Basic Theory in Construction of Boolean Functions with Maximum Possible Annihilator Immunity*

Deepak Kumar Dalai, Subhamoy Maitra and Sumanta Sarkar
Applied Statistics Unit, Indian Statistical Institute
203 B T Road, Kolkata 700 108, INDIA
Email: {deepak_r, subho, sumanta_r}@isical.ac.in

Abstract

So far there is no systematic attempt to construct Boolean functions with maximum annihilator immunity. In this paper we present a construction keeping in mind the basic theory of annihilator immunity. This construction provides functions with the maximum possible annihilator immunity and the weight, nonlinearity and algebraic degree of the functions can be properly calculated under certain cases. The basic construction is that of symmetric Boolean functions and applying linear transformation on the input variables of these functions, one can get a large class of non-symmetric functions too. Moreover, we also study several other modifications on the basic symmetric functions to identify interesting non symmetric functions with maximum annihilator immunity. In the process we also present an algorithm to compute the Walsh spectra of a symmetric Boolean function with $O(n^2)$ time and $O(n)$ space complexity.

Keywords: Algebraic Attack, Algebraic Degree, Algebraic Immunity, Annihilator, Annihilator Immunity, Balancedness, Boolean Functions, Krawtchouk Polynomials, Nonlinearity, Symmetric Boolean Functions.

1 Introduction

Algebraic attack (that uses overdefined systems of multivariate equations to recover the secret key) has received a lot of attention recently [1,2,8,9,11–13,19,23] in studying security of the cryptosystems. This adds a new cryptographic property for designing Boolean functions to be used as building blocks in cryptosystems which is known as algebraic immunity [3–5, 7, 14, 15, 23]. Later, in Remark 1, we will discuss some problem about the

[*] We use the term "Annihilator Immunity" instead of "Algebraic Immunity" referred in the recent papers [3–5, 7, 14, 15]. Please see Remark 1 for the details of this notational change.

term “algebraic immunity” and use the term “annihilator immunity” instead of the earlier term.

Given an n-variable Boolean function f, different cases related to low degree multiples of f have been studied in [12, 23]. The main objective is to find out minimum (or low) degree annihilators of f and $1 + f$, i.e., to find out minimum (or low) degree n-variable nonzero functions g_1, g_2 such that $f * g_1 = 0$ and $(1 + f) * g_2 = 0$. To mount the algebraic attack, one needs the low degree linearly independent annihilators [12, 23] of f and $1 + f$.

Though there are increasing interest in construction of Boolean functions with good annihilator immunity [3–5, 7, 14, 15], so far there is only one construction method [15] that can achieve the maximum possible annihilator immunity $\lceil \frac{n}{2} \rceil$ for an n-variable function. The heart of the construction in [15] was a function ϕ_{2k} on even $(2k)$ number of variables with maximum possible annihilator immunity k. The main problem with ϕ_{2k} is that no clear intuition has been provided how one can land into such a complicated structure. Further, the other cryptographic properties, such as weight, nonlinearity or algebraic degree of the function ϕ_{2k} are yet to be answered and only a few experimental results have been provided in [15] for $k = 1, \ldots, 8$. Also the functions ϕ_{2k} are not balanced.

In this paper we explain a generic construction idea of functions with maximum annihilator immunity that comes from the basic theory. Most importantly, the cryptographic properties of our constructions, such as nonlinearity, algebraic degree etc., are theoretically proved for certain subcases that could not be done for the construction in [15]. Interestingly, for this subcase of our construction with even n, the weight and nonlinearity (both $2^{n-1} - \binom{n-1}{\frac{n}{2}}$, we provide exact proofs) matches with the experimental data provided on ϕ_{2k} in [15] (no proof). However our functions (in the subcase) are not linear transformation of ϕ_{2k} as the algebraic degree of our construction $(2^{\lfloor \log_2 n \rfloor})$ is different from the experimental results available in [15].

