The Relationship Between Multiplicative Complexity and Nonlinearity*

Joan Boyar and Magnus Gausdal Find

Department of Mathematics and Computer Science
University of Southern Denmark

Abstract. We consider the relationship between nonlinearity and multiplicative complexity for Boolean functions with multiple outputs, studying how large a multiplicative complexity is necessary and sufficient to provide a desired nonlinearity. For quadratic circuits, we show that there is a tight connection between error correcting codes and circuits computing functions with high nonlinearity. Using known coding theory results, the lower bound proven here, for quadratic circuits for functions with \( n \) inputs and \( n \) outputs and high nonlinearity, shows that at least \( 2^{3.32^2} \) AND gates are necessary. We further show that one cannot prove stronger lower bounds by only appealing to the nonlinearity of a function; we show a bilinear circuit computing a function with almost optimal nonlinearity with the number of AND gates being exactly the length of such a shortest code. For general circuits, we exhibit a concrete function with multiplicative complexity at least \( 2^n - 3 \).

1 Definitions and Preliminaries

Let \( \mathbb{F}_2 \) be the finite field of order 2 and \( \mathbb{F}_2^n \) the \( n \)-dimensional vector space over \( \mathbb{F}_2 \). We denote by \( [n] \) the set \( \{1, \ldots, n\} \). An \( (n, m) \)-function is a mapping from \( \mathbb{F}_2^n \) to \( \mathbb{F}_2^m \) and we refer to these as the Boolean functions.

It is well known that every \( (n, 1) \)-function \( f \) can be written uniquely as a multilinear polynomial over \( \mathbb{F}_2 \)

\[
f(x_1, \ldots, x_n) = \sum_{X \subseteq [n]} \alpha_X \prod_{i \in X} x_i.
\]

This polynomial is called the Zhegalkin polynomial or the algebraic normal form of \( f \). For the rest of this paper most, but not all, arithmetic will be in \( \mathbb{F}_2 \). We trust that the reader will find it clear whether arithmetic is in \( \mathbb{F}_2, \mathbb{F}_{2^n}, \) or \( \mathbb{R} \) when not explicitly stated, and will not address it further.

The degree of \( f \) is the largest \( |X| \) such that \( \alpha_X = 1 \). For an \( (n, m) \)-function \( f \), we let \( f_i \) be the \( (n, 1) \)-function defined by the \( i \)th output bit of \( f \), and say

* This is a preliminary version of the paper with the same title. The final and official version will be published in the proceedings of MFCS'2014 and will be available via www.springerlink.com.
that the degree of \( f \) is the largest degree of \( f_i \) for \( i \in [m] \). A function is \textit{affine} if it has degree 1, and \textit{quadratic} if it has degree 2. For \( T \subseteq [m] \) we let

\[
fr = \sum_{i \in T} f_i,
\]

and for \( v \in \mathbb{F}_2^n \) we let \(|v|\) denote the \textit{Hamming weight} of \( v \), that is, the number of nonzero entries in \( v \), and let \(|u + v|\) be the \textit{Hamming distance} between the two vectors \( u \) and \( v \).

We will use several facts on the nonlinearity of Boolean functions. We refer to the two chapters [6,7] by Carlet for proofs and references. The nonlinearity of an \((n, m)\)-function \( f \) is the Hamming distance to the closest affine function, more precisely

\[
NL(f) = 2^n - \max_{a \in \mathbb{F}_2^n, b \in \mathbb{F}_2} |\{ x \in \mathbb{F}_2^n \mid \langle a, x \rangle + b = f(x) \} | ,
\]

where \( \langle a, x \rangle = \sum_{i=1}^n a_i x_i \). For an \((n, m)\)-function \( f \), the nonlinearity is defined as

\[
NL(f) = \min_{T \subseteq [m], T \neq \emptyset} \{ NL(f_T) \}.
\]

The nonlinearity of an \((n, m)\)-function is always between 0 and \( 2^n - 1 - 2^{n-1} \). The \((n, m)\)-functions meeting this bound are called \textit{bent functions}. Bent \((n, 1)\)-functions exist if and only if \( n \) is even. A standard example of a bent \((n, 1)\)-function is the \textit{inner product}, on \( n = 2^k \) variables, defined as:

\[
IP_{2k}(x_1, \ldots, x_k, y_1, \ldots, y_k) = \langle x, y \rangle .
\]

This function is clearly quadratic. If we identify \( \mathbb{F}_2^n \) with \( \mathbb{F}_{2^n} \), a standard example of a bent \((2n, n)\)-function is the \textit{finite field multiplication} function:

\[
f(x, y) = x \cdot y
\]

where multiplication is in \( \mathbb{F}_{2^n} \).

