.9) and (4.10). □

4.4. Squarefree and irreducible polynomials. Using the simple sieve method as above one can prove the following result.

LEMMA 4.19. *Let $w \in \mathbb{F}_2[x]$ be a polynomial of degree l such that $w(0) = w(1) = 1$. Then for any $k \geq l - 3$, there are*

$$Q_k = \frac{8}{9}\, 2^k + O(2^{k/2})$$

squarefree polynomials of the form $w + T(T+1)^2 q$ with $q \in \mathbb{F}_2[x]$ of degree less than k.

PROOF. We denote by $\mathcal{W}$ the set of squarefree polynomials $m \in \mathbb{F}_2[x]$ with $m(0) = m(1) = 1$, or, equivalently, with $\gcd(m, x(x+1)) = 1$. For any $m \in \mathcal{W}$ with $\deg m \leq k/2$ the congruence

$$w + x(x+1)^2 q \equiv 0 \bmod m^2$$

has precisely $2^{k-2\deg m}$ solutions $q \in \mathbb{F}_2[x]$ with $\deg q < k$. There are not more than $2^{k+3-2\deg m}$ solutions for polynomials $m \in \mathcal{W}$ with $\deg m \leq \max\{l/2, (k+3)/2\} = (k+3)/2$. It is also clear that if $m \notin \mathcal{W}$ or if $\deg m > (k+3)/2$, then there are no solutions.

Using the same arguments as in the proofs of previous statements we obtain

$$\begin{aligned} Q_k &= \sum_{\substack{\deg m \leq k/2 \\ m \in \mathcal{W}}} \mu(m) 2^{k-2\deg m} + O\left(\sum_{\substack{k/2 \leq \deg m \leq (k+3)/2 \\ m \in \mathcal{W}}} 2^{k+3-2\deg m} \right) \\ &= 2^k \sum_{m \in \mathcal{W}} \mu(m) 2^{-2\deg m} + O(2^{k/2}) = 2^k \prod_{\substack{w \in \mathcal{I} \\ \deg w \geq 2}} \left(1 - 2^{-2\deg w}\right) + O(2^{k/2}) \\ &= \frac{16}{9}\, 2^k \prod_{w \in \mathcal{I}} \left(1 - 2^{-2\deg w}\right) + O(2^{k/2}), \end{aligned}$$

and from Lemma 4.1 we obtain the desired result. □

Finally we need the well-known estimate on the number of irreducible polynomials of given degree; see for example the inequality (3.37) in Berlekamp (1968).

LEMMA 4.20. *For any integer $k \geq 1$ there exist $2^k k^{-1} + O(2^{k/2} k^{-1})$ irreducible monic polynomials of degree k in $\mathbb{F}_2[x]$.*

5. Complexity lower bounds for arithmetic problems for binary polynomials

At this point we are able to derive our main results about the complexity of irreducibility, squarefreeness, and coprimality in various models.

We first consider bounds on the decision tree size.

THEOREM 5.1. *For the squarefreeness function g and the coprimality function h we have*

$$M(g) \geq \frac{1}{3}\,2^n + o(2^n) \qquad \text{and} \qquad M(h) \geq \frac{1}{3}\,2^n + o(2^n).$$

PROOF. If we take $w = (1, \ldots, 1)$ in Lemma 3.2, we obtain the bound $M(\varphi) \geq 2^n |c(\varphi)|$. The two bounds then follow from Lemmas 4.4 and 4.5. □

The worst case decision tree depth is $D(g) = D(h) = n$ by Lemmas 3.3, 4.4, and 4.5. The inequality (2.4) and Theorem 5.1 imply a similar bound for the average depth.