We also provide a large class of balanced Boolean functions with maximum possible annihilator immunity having nonlinearity $\geq 2^{n-1} - \binom{n}{\frac{n}{2}}$. Under experimental set up, with a simple heuristic, we show that actually one can achieve much better nonlinearity than this lower bound (in fact very close to $2^{n-1} - \binom{n-1}{\frac{n}{2}}$). For odd n our construction provides balanced functions directly with nonlinearity $2^{n-1} - \binom{n-1}{\frac{n-1}{2}}$ and algebraic degree $(2^{\lfloor \log_2 n \rfloor})$.

As our basic construction starts from symmetric Boolean functions and the Walsh spectra of Boolean functions are related to Krawtchouk Polynomials, we need to use the properties of Krawtchouk Polynomials extensively. In the process we identify an interesting inequality as explained in Lemma 5. Further, we present an algorithm for calculating the Walsh spectra of symmetric Boolean functions. This requires $O(n^2)$ time and $O(n)$ space complexity for a symmetric Boolean function on n variables. To the best of our knowledge, this algorithm is novel and it is not easy to improve it further. The algorithm does not only use the direct relationship between Walsh spectra of symmetric Boolean function and Krawtchouk polynomial, but we need to integrate different properties of Krawtchouk polynomial to get the optimized algorithm.

It is well known that the annihilator immunity (also algebraic degree and nonlinearity) of a Boolean function is invariant under linear transformation on the input variables. Thus one can easily apply linear transformation to get a wider class of functions (which are not symmetric) from our construction achieving the maximum possible annihilator immunity (with same algebraic degree and nonlinearity).

Note that there is much scepticism towards using symmetric Boolean functions in cryptosystems. Moreover, the other cryptographic properties of the Boolean functions (whether symmetric or not) that we consider here are not very good (these functions may be composed with other functions with different cryptographic properties to get a secondary construction and those functions may be used in a cryptosystem). Thus, in no way, we are proposing these functions for direct use in cryptosystems. The motivation of this paper is systematic theoretical study of Boolean functions with maximum possible annihilator immunity (see also Remark 1).

The organization of the paper is as follows. In the following section we present the basic theory behind the construction of Boolean functions with maximum possible annihilator immunity and present a specific construction. In Section 3, we consider symmetric functions with maximum possible annihilator immunity and calculate the algebraic degree and nonlinearity of the functions. Some extensions and comparison of the parameters with a very recent construction method [15] is presented in Section 4. In Section 5, we discuss the algorithm for calculating Walsh spectra of a symmetric Boolean function. Section 6 concludes the paper.

2 Construction using the Basic Theory

Let us denote the set of n-variable Boolean functions by B_n. The support of a Boolean function $f \in B_n$ is defined as $supp(f) = \{(x_1, \ldots, x_n) | f(x_1, \ldots, x_n) = 1\}$. The weight of a function $f \in B_n$ is $wt(f) = |supp(f)|$. A function $f \in B_n$ is balanced if $wt(f) = 2^{n-1}$.

Any $f \in B_n$ can be uniquely represented as a multivariate polynomial over $GF(2)$, called the algebraic normal form (ANF), as

$$f(x_1, \ldots, x_n) = a_0 + \sum_{1 \le i \le n} a_i x_i + \sum_{1 \le i < j \le n} a_{i,j} x_i x_j + \ldots + a_{1,2,\ldots,n} x_1 x_2 \ldots x_n,$$

where the coefficients $a_0, a_{i,j}, \ldots, a_{1,2,\ldots,n} \in \{0, 1\}$. The algebraic degree, $\deg(f)$, is the number of variables in the highest order term with non zero coefficient. A Boolean function is affine if there exists no term of degree > 1 in the ANF and the set of all affine functions is denoted by A_n. An affine function with constant term equal to zero is called a linear function.