If \( n = m \), \( NL(f) \) is between 0 and \( 2^n - 1 - 2^{n-1} \) [3], and functions meeting this bound are called \textit{almost bent}. These exist only for odd \( n \). As remarked by Carlet, this name is a bit misleading since the name indicates that they are suboptimal, which they are not. Again, if we identify \( \mathbb{F}_2^n \) and \( \mathbb{F}_{2^n} \), for \( 1 \leq i \leq \frac{n-1}{2} \) and \( \gcd(i, n) = 1 \), the so called \textit{Gold functions} defined as

\[
G(x) = x^{2^i+1} = x \cdot (x^{2^i})
\]

are almost bent. This function is quadratic since the mapping \( x \mapsto x^2 \) is affine in \( \mathbb{F}_2^n \), and each output bit of finite field multiplication is quadratic in the inputs, see also [7].

An \textit{XOR-AND circuit} is a Boolean circuit where each of the gates is either \( \oplus \) (XOR, addition in \( \mathbb{F}_2 \)), \( \wedge \) (AND, multiplication in \( \mathbb{F}_2 \)) or the constant 1. In this paper we are mainly concerned with the number of \( \wedge \) gates, so we allow \( \oplus \)-gates
to have unbounded fan-in while \textasciitilde\text{-gates have fan-in 2. A circuit is \textit{quadratic} if every AND gate computes a quadratic function. A quadratic circuit is \textit{bilinear} if the input is partitioned into two sets, and each input to an AND gate is a linear combination of variables from one of these two sets, with the other input using the opposite set of the partition.

The \textit{multiplicative complexity} of an \((n, m)\)-function, \(f\), is the smallest number of AND gates in any XOR-AND circuit computing \(f\). Some relations between nonlinearity and multiplicative complexity are known. In particular, if an \((n, 1)\)-function is to have a certain nonlinearity, it is known exactly how many AND gates are necessary and sufficient.

**Corollary 1 ([1]).** If the \((n, 1)\)-function, \(f\), has multiplicative complexity \(M\), it has nonlinearity at most \(2^n - 1 - 2^{n-M} - 1\). Furthermore this is tight: for \(M \leq \frac{n}{2}\), there exists a simple quadratic function with this nonlinearity.

The upper bound holds for all \(M\), but gives something nontrivial only when \(M < \frac{n}{2}\).

### 1.1 Linear Codes

Most bounds in this paper will come from coding theory. In this subsection, we briefly review the necessary facts. For more information, see chapter 17 in [10] or the older but comprehensive [20].

A \textit{linear (error correcting) code} of length \(s\) is a linear subspace, \(C\) of \(\mathbb{F}_2^s\). The \textit{dimension} of a code is the dimension of the subspace, \(C\), and the elements of \(C\) are called \textit{codewords}. The (minimum) \textit{distance} \(d\) of \(C\) is defined as

\[
d = \min_{x \neq y \in C} |x + y|.
\]

The following fact is well known

**Proposition 1.** For every linear code, \(C\), the distance is exactly the minimum weight among non-zero codewords.

Let \(L(m, d)\) be the length of the shortest linear \(m\)-dimensional code over \(\mathbb{F}_2\) with distance \(d\). We will use lower and upper bounds on \(L(m, d)\). One lower bound is the following [16], see also [20], page 563.

**Theorem 1 (McEliece, Rodemich, Rumsey, Welch).** For \(0 < \delta < 1/2\), let \(C \subseteq \{0,1\}^s\) be a linear code with dimension \(m\) and distance \(\delta s\). Then the rate \(R = \frac{m}{s}\) of the code satisfies \(R \leq \min_{0 \leq u \leq 1-2\delta} B(u, \delta)\), where \(B(u, \delta) = 1 + h(u^2) - h(u^2 + 2\delta u + 2\delta), h(x) = H_2 \left( \frac{1-\sqrt{x}}{2} \right)\), and \(H_2(x) = -x \log x - (1-x) \log(1-x)\).

An upper bound is the following, see [10].