THEOREM 5.2. *For the squarefreeness function g and the coprimality function h we have*

$$\overline{D}(g) \geq n - \log_2 3 + o(1) \qquad \text{and} \qquad \overline{D}(h) \geq n - \log_2 3 + o(1).$$

For the exact real degrees of g and h, Lemmas 3.3, 4.4, and 4.5 immediately imply $\Delta(g) = \Delta(h) = n$. Corollary 3.7 yields a linear lower bound also on the approximate real degrees $\delta(g)$ and $\delta(h)$:

THEOREM 5.3. *For the squarefreeness function g and the coprimality function h, we have*

$$\delta(g) \geq Cn, \qquad \delta(h) \geq Cn,$$

where $C > 0$ is an effectively computable absolute constant.

For formula size, lower bounds of order $\Omega(n^2)$ follow from Lemma 3.8 together with Lemmas 3.1, 4.4, and 4.5. The asymptotic formulas derived in Theorems 4.7 and 4.14 provide the following explicit bounds.

THEOREM 5.4. *For the squarefreeness function g and the coprimality function h we have*

$$L(g) \geq 4\gamma^2 n^2 \qquad \text{and} \qquad L(h) \geq 4\gamma^2 n^2,$$

where γ is defined by (4.6).

Using the same arguments as in the proofs of Lemmas 4.4 and 4.5, one can show that

$$E(g) \sim E(h) \sim 1 - 2 \prod_{w \in \mathcal{I}_0} \left(1 - 2^{-2 \deg w}\right).$$

Hence by Lemma 4.1, $E(g) \sim E(h) \sim -1/3$, and we can obtain explicit values for the constants in the bounds of Theorem 5.4.

In the circuit model, we have the following result.

THEOREM 5.5. *For any odd prime p, the irreducibility function f, the squarefreeness function g and the coprimality function h do not belong to $AC^0[p]$.*

PROOF. First of all we remark that for a polynomial $u \in \mathcal{M}_n$ we have $\gcd(u, x+1) = 1 \iff u(1) = 1 \iff$ parity $(u_1, \ldots, u_n) = 0$. Therefore, from Lemma 3.9 (with $d = 2$) we obtain the desired result for the function h.

If $\gcd(u, x+1) = 1$, then every irreducible polynomial of degree $n+2$ has a unique representation of the form $u + x(x+1)q$ with $\deg q = n$, and there is no such representation if $\gcd(u, x+1) = x+1$.

Now, to test whether $\gcd(u, x+1) = 1$, which as we have seen is equivalent to the parity of the vector $(u_1, \ldots, u_n)$, we test irreducibility of $u + x(x+1)q$ for n^3 random polynomials $q \in \mathbb{F}_2[x]$ of degree n. If $\gcd(u, x+1) = x+1$, then the results of all tests are 'No'. We see from Lemma 4.20 that otherwise with probability at least $1 - 2^{-2n}$ at least one of the tests will return 'Yes'. Now, as in the standard argument of Adleman (1978), there must be at least one set of n^3 polynomials $q \in \mathbb{F}_2[x]$ with $\deg q = n$ such that for all polynomials u of the above form, the corresponding n^3 tests all return 'No' if and only if $\gcd(u, x+1) = x+1$. Because $u + x(x+1)q$ can be computed by a circuit from AC^0, Lemma 3.9 shows that the function f does not belong to $\mathrm{AC}^0[p]$.

Finally, we remark that

$$(x+1)^2 \mid u(x^2) = u^2 \iff x+1 \mid u \iff \text{parity } (u_1, \ldots, u_n) = 1.$$

If $(x+1)^2 \mid u(x^2)$, then the polynomial $u(x^2) + x(x+1)^2 q$ is not squarefree for any $q \in \mathbb{F}_2[x]$. Otherwise from Lemma 4.19 we see that there are at

least $8 \cdot 2^{2n+1}/9 \geq 2^{2n}$ polynomials $q \in \mathbb{F}_2[x]$ of degree at most $2n$ such that $u(x^2) + x(x+1)^2 q$ is squarefree. Repeating the previous arguments, we obtain the desired result for the function g. □

6. Concluding remarks

We remark that using Corollary 2.5 and Lemma 3.8 of Nisan & Szegedy (1994), one can estimate $\Delta(g)$, $\Delta(h)$, $\delta(g)$, and $\delta(h)$ directly from the linear bounds that we have on the average sensitivity. However, this gives only a cn lower bound for $\Delta(g)$ and $\Delta(h)$ (for some $c < 1$) rather than the tight bound of n proved here, and an $\Omega(n^{1/2})$ lower bound for $\delta(g)$ and $\delta(h)$, compared to the linear bound proved here. This approach has been used in Bernasconi *et al.* (2000) for studying the analogue $\widetilde{g}$ of the function g over the integers. Unfortunately the highest order Fourier coefficient of $\widetilde{g}$ seems to be quite small, and thus Corollary 3.7 is not useful for this function.