A Boolean function should be of high algebraic degree to be cryptographically secure [17]. Further, to resist algebraic attack, the function should not have a low degree multiple [12,23]. It is shown [12] that given any n-variable Boolean function f, it is always possible to get a Boolean function g with degree at most $\lceil \frac{n}{2} \rceil$ such that $f * g$ is of degree at most $\lceil \frac{n}{2} \rceil$. Here the functions are considered to be multivariate polynomials over $GF(2)$ and $f * g$ is the polynomial multiplication over $GF(2)$. Thus while choosing an f, the cryptosystem designer should be careful that it should not happen that degree of $f * g$

falls much below $\lceil \frac{n}{2} \rceil$. Towards defining annihilator immunity [12,14,15,23], it is now clear that one needs to consider the annihilators of both $f, 1+f$. In that line we present the following definition.

Definition 1

1. *Given $f \in B_n$, a nonzero function $g \in B_n$ is called an annihilator of f if $f * g = 0$. By $AN(f)$ we mean the set of annihilators of f.*

2. *Given $f \in B_n$, the annihilator immunity of f, denoted by $\mathcal{AI}_n(f) = \deg(g)$, where $g \in B_n$ is the minimum degree nonzero function such that either $f * g = 0$ or $(1 + f) * g = 0$.*

It is known [12,23] that for $f \in B_n$, $\mathcal{AI}_n(f) \leq \lceil \frac{n}{2} \rceil$ and in this paper we present constructions achieving the maximum value.

Remark 1 *At this point we like to discuss on the term "algebraic immunity" of a Boolean function. Recently there are many works in the area of algebraic attacks and some of the initial and important papers are [11–13]. It is now clear that a Boolean function or its complement, used in a cryptosystem, should not have a low degree annihilator. However, the algebraic normal form (ANF) of the annihilators are also important. It may very well happen that an annihilator with higher degree may have a few terms and on the other hand an annihilator with lower degree may have many more terms in the ANF and in certain cases, it may be better to use the high degree annihilator with fewer terms than the low degree annihilator with more terms for the algebraic attack. Thus increase in the degree of annihilator (of the Boolean function) may not be the only measure in terms of resistance of a cryptosystem (that uses the Boolean function) against algebraic attack. Based on the existing research so far, it is difficult to formalize or quantify the measure of resistance of a Boolean function used in a cryptosystem against algebraic attack. It clearly depends on how the Boolean function is used in the construction of cryptosystem and how the algebraic attack is designed against the complete scheme. These arguments go against using the term "algebraic immunity".*

On the other hand, if one just concentrates on a Boolean function, then it is meaningful to consider the annihilators of $f, 1+f$ to study its resistance against algebraic attack and one would always like to get a Boolean function f, such that both f and $1+f$ do not have any annihilator with degree $< \lceil \frac{n}{2} \rceil$. Further, if one considers the algebraic degree of an n-variable Boolean function, then it may very well happen that the function $f(x_1, x_2, \ldots, x_n)$ is of very good algebraic degree, but if one conditions one variable, say $f(x_1 = 0, x_2, \ldots, x_n)$, the degree falls drastically. However, this is not true in terms of algebraic immunity. It can be checked that if f has algebraic immunity t, then after conditioning any k variables, the algebraic immunity of the subfunction on $n-k$ variables will be $\geq t-k$. This is clearly a stronger property than the algebraic degree of a Boolean function. Based on these arguments and as the term has already been appeared in many papers [3–5, 7, 14, 15], one may be tempted to use the term "algebraic immunity".

To get out of this confusion, in this paper we use the term "annihilator immunity" as this clearly quantifies the measure how good a Boolean function is in terms of not having low degree annihilators, and we feel this is also a properly definable necessary (may not be sufficient) condition for a Boolean function with respect to the resistance against algebraic attack. As long as no better (and properly quantifiable) definition related to Boolean function is proposed in terms of resistance against algebraic attack, "annihilator immunity" of Definition 1(2) remains an important topic to study in the field of cryptographically significant Boolean functions.