**Theorem 2 (Gilbert-Varshamov).** A linear code \(C \subseteq \{0,1\}^s\) of dimension \(m\) and distance \(d\) exists provided that \(\sum_{i=0}^{d-2} \binom{s-1}{i} < 2^{s-m}\).
2 Introduction

In several practical settings, such as homomorphic encryption and secure multiparty computation (see e.g. [23] and [13]), the number of AND gates is significantly more important than the number of XOR gates, hence one is interested in \((n, m)\)-functions with as few AND gates as possible.

Encryption functions should have high nonlinearity to be resistant against linear and differential attacks (see again [7] and the references therein). This is an explicit design criteria for modern cryptographic systems, such as AES, [9], which has been used has a benchmark for several implementations of homomorphic encryption. A natural question to ask is how these nonlinearity and multiplicative complexity are related to each other: how large does one measure need to be in order for the other to have at least a certain value? As stated in Section 1, for every desired nonlinearity, it is known exactly how many AND gates are necessary and sufficient for an \((n, 1)\)-function to achieve this. We study this same question for functions with multiple bits of output.

Our Contributions Let \(f\) be an \((n, m)\)-function with nonlinearity \(2^n - 1 - 2^n - M - 1\). We show that any quadratic circuit with \(s\) AND gates computing \(f\) defines an \(m\)-dimensional linear code \(F_s\) with distance \(M\), so lower bounds on the size of such codes show lower bounds on the number of AND gates in such a circuit. In particular this implies that any quadratic circuit computing an almost bent function must have at least \(L(n, \frac{n-1}{2})\) AND gates, and that any quadratic function from \(2n\) bits to \(n\) bits with optimal nonlinearity requires quadratic circuits with \(L(n, n)\) AND gates. Since the finite field multiplication function is bent, the \(L(n, n)\) lower bound applies, so the well known result in [3,15], described in the section 2.1, follows immediately as a corollary.

On the other hand, we show that appealing only to the nonlinearity of a function cannot lead to much stronger lower bounds on the multiplicative complexity, by showing the existence of quadratic (in fact, bilinear) circuits with \(L(n, \frac{n}{2})\) AND gates computing a function from \(n\) bits to \(n\) bits with nonlinearity at least \(2^n - 2^{\frac{n}{2} + 3\sqrt{n}}\) which is close to the optimum.

Although almost all Boolean functions with \(n\) inputs and one output have multiplicative complexity at least \(2^n/2 - O(n)\) [2], no concrete function of this type has been shown to have multiplicative complexity more than \(n - 1\). We give a concrete function with \(n\) inputs and \(n\) outputs with multiplicative complexity at least \(2n - 3\).

Using known coding theory bounds, the lower bound proven here, for quadratic circuits for functions with \(n\) inputs and \(n\) outputs and high nonlinearity, shows that at least \(2.32n\) AND gates are necessary. Using a known upper bound on \(L(n, \frac{n}{2})\) gives that circuits for \((n, n)\)-functions with nonlinearity at least \(2^n - 2^{\frac{n}{2} + 3\sqrt{n}}\) can be designed using at most \(2.95n\) AND gates. This is a factor less than 6 times larger than the multiplicative complexity of \((n, 1)\)-functions with similar nonlinearity.
2.1 Related Results

To the best of our knowledge, our lower bound of $2.32n$ AND gates is the largest lower bound on the number of AND gates for quadratic circuits. Previous results showing relations between error correcting codes and bilinear and quadratic circuits include the work of [3,15] where it is shown that a bilinear or quadratic circuit computing finite field multiplication of two $\mathbb{F}_q^n$ elements induces an error correcting code over $\mathbb{F}_q$ of dimension $n$ and distance $n$. For $q = 2$, Theorem 1 implies that such a circuit must have at least $3.52n$ multiplications (AND gates).

If $n$ is the number of input bits, this corresponds to a lower bound of $1.76n$. For $q > 2$, the gates (or lines in a straight-line program) have field elements as inputs, and the total number of multiplications and divisions is counted. Kaminski and Bshouty show a lower bound of $3n - o(n)$ for bilinear circuits [12] and extend it to general circuits [4]. This proof is not based on coding theoretic techniques, but rather the study of Hankel matrices related to the bilinear transformation.