The only nontrivial lower bound on the complexity of irreducibility testing is given by Theorem 5.5, and at the moment we do not see how to extend other results to this function.

QUESTION 6.1. *Obtain analogs of Theorems 5.1, 5.2, 4.7, and 5.4 for the irreducibility function f.*

Although our results are similar to those of Allender *et al.* (2001), Bernasconi *et al.* (1999, 2000, 2001), and Bernasconi & Shparlinski (1999), we still have not been able to establish complete analogs of the results of Allender *et al.* (2001). Namely it is shown there that the integer primality, squarefreeness, and coprimality functions are hard for the complexity class TC^0. In contrast, we are able only to show that the analogous irreducibility, squarefreeness, and coprimality problems f, g, and h over $\mathbb{F}_2[x]$ are not in $\mathrm{AC}^0[p]$ for any odd prime p. In particular, we cannot rule out the possibility that these problems are in $\mathrm{AC}^0[2]$.

7. Acknowledgment

We would like to thank the anonymous referees for many remarks that improved the presentation.

Allender's work was supported in part by NSF grant CCR-9734918. Part of Damm's work was done while he was with Universität Trier, supported by

DFG grant Me 1077/14-1. The work of von zur Gathen was partly supported by DFG SFB 376 "Massive Parallelität: Algorithmen, Entwurfsmethoden, Anwendungen", that of Saks by NSF grant CCR-9988526, and that of Shparlinski by ARC grant A69700294.

References

L. M. Adleman, Two theorems on random polynomial time. In *Proceedings of the 19th Annual IEEE Symposium on Foundations of Computer Science,* Ann Arbor MI, 1978, 75–83.

Manindra Agrawal, Neeraj Kayal, and Nitin Saxena, PRIMES is in P. Preprint, 2002.

M. Ajtai, Σ_1^1 formulae on finite structures. *Annals of Pure and Applied Logic* **24** (1983). 1-48.

E. Allender, M. Saks, and I. E. Shparlinski, A lower bound for primality. *Journal of Computer and System Sciences* **62** (2001), 356–366.

Eric Bach and Jeffrey Shallit, *Algorithmic Number Theory, Vol.1: Efficient Algorithms.* MIT Press, Cambridge MA, 1996.

Elwin R. Berlekamp, *Algebraic Coding Theory.* McGraw-Hill, New York, 1968.

A. Bernasconi and I. E. Shparlinski, Circuit complexity of testing square-free numbers. In *Advances in Cryptology: Proceedings of CRYPTO '99,* Santa Barbara CA, Lecture Notes in Computer Science **1563**, 47–56. Springer-Verlag, Berlin, 1999.

A. Bernasconi, B. Codenotti, and J. Simon, On the fourier analysis of boolean functions. Preprint, 1996.

A. Bernasconi, C. Damm, and I. Shparlinski, On the average sensitivity of testing squarefree numbers. In *Proceedings of the 5th International Computing and Combinatorics Conference,* Tokyo Japan, Lecture Notes in Computer Science **1627**, Berlin, 1999, Springer-Verlag, 291–299.

A. Bernasconi, C. Damm, and I. E. Shparlinski, The average sensitivity of square-freeness. *computational complexity* **9** (2000), 39–51.

Anna Bernasconi, Carsten Damm, and Igor Shparlinski, Circuit and decision tree complexity of some number theoretic problems. *Information and Computation* **168**(2) (2001), 113–124.