The idea of our construction comes from the following.

Construction 1 *Let $f, f_1, f_2 \in B_n$ with the following conditions.*

1. *There is no annihilator of f_1, f_2 having degree $< \lceil \frac{n}{2} \rceil$.*
2. *$supp(f) \supseteq supp(f_2)$ and $supp(1+f) \supseteq supp(f_1)$.*

Then we have the following important result.

Lemma 1 *Let $f \in B_n$ be a function as described in Construction 1. Then $\mathcal{AI}_n(f) = \lceil \frac{n}{2} \rceil$.*

Proof: As $supp(1+f) \supseteq supp(f_1)$, $AN(1+f) \subseteq AN(f_1)$ and as $supp(f) \supseteq supp(f_2)$, $AN(f) \subseteq AN(f_2)$. Since there is no annihilator of f_1, f_2 having degree $< \lceil \frac{n}{2} \rceil$, neither f nor $1+f$ can have any annihilator of degree $< \lceil \frac{n}{2} \rceil$. Thus $\mathcal{AI}_n(f) = \lceil \frac{n}{2} \rceil$. ■

Now we present the other direction.

Lemma 2 *Let $f \in B_n$ and $\mathcal{AI}_n(f) = \lceil \frac{n}{2} \rceil$. Then there exist $f_1, f_2 \in B_n$ with $supp(f_1) \subseteq supp(1+f)$ and $supp(f_2) \subseteq supp(f)$ such that $wt(f_1) = wt(f_2) = \sum_{i=0}^{\lceil \frac{n}{2} \rceil - 1} \binom{n}{i}$ and f_1, f_2 have no annihilator of degree $< \lceil \frac{n}{2} \rceil$.*

Proof: Since $\mathcal{AI}_n(f) = \lceil \frac{n}{2} \rceil$, f has no annihilator of degree $< \lceil \frac{n}{2} \rceil$. That is, there cannot be any $g(x_1, \ldots, x_n) = a_0 + \sum_{i=0}^{n} a_i x_i + \cdots + \sum_{1 \le i_1 \ldots \le i_{\lceil \frac{n}{2} \rceil - 1} \le n} a_{i_1 \ldots i_{\lceil \frac{n}{2} \rceil - 1}} x_{i_1} \cdots x_{i_{\lceil \frac{n}{2} \rceil - 1}}$ such that $g(x_1, \ldots, x_n) = 0$ where $f(x_1, \ldots, x_n) = 1$. That is there is no nonzero solution of the system of homogeneous linear equations $g(x_1, \ldots, x_n) = 0$ for $(x_1, \ldots, x_n) \in supp(f)$ on a_i's, i.e., this system has full rank $(\sum_{i=0}^{\lceil \frac{n}{2} \rceil - 1} \binom{n}{i})$. So, there must be $\sum_{i=0}^{\lceil \frac{n}{2} \rceil - 1} \binom{n}{i}$ many linearly independent equations. Now we construct f_2 such that $supp(f_2)$ is the set of input vectors corresponding to $\sum_{i=0}^{\lceil \frac{n}{2} \rceil - 1} \binom{n}{i}$ many linearly independent equations. So, f_2 has no annihilator of degree $< \lceil \frac{n}{2} \rceil$. Similarly, we can construct f_1 considering $(1+f)$ has no annihilator of degree $< \lceil \frac{n}{2} \rceil$. ■

Based on Lemma 1 and Lemma 2, we get a clear idea of a construction strategy for a function with maximum possible annihilator immunity.

For odd n, there is no option other than $f_1 = f$ and $f_2 = 1+f$ to have maximum annihilator immunity for f, since $wt(f_1) + wt(f_2) = 2 \sum_{i=0}^{\lceil \frac{n}{2} \rceil - 1} \binom{n}{i} = 2^n$. This fact also follows from [14, Corollary 1] that a function on odd number of variables must be balanced (weight

2^{n-1} for n-variable function) to achieve the maximum possible annihilator immunity. Also recently it has been shown [6] that for balanced functions on odd number of variables, it is enough to consider the annihilators of f (the case for $1+f$ will automatically be deduced) in terms of maximum annihilator immunity. The exact result is as follows.