Suppose some $(n,m)$-function $f$ has a certain nonlinearity $D$. If we identify $f_1,\ldots,f_m,x_1,\ldots,x_n$ and the constant 1 with their truth tables as vectors in $\mathbb{F}_2^n$, then $C = \text{span}\{f_1,\ldots,f_m,x_1,\ldots,x_n,1\}$ is a code in $\mathbb{F}_2^n$ with dimension $n + m + 1$ and distance $D$, and limitations and possibilities for codes transfer to results on nonlinearity (see the survey [7] and the references therein). However this says nothing about the multiplicative complexity of the function $f$.

The structure of quadratic circuits has itself been studied by Mirwald and Schnorr [17]. Among other things they show that for quadratic $(n,1)$- and $(n,2)$-functions, quadratic circuits are optimal. It is still not known whether this is true for $(n,m)$-functions in general.

3 Lower Bounds on Multiplicative Complexity

The multiplicative complexity of an $(n,1)$-function is between 0 and $(1 + o(1)) 2^{n/2}$ [18] (see also [11]), and almost all such functions have multiplicative complexity at least $2^{n/2} - O(n)$ [2]. However, there is no value of $n$ where a concrete $(n,1)$-function has been exhibited with a proof that more than $n - 1$ AND gates are necessary to compute it. A lower bound of $n - 1$ follows by the simple degree bound\footnote{Notice that despite the name, this is not the same as Strassen’s degree bound as described in [21] and Chapter 8 of [5].} a function with degree $d$ has multiplicative complexity at least $d - 1$ [19]. Here we show that repeated use of the degree bound gives a concrete $(n,n)$-function, exhibiting a lower bound of $2n - 3$. To the best of our knowledge this is the first example of lower bound on the multiplicative complexity for $(n,n)$-functions.

Theorem 3. The $(n,n)$-function $f$ defined as $f_i(x) = \prod_{j \in [n]\setminus \{i\}} x_j$, has multiplicative complexity at least $2n - 3$.

Proof. Consider the first AND gate, $A$, with degree at least $n - 1$. Such a gate exists since the outputs have degree $n - 1$. By the degree bound, $A$ must have at
least \( p \geq n - 3 \) AND gates with degree at most \( n - 2 \) in its subcircuit. Call these AND gates \( A_1, \ldots, A_p \). None of these AND gates can be an output gate. Suppose there are \( q \) additional AND gates (including \( A \)), where some of these must have degree at least \( n - 1 \). Call these AND gates \( B_1, \ldots, B_q \). Then, for every \( i \in [n] \), there exist \( P_i \subseteq [p] \) and \( Q_i \subseteq [q] \) such that \( f_i = \sum_{j \in P_i} A_j + \sum_{j \in Q_i} B_j \). We can think of each \( B_j \) as a vector in \( \mathbb{F}_2^n \), where the \( i \)th coordinate is 1 if the term \( \prod_{k \in [n]\setminus\{i\}} x_k \) is present in the Zhegalkin polynomial of the function computed by \( B_j \). Since each \( A_j \) has degree at most \( n - 2 \), all the \( A_j \) are zero vectors in this representation, so \( \text{span}(A_1, \ldots, A_p, B_1, \ldots, B_q) = \text{span}(B_1, \ldots, B_q) \). It follows that \( \{f_1, \ldots, f_n\} \subseteq \text{span}(B_1, \ldots, B_q) \). Since

\[ n = \dim(\{f_1, \ldots, f_n\}) \leq \dim(\text{span}(B_1, \ldots, B_q)) \leq q, \]

we conclude that the circuit has at least \( q + p \geq 2n - 3 \) AND gates. \( \square \)

4 Nonlinearity and Multiplicative Complexity

This section is devoted to showing a relation between the nonlinearity and the multiplicative complexity of quadratic circuits. We first show a connection between nonlinearity, multiplicative complexity and certain linear codes. Applying this connection, Theorem 1 gives a bound on any quadratic \((n, m)\)-function.

**Theorem 4.** Let the \((n, m)\)-function, \( f \), have \( \text{NL}(f) \geq 2^{n-1} - 2^{n-M-1} \), where \( M \leq \frac{n}{2} \). Then a quadratic circuit with \( s \) AND gates computing \( f \) exhibits an \( m \)-dimensional linear code over \( \mathbb{F}_2^s \) with distance \( M \).