R. B. Boppana, The average sensitivity of bounded-depth circuits. *Information Processing Letters* **63** (1997), 257–261.

H. Cohen, *A course in computational algebraic numbertheory.* Springer-Verlag, Berlin, 1997.

D. Coppersmith and I. E. Shparlinski, On polynomial approximation of the discrete logarithm and the Diffie–Hellman mapping. *Journal of Cryptology* **10** (1998), 233–260.

Martin Dietzfelbinger, Mirosław Kutylowski, and Rüdiger Reischuk, Feasible time-optimal algorithms for boolean functions on exclusive-write parallel random-access machines. *SIAM Journal on Computing* **25** (1996), 1196–1230.

F. E. Fich, The complexity of computation on the parallel random access machine. In *Handbook of Theoretical Comp. Sci.*, 757–804. Elsevier, Amsterdam, 1990.

M. Furst, J. Saxe, and M. Sipser, Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory* **17** (1984), 13–27.

Joachim von zur Gathen and Jürgen Gerhard, *Modern Computer Algebra.* Cambridge University Press, Cambridge, UK, first edition, 1999. Second edition 2003.

S. Jukna, A. Razborov, P. Savicky, and I. Wegener, On p versus np ∩ co-np for decision trees and read-once branching programs. *computational complexity* **8** (1999), 357–370.

E. Kushilevitz and Y. Mansour, Learning decision trees using the fourier spectrum. *SIAM Journal on Computing* **22** (1993), 1331–1348.

N. Linial, Y. Mansour, and N. Nisan, Constant depth circuits, Fourier transforms, and learnability. *Journal of the ACM* **40** (1993), 607–620.

Noam Nisan, Crew prams and decision trees. In *Proceedings of the Twenty-first Annual ACM Symposium on the Theory of Computing,* Seattle WA. ACM Press, 1989, 327–335.

N. Nisan and M. Szegedy, On the degree of boolean functions as real polynomials. *computational complexity* **4** (1994), 301–313.

Ian Parberry and Pei Yuan Yan, Improved upper and lower time bounds for parallel random access machines without simultaneous writes. *SIAM Journal on Computing* **20**(1) (1991), 88–99.

E. Plaku and I. E. Shparlinski, On polynomial representations of boolean functions related to some number theoretic problems. In *Foundations of software technology and theoretical computer science: 21th conference*, Lecture Notes in Computer Science **2245**, Berlin, 2001, Springer-Verlag, 305–316.

I. E. Shparlinski, *Number theoretic methods in cryptography: Complexity lower bounds*. Birkhäuser Verlag, 1999a.

Igor E. Shparlinski, *Finite Fields: Theory and Computation*. Mathematics and Its Applications. Kluwer Academic Publishers, Dordrecht/Boston/London, 1999b.

R. Smolensky, Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proceedings of the Nineteenth Annual ACM Symposium on the Theory of Computing,* New York. ACM Press, 1987, 77–82.

Ingo Wegener, *The Complexity of Boolean Functions*. Wiley-Teubner Series in Computer Science. B. G. Teubner, Stuttgart, and John Wiley & Sons, 1987.

Eric Allender
Department of Computer Science
Rutgers University
Piscataway, NJ 08854-8019, USA
allender@cs.rutgers.edu

Anna Bernasconi
Dipartimento di Informatica
Università di Pisa
Pisa 56125, Italy
annab@di.unipi.it

Carsten Damm
Institut für Numerische und Angewandte Mathematik
Universität Göttingen
D-37083 Göttingen, Germany
damm@math.uni-goettingen.de

Joachim von zur Gathen
Fakultät für Elektrotechnik, Informatik und Mathematik
Universität Paderborn
D-33095 Paderborn, Germany
gathen@uni-paderborn.de

Michael Saks
Mathematics Department
Rutgers University
Piscataway, NJ 08854-8019, USA
saks@math.rutgers.edu

Igor Shparlinski
Department of Computing
Macquarie University
NSW 2109, Australia
igor@comp.mq.edu.au