Proposition 1 [6] *Let $\psi \in B_n$ (n odd) be balanced function and it does not have any annihilator with algebraic degree $< \lceil \frac{n}{2} \rceil$. Then $1+\psi$ has no annihilator with algebraic degree $< \lceil \frac{n}{2} \rceil$. Consequently, $\mathcal{AI}_n(\psi) = \lceil \frac{n}{2} \rceil$.*

However, for even n, $wt(f_1)+wt(f_2) = 2\sum_{i=0}^{\lceil \frac{n}{2} \rceil - 1} \binom{n}{i} = 2^n - \binom{n}{\frac{n}{2}}$. So, a part of remaining $\binom{n}{\frac{n}{2}}$ output points can be chosen randomly to get different functions f without affecting the annihilator immunity. Hence for even n case this restriction is not as strict as odd n case.

2.1 A construction for maximum Annihilator Immunity

Let us now present the application of the basic theory for a concrete construction of functions having optimal annihilator immunity.

Construction 2 *Let $f \in B_n$.*

1. *If n is odd then*

$$\begin{aligned} f(x_1,\ldots,x_n) &= 0 \textit{ for } wt(x_1,\ldots,x_n) \le \lfloor \frac{n}{2} \rfloor, \\ &= 1 \textit{ for } wt(x_1,\ldots,x_n) \ge \lceil \frac{n}{2} \rceil. \end{aligned}$$

2. *If n is even then*

$$\begin{aligned} f(x_1,\ldots,x_n) &= 0 \textit{ for } wt(x_1,\ldots,x_n) < \frac{n}{2}, \\ &= 1 \textit{ for } wt(x_1,\ldots,x_n) > \frac{n}{2}, \\ &= b \in \{0,1\} \textit{ for } wt(x_1,\ldots,x_n) = \frac{n}{2}. \end{aligned}$$

Lemma 3 *Define two functions $f_1, f_2 \in B_n$ as follows.*

$$\begin{aligned} f_1(x_1,\ldots,x_n) &= 1 \textit{ for } wt(x_1,\ldots,x_n) < \lceil \frac{n}{2} \rceil, \\ &= 0 \textit{ for } wt(x_1,\ldots,x_n) \ge \lceil \frac{n}{2} \rceil. \end{aligned}$$

$$\begin{aligned} f_2(x_1,\ldots,x_n) &= 0 \textit{ for } wt(x_1,\ldots,x_n) \le \lceil \frac{n}{2} \rceil, \\ &= 1 \textit{ for } wt(x_1,\ldots,x_n) > \lceil \frac{n}{2} \rceil. \end{aligned}$$

Then f_1, f_2 have no annihilator of degree $< \lceil \frac{n}{2} \rceil$.

Proof: We first show that f_1 has no annihilator of degree $< \lceil \frac{n}{2} \rceil$. Suppose f_1 has a nonzero annihilator $g \in B_n$ having degree $< \lceil \frac{n}{2} \rceil$ of the form

$$a_0 + \sum_{i=0}^{n} a_i x_i + \cdots + \sum_{1 \le i_1 < \ldots < i_{\lceil \frac{n}{2} \rceil - 1} \le n} a_{i_1, \ldots, i_{\lceil \frac{n}{2} \rceil - 1}} x_{i_1} \cdots x_{i_{\lceil \frac{n}{2} \rceil - 1}},$$

where a's are in $\{0, 1\}$, but not all of them are zero. As g is an annihilator of f_1, $g(x_1, \ldots, x_n) = 0$ when $f_1(x_1, \ldots, x_n) = 1$. Hence solving the system of homogeneous linear equations (considering a's as the variables) formed by $g(x_1, \ldots, x_n) = 0$ when $f_1(x_1, \ldots, x_n) = 1$, we must get a nontrivial (not all zero) solution on a's.