*Proof.* Let \( C \) be a quadratic circuit with \( s \) AND gates computing \( f \), and let \( A_1, \ldots, A_s \) be the AND gates. Since \( C \) is quadratic, for each \( i \in [m] \) there exist \( S_i \subseteq [s] \) and \( X_i \subseteq [n] \) such that \( f_i \) can be written as

\[ f_i = \sum_{j \in S_i} A_j + \sum_{j \in X_i} x_j. \]

Without loss of generality, we can assume that \( X_i = \emptyset \) for all \( i \), since both nonlinearity and multiplicative complexity are invariant under the addition of affine terms. For each \( i \in [m] \), we define the vector \( v_i \in \mathbb{F}_2^n \), where \( v_{i,j} = 1 \) if and only if there is a directed path from \( A_j \) to the \( i \)th output. By the nonlinearity of \( f \), we have that for each \( i \in [m] \),

\[ \text{NL}(f_i) \geq 2^{n-1} - 2^{n-M-1}. \]

Applying Corollary 1 the multiplicative complexity of \( f_i \) is at least \( M \), hence \( |v_i| \geq M \). Similarly, for any nonempty \( T \subseteq [m] \) we can associate a vector \( v_T \) by setting

\[ v_T = \sum_{i \in T} v_i. \]
Since the circuit is quadratic, it holds that if $|v_T| \leq p$, the multiplicative complexity of $f_T = \sum_{i \in T} f_i$ is at most $p$. Applying the definition of nonlinearity to $f_T$, $NL(f_T) \geq 2^{n-1} - 2^n - M^{-1}$. Corollary 4 implies that the multiplicative complexity of $f_T$ is at least $M$, so we have that $|v_T| \geq M$ when $T \neq \emptyset$.

In conclusion, every nonzero vector in the $m$ dimensional vector space $\mathcal{C} = \text{span}_{\mathbb{F}_2}\{v_1, \ldots, v_m\}$ has Hamming weight at least $M$. By Proposition 1, $\mathcal{C}$ is a linear code with dimension $m$ and distance at least $M$.

Applying this theorem to quadratic almost bent functions, we have that a quadratic circuit computing such a function has at least $L(n, \frac{n-1}{2})$ AND gates. Combining this with Theorem 1, calculations show:

**Corollary 2.** Any quadratic circuit computing an almost bent $(n, n)$-function has at least $L(n, \frac{n-1}{2})$ AND gates. For sufficiently large $n$, $L(n, \frac{n-1}{2}) > 2.32n$.

The corollary above applies to e.g. the almost bent Gold functions $G$ defined in Eqn. (2). For bent $(2n, n)$-functions, using Theorem 4 with $M = n$ and applying Theorem 1, calculations show:

**Corollary 3.** A quadratic circuit computing any bent $(2n, n)$-function has at least $L(n, n)$ AND gates. For sufficiently large $n$, $L(n, n) > 3.52n$.

This applies to e.g. the finite field multiplication function as defined in Eqn. 1, reproving the known result on multiplicative complexity for quadratic circuits for field multiplication mentioned in Section 2.1.

For both Corollaries 2 and 3 any improved lower bounds on codes lengths would give an improved lower bound on the multiplicative complexity. For Corollary 2 this technique cannot prove better lower bounds than $L(n, \frac{n-1}{2})$. Theorem 2 implies that $L(n, \frac{n-1}{2}) \leq 2.95n$. Below we show that this is not merely a limitation of the proof strategy; there exist quadratic circuits with $L(n, \frac{n-1}{2})$ AND gates with nonlinearity close to the optimal. To the best of our knowledge this is the first example of highly nonlinear $(n, n)$-functions with linear multiplicative complexity, and therefore it might be a useful building block for cryptographic purposes.

Before proving the next theorem, we need a technical lemma on the probability that a random matrix has small rank. A simple proof of this can be found in e.g. [14].

**Lemma 1 (Komargodski, Raz, Tal).** A random $k \times k$ matrix has rank at most $d$ with probability at most $2^k - (k-d)^d$.

**Theorem 5.** There exist $(n, n)$-functions with multiplicative complexity at most $L(n, \frac{n-1}{2})$ and nonlinearity at least $2^{n-1} - 2^\frac{n}{3} + 3\sqrt{n} - 1$.

**Proof.** For simplicity we show the upper bound for $L(n, \frac{n}{2})$ AND gates. It is elementary to verify that it holds for $L(n, \frac{n-1}{2})$ AND gates as well. We give a probabilistic construction of a quadratic (in fact, bilinear) circuit with $s = L(n, \frac{n}{2})$
AND gates, then we show that with high probability, the function computed by this circuit has the desired nonlinearity.

For the construction of the circuit, we first define the value computed by the $i$th AND gate as $A_i(x) = L_i(x)R_i(x)$ where $L_i$ is a random sum over $x_1, \ldots, x_{n/2}$ and $R_i$ is a random sum over $x_{n/2+1}, \ldots, x_n$. In the following, we will identify sums over $x_1, \ldots, x_n$ with vectors in $\mathbb{F}_2^n$ and sums over $A_1, \ldots, A_s$ with vectors in $\mathbb{F}_2^s$.

Let $C$ be an $n$-dimensional code of length $L(n, \frac{n}{2})$ with distance $\frac{n}{2}$ and let $y_1, \ldots, y_n \in \mathbb{F}_2^n$ be a basis for $C$. Now we define the corresponding sums over $A_1, \ldots, A_s$ to be the outputs computed by the circuit. This completes the construction of the circuit. Now fix $r(x) \in \text{span}_{\mathbb{F}_2} \{y_1, \ldots, y_n\}$, $r \neq 0$. We want to show that $r$ has the desired nonlinearity with high probability. By an appropriate relabeling of the AND gates, we can write $r(x) = \sum_{i=1}^q A_i(x) = \sum_{i=1}^q L_i(x)R_i(x)$ (3)

for some $q \geq \frac{n}{2}$. We now assume that

$$t = \text{rk}\{R_1, \ldots, R_q\} \geq \frac{n}{2} - \frac{3\sqrt{n}}{2}. \quad (4)$$

At the end of the proof, we will show that this is true with high probability. Again by an appropriate relabeling, we let $\{R_1, \ldots, R_q\}$ be a basis of $\text{span}\{R_1, \ldots, R_q\}$. If $q > t$, for $j > t$, we can write $R_j = \sum_{i=1}^t \alpha_{j,i}R_i$. In particular for $j = q$, we can substitute this into (3) and obtain

$$r(x) = \sum_{i=1}^{q-1} (L_i(x) + \alpha_{q,i}L_q(x)) R_i(x)$$

where we let $\alpha_{q,i} = 0$ for $i > t$. If $\{L_1, \ldots, L_q\}$ are independently, uniformly randomly distributed, then so are $\{L_1 + \alpha_{q,1}L_q, \ldots, L_{q-1} + \alpha_{q,q-1}L_q\}$. Continuing this process, we get that for $\frac{n}{2} \geq t \geq \frac{n}{2} - \frac{3\sqrt{n}}{2}$, there are sums $L_1', \ldots, L_t', R_1', \ldots, R_t'$ such that

$$r(x) = \sum_{i=1}^t L_i'(x)R_i'(x)$$

where the $\{L_1', \ldots, L_t'\}$ are independently, uniformly random and the $\{R_1', \ldots, R_t'\}$ are linearly independent. We now further assume that

$$u = \text{rk}\{L_1', \ldots, L_t'\} \geq t - \frac{3\sqrt{n}}{2}. \quad (5)$$

Again, we will show at the end of this proof that this is true with high probability. Applying a similar procedure as above, we get that for some

$$u \geq t - \frac{3\sqrt{n}}{2} \geq \frac{n}{2} - 3\sqrt{n}.$$
there exist sums $\tilde{L}_1, \ldots, \tilde{L}_u$ and $\tilde{R}_1, \ldots, \tilde{R}_u$, such that

$$r(x) = \sum_{i=1}^u \tilde{L}_i(x)\tilde{R}_i(x),$$

where all $\tilde{L}_1, \ldots, \tilde{L}_u$ and all $\tilde{R}_1, \ldots, \tilde{R}_u$ are linearly independent. Thus, there exists a linear bijection $(x_1, \ldots, x_n) \mapsto (z_1, \ldots, z_n)$ with $z_1 = \tilde{L}_1, \ldots, z_u = \tilde{L}_u, z_{u+1} = \tilde{R}_1, \ldots, z_{2u} = \tilde{R}_u$, such that

$$\tilde{r}(z) = z_1z_{u+1} + \ldots, z_{2u}$$

where $r$ and $\tilde{r}$ are equivalent up to a linear bijection on the inputs. Since nonlinearity is invariant under linear bijections, we just need to determine the nonlinearity of $\tilde{r}$. Given the high nonlinearity of $IP_n$, it is elementary to verify that

$$NL(\tilde{r}) = 2^{n-2u} \left(2^{2u-1} - 2^{u-1} \right) = 2^{n-1} - 2^{n-u-1}.$$ 

If $u \geq \frac{n}{2} - 3\sqrt{n}$, this is at least $2^{n-1} - 2^{\frac{n}{2}+3\sqrt{n}-1}$.