Let us consider an input $(x_1, \ldots, x_n)$, where $x_{i_1}, \ldots, x_{i_t}$ are 1 ($t < \lceil \frac{n}{2} \rceil$) and the rest are 0 with $f_1(x_1, x_2, \ldots, x_n) = 1$. Then for this input, we have the homogeneous linear equation of the form $\sum_{I \subseteq \{i_1, \ldots, i_t\}} a_I = 0$, i.e., $a_{i_1, \ldots, i_t} = \sum_{I \subset \{i_1, \ldots, i_t\}} a_I$.

Since $f_1(0, \ldots, 0) = 1$, we must have $g(0, \ldots, 0) = 0$, i.e., $a_0 = 0$. As $f_1(x) = 1$ for $wt(x) = 1$, we have $a_i = a_0 = 0$. Following the same process we have all a's in g are 0. Thus g becomes a zero function, which is a contradiction as we have started with nonzero g. Thus f_1 has no annihilator of degree $< \frac{n}{2}$.

Now we show that f_2 has no annihilator of degree $< \lceil \frac{n}{2} \rceil$. Suppose f_2 has an annihilator h of degree $< \lceil \frac{n}{2} \rceil$. That is, $f_2(x_1, \cdots, x_n) * h(x_1, \cdots, x_n) = 0$. Note that $f_1(x_1, \cdots, x_n) = f_2(1 + x_1, \cdots, 1 + x_n)$, i.e., $f_2(x_1, \cdots, x_n) = f_1(1 + x_1, \cdots, 1 + x_n)$. Thus, $f_1(1+x_1, \cdots, 1+x_n) * h(x_1, \cdots, x_n) = 0$. Define h' as $h'(x_1, \cdots, x_n) = h(1+x_1, \cdots, 1+x_n)$, i.e., $h(x_1, \cdots, x_n) = h'(1 + x_1, \cdots, 1 + x_n)$. This gives $\deg(h') = \deg(h) < \lceil \frac{n}{2} \rceil$. Hence, we have $f_1(1+x_1, \cdots, 1+x_n) * h'(1+x_1, \cdots, 1+x_n) = 0$, i.e., $f_1(x_1, \ldots, x_n) * h'(x_1, \ldots, x_n) = 0$. So, f_1 has an annihilator of degree $< \lceil \frac{n}{2} \rceil$, which is a contradiction. ■

Thus we get the following theorem.

Theorem 1 *Let $f(x_1, \ldots, x_n) \in B_n$ constructed by Construction 2. Then $\mathcal{AI}_n(f) = \lceil \frac{n}{2} \rceil$.*

Proof: First we prove for odd n. Here $supp(1 + f) = supp(f_1)$ and $supp(f) = supp(f_2)$, where f_1, f_2 are as described in Lemma 3. Thus from Lemma 1 we have the proof for odd n. Now we will prove for n even. It can be checked that $supp(1 + f) \supseteq supp(f_1)$ and $supp(f) \supseteq supp(f_2)$, where f_1, f_2 are as described in Lemma 3. This, using Lemma 1, gives the proof for n even. ■

3 Algebraic Degree and Nonlinearity for a subcase (Symmetric Functions) of Construction 2

Given the function f in Construction 2, we can consider a special case where f is as follows.