Now it remains to show that the probability of either (4) or (5) occurring is so small that a union bound over all the $2^n - 1$ choices of $r$ gives that with high probability, every $r \in \text{span}\{y_1, \ldots, y_n\}$ has at least the desired nonlinearity.

For (4), we can think of the $q \geq \frac{n}{2}$ vectors $R_1, \ldots, R_q$ as rows in a $q \times \frac{n}{2}$ matrix. We will consider the upper left $\frac{n}{2} \times \frac{n}{2}$ submatrix. By Lemma 1 this has rank at most $\frac{n}{2} - \frac{3\sqrt{n}}{2}$ with probability at most

$$2^{\frac{n}{2} - \left(\frac{n}{2} - \frac{3\sqrt{n}}{2} \right)^2} = 2^{\frac{n}{2} - \frac{9n}{4}} = 2^{-\frac{7n}{4}}$$

Similarly for (5) we can consider the $\frac{n}{2} \geq t \geq \frac{t}{2} - \frac{3\sqrt{n}}{2}$ vectors $L'_1, \ldots, L'_t$ as the rows in a $t \times \frac{n}{2}$ matrix. Consider the top left $t \times t$ submatrix. Again, by Lemma 1 the probability of this matrix having rank at most $t - \frac{3\sqrt{n}}{2}$ is at most

$$2^t - \left(t - \frac{3\sqrt{n}}{2} \right)^2 \leq 2^t - \frac{9n}{4} = 2^{-\frac{7n}{4}}$$

There are $2^n - 1$ choices of $r$, so by the union bound, the total probability of at least one of (4) or (5) failing for a least one choice is at most $2 \cdot (2^n - 1) \cdot 2^{-\frac{7n}{4}}$, which tends to zero, so in fact the described construction will have the desired nonlinearity with high probability.

We should note that it is not hard to improve in the constants in the proof and show that in fact the described function has nonlinearity at least $2^{n-1} - 2^{\frac{n}{2}+c\sqrt{n}}$ for some constant $c < 3$. However, the proof given does not allow improvement to e.g. $c = 2$. 

9
5 Open Problems

Strassen [22] (see also [5], Proposition 14.1, p. 351) proved that for an infinite field, $\mathbb{K}$, if the quadratic function $F: \mathbb{K}^n \rightarrow \mathbb{K}^m$ can be computed with $M$ multiplications/divisions, then it can be computed in $M$ multiplications by a quadratic circuit. However, it is unknown whether a similar result holds for finite fields in particular for $\mathbb{F}_2$. Mirwald and Schnorr [17] showed that for quadratic $(n, 1)$- and $(n, 2)$-functions, quadratic circuits are optimal. It is still not known whether this is true for $(n, m)$-functions in general. It would be very interesting to determine if the bounds proven here for quadratic circuits also hold for general circuits.

When inspecting the proof of Theorem 1 one can make a weaker assumption on the circuit than it being quadratic. For example, it is sufficient if it holds that for every AND gate, $A$, there is a unique AND gate, $A'$ (which might be equal to $A$), such that every path from $A$ to an output goes through $A'$. Can one find a larger, interesting class of circuits where the proof holds?

The function defined in Theorem 3 has multiplicative complexity at least $2n - 3$ and at most $3n - 6$. What is the exact value?

References

1. Boyar, J., Find, M., Peralta, R.: Four measures of nonlinearity. In: Spirakis, P.G., Serna, M.J. (eds.) CIAC 2013. LNCS, vol. 7878, pp. 61–72. Springer (2013), eprint with correction available at the Cryptology ePrint Archive, Report 2013/633, 2013, http://eprint.iacr.org/
2. Boyar, J., Peralta, R., Pochuev, D.: On the multiplicative complexity of Boolean functions over the basis ($\land$, $\oplus$, 1). Theor. Comput. Sci. 235(1), 43–57 (2000)
3. Brown, M.R., Dobkin, D.P.: An improved lower bound on polynomial multiplication. IEEE Trans. Computers 29(5), 337–340 (1980)
4. Bshouty, N.H., Kaminski, M.: Polynomial multiplication over finite fields: from quadratic to straight-line complexity. Computational Complexity 15(3), 252–262 (2006)
5. Bürgisser, P., Clausen, M., Shokrollahi, M.A.: Algebraic Complexity Theory, Grundlehren der mathematischen Wissenschaften, vol. 315. Springer (1997)
6. Carlet, C.: Boolean functions for cryptography and error correcting codes. In: Crama, Y., Hammer, P.L. (eds.) Boolean Models and Methods in Mathematics, Computer Science, and Engineering, chap. 8, pp. 257–397. Cambridge, UK: Cambridge Univ. Press (2010)
7. Carlet, C.: Vectorial Boolean functions for cryptography. In: Crama, Y., Hammer, P.L. (eds.) Boolean Models and Methods in Mathematics, Computer Science, and Engineering, chap. 9, pp. 398–469. Cambridge, UK: Cambridge Univ. Press (2010)
8. Chabaud, F., Vaudenay, S.: Links between differential and linear cryptoanalysis. In: Santis, A.D. (ed.) EUROCRYPT ’94. pp. 356–365. LNCS, Springer (1995)
9. Daemen, J., Rijmen, V.: The Design of Rijndael: AES-The Advanced Encryption Standard. Security and Cryptology, Springer (2002)
10. Jukna, S.: Extremal Combinatorics: with Applications in Computer Science. Texts in Theoretical Computer Science, Springer, second edn. (2011)
11. Jukna, S.: Boolean Function Complexity: Advances and Frontiers. Springer Berlin Heidelberg (2012)
Theorem 3 demonstrated a lower bound of $2^n - 3$ on the multiplicative complexity of the $(n,n)$-function $f$ defined as
\[
f_i(x) = \prod_{j \in [n] \setminus \{i\}} x_j.
\]

In the Open Problems section, an upper bound of $3n - 6$ was claimed. This can be seen from the following construction:

1. Use $n - 2$ AND gates with the following outputs:
\[
A = \{x_1x_2, x_1x_2x_3, ..., x_1x_2 \ldots x_{n-1}\}.
\]
2. Use $n - 2$ AND gates with the following outputs:
\[
B = \{x_2x_3 \ldots x_n, x_3x_4 \ldots x_n, ..., x_{n-4}x_{n-3}x_{n-2}x_{n-1}\}.
\]
3. Use $n - 4$ AND gates to AND together the $i$th element of $A$ with the $i + 2$nd element of $B$, for $1 \leq i \leq i - 4$. 

Appendix
4. Compute \( x_1 \land x_3 x_4 \ldots x_n \) and \( x_1 x_2 \ldots x_{n-2} \land x_n \).

**Proof (Of Corollary 2).** Recall that almost bent functions have nonlinearity
\( 2^{n-1} - 2^{n-2} \), so in terms of Theorem 4 we have \( M = \frac{n-1}{2} \). Suppose for the sake
of contradiction that \( L(n, \frac{n-1}{2}) \leq 2.32n \) for infinitely many values of \( n \). Then
we have an \( n \)-dimensional code on \( \mathbb{F}_2^{2.32n} \) with distance \( \frac{n-1}{2} \). In Theorem 1 for
sufficiently large \( n \) we can let
\[
\delta = \frac{1}{2 \cdot 2.32} - 0.0001 \approx 0.2154.
\]
The rate of the code is \( R = \frac{1}{2 \cdot 2.32} > 0.431 \). Choosing \( u = 0.32 \) shows that the rate \( R \) satisfies
\[
R \leq B(0.32, 0.2155) < 0.42,
\]
contradicting the assumption that \( L(n, \frac{n-1}{2}) \leq 2.32n \). \( \square \)

The calculations below using Theorem 1 are not new; they are only included
for completeness.

**Proof (Of Corollary 3).** Analogously to the proof above, in terms of Theorem 1
we have \( m = n \) and \( M = n \). Again suppose for the sake of contradiction that
\( L(n, n) \leq 3.52 \) for infinitely many values of \( n \). In Theorem 1 for sufficiently large
\( n \) we have that
\[
R = \frac{1}{3.52} \approx 0.28409.
\]
Choosing \( u = 0.4 \) in Theorem 1 shows that the rate \( R \) satisfies
\[
R \leq B(0.4, 0.28409) \approx 0.2826005815 < 0.284..
\]
contradicting the assumption that \( L(n, n) \leq 3.52 \). \( \square \)