Construction 3

$$\begin{aligned} f(x_1, \ldots, x_n) &= 0 \textit{ for } wt(x_1, \ldots, x_n) \le \lfloor \frac{n}{2} \rfloor, \\ &= 1 \textit{ for } wt(x_1, \ldots, x_n) > \lfloor \frac{n}{2} \rfloor. \end{aligned}$$

Note that in this case f is a symmetric Boolean function. A Boolean function is called symmetric if it outputs the same value for all the inputs of same weight. Thus it is clear that one can represent an n-variable symmetric Boolean function $f(x_1, \ldots, x_n)$ in a reduced form by $n+1$ bits string re_f such that $re_f(i) = f(x_1, \ldots, x_n)$ when $wt(x_1, \ldots, x_n) = i$. It is also clear that in the algebraic normal form, a symmetric Boolean function will either contain all the terms of the same degree monomial or none of them. Thus we can present the algebraic normal form in a reduced form by $n+1$ bits string ra_f such that $ra_f(i) = 1$, when all the i degree monomials are present and $ra_f(i) = 0$, when all the i degree monomials are absent. Thus for an n-variable symmetric Boolean function f, both re_f, ra_f can be seen as mappings from $\{0, 1, \ldots, n\}$ to $\{0, 1\}$.

Now we exactly calculate the algebraic degree, weight and nonlinearity of the functions in Construction 3.

3.1 Algebraic Degree

The relationship between re_f, ra_f have been presented in [22, Theorem 3] as

$$re_f(i) = (\sum_{k=0}^{i} ra_f(k) \binom{i}{k}) \bmod 2, \tag{1}$$

where $0 \le i \le n$. From [10, Page 85], for two integer sequences p, q,

$$p_i = \sum_{k=0}^{i} q_k \binom{i}{k} \text{ iff } q_i = \sum_{k=0}^{i} p_k (-1)^{i-k} \binom{i}{k}. \tag{2}$$

From Equation 1 and Equation 2 we get

Proposition 2 $ra_f(i) = (\sum_{k=0}^{i} re_f(k) \binom{i}{k}) \bmod 2.$

Proposition 3 *Suppose n and k are nonnegative integers with $n \ge k$. Let $n = 2^t + l$ where $0 \le l < 2^t$ and $t \ge 0$. Then we have*

1. *Let $k = 2^t + l_1$ where $l_1 \le l$. Then $\binom{n}{k}$ is even iff $\binom{l}{l_1}$ is even.*

2. *Let $k = 2^{t-1} + l_2$ where $l_2 < 2^{t-1}$. Then $\binom{n}{k}$ is even.*

Proof: Define $T(x) = a$ for any integer $x = 2^a b$ where b is an odd integer. It can be checked that $T((2^m)!) = \sum_{i=0}^{m-1} 2^i = 2^m - 1$. For $0 \le j < 2^m$, $T((2^m + j)!) = T((2^m)!) + T((2^m + 1)(2^m + 2) \cdots (2^m + j)) = 2^m - 1 + T(1 \cdot 2 \cdots j) = 2^m - 1 + T(j!)$.

For item 1, we have $k = 2^t + l_1$ where $l_1 \le l$. So, $T(\binom{n}{k}) = T(n!) - (T(k!) + T((n-k)!)) = T((2^t + l)!) - (T((2^t + l_1)!) + T((l - l_1)!)) = 2^t - 1 + T(l!) - (2^t - 1 + T(l_1!) + T((l - l_1)!)) = T(l!) - (T(l_1!) + T((l - l_1)!)) = T(\binom{l}{l_1})$.

For item 2, we have $k = 2^{t-1} + l_2$ where $l_2 < 2^{t-1}$. So, $T(\binom{n}{k}) = T((2^t + l)!) - (T((2^{t-1} + l_2)!) + T((2^{t-1} + l - l_2)!)) = 2^t - 1 + T(l!) - (2^{t-1} - 1 + T(l_2!) + 2^{t-1} - 1 + T((l - l_1)!)) = 1 + T(l!) - (T(l_1!) + T((l - l_1)!)) = 1 + T(\binom{l}{l_1}) \geq 1$. So, $\binom{n}{k}$ is even. ■

The following result provides the algebraic normal form and degree of f.

Theorem 2 *Let $f \in B_n$ a symmetric function as given in Construction 3. Then,*

1. *$ra_f(i) = 0$ for $i \leq \lfloor \frac{n}{2} \rfloor$,*
2. *$ra_f(\lfloor \frac{n}{2} \rfloor + 1) = 1$,*
3. *$ra_f(i) = \sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^{i} \binom{i}{k} \bmod 2$, for $i \geq \lfloor \frac{n}{2} \rfloor + 2$,*
4. *$\deg(f) = 2^{\lfloor \log_2 n \rfloor}$.*

Proof: Given the function f, it is clear that $re_f(i) = 0$ for $0 \leq i \leq \lfloor \frac{n}{2} \rfloor$ and $re_f(i) = 1$ for $\lfloor \frac{n}{2} \rfloor + 1 \leq i \leq n$. Thus from $ra_f(i) = (\sum_{k=0}^{i} re_f(k) \binom{i}{k}) \bmod 2$ (Proposition 2), we get $ra_f(i) = 0$ for $i \leq \lfloor \frac{n}{2} \rfloor$ and $ra_f(\lfloor \frac{n}{2} \rfloor + 1) = 1$. So we get the proofs of items 1 and 2.

The item 3 follows from Proposition 2 considering the result from item 1 and using $re_f(k) = 1$ for $k \geq \lfloor \frac{n}{2} \rfloor + 1$.

Suppose $t = \lfloor \log_2 n \rfloor$ and $l = n - 2^t$, i.e., $n = 2^t + l$ where $0 \leq l < 2^t$ and $t \geq 0$. For item 4 we need to show that $ra_f(i) = 1$ for $i = 2^t = 2^{\lfloor \log_2 n \rfloor}$ and $ra_f(i) = 0$ for all $i > 2^t$. Now $ra_f(i) = \sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^{i} \binom{i}{k} \bmod 2$. Here $n = 2^t + l$, i.e., $\lfloor \frac{n}{2} \rfloor + 1 = 2^{t-1} + \lfloor \frac{l}{2} \rfloor + 1$. Suppose $i = 2^t + l_1$ where $0 \leq l_1 \leq l$. So following the fact $\binom{i}{k} = 0 \bmod 2$ for $2^{t-1} \leq k < 2^t$ in Proposition 3 (item 2) we have $ra_f(i) = \sum_{k=2^t}^{i} \binom{i}{k} \bmod 2$. Then $ra_f(i) = \sum_{j=0}^{l_1} \binom{2^t + l_1}{2^t + j} \bmod 2$ as $i = 2^t + l_1$. Then following Proposition 3 (item 1) we have $ra_f(i) = \sum_{j=0}^{l_1} \binom{l_1}{j} \bmod 2 = 2^{l_1} \bmod 2$. Thus, $ra_f(2^t) = 1$ as $l_1 = 0$ and $ra_f(i) = 0$ for $i > 2^t$ as $l_1 > 0$. ■

3.2 Nonlinearity

In this section we will analyse the nonlinearity of the function f as explained in Construction 3. Nonlinearity is one of the most important cryptographic properties of Boolean functions which is used in cryptosystems to prevent linear attacks [17]. Moreover, this property is also very interesting from combinatorial point of view.

The nonlinearity of an n-variable function f, denoted as $nl(f)$, is the minimum distance from the set of all n-variable affine functions, i.e.,

$$nl(f) = \min_{g \in A(n)} (d(f, g)).$$

Walsh transform is a very useful tool in analysing Boolean functions. Let $x = (x_1, \ldots, x_n)$ and $\omega = (\omega_1, \ldots, \omega_n)$ both belonging to $\{0, 1\}^n$ and $x \cdot \omega = x_1\omega_1 + \ldots + x_n\omega_n$. Let $f(x)$ be a Boolean function on n variables. Then the *Walsh transform* of $f(x)$ is an integer valued function over $\{0,1\}^n$ which is defined as

$$W_f(\omega) = \sum_{x \in \{0,1\}^n} (-1)^{f(x)+x\cdot\omega}.$$

A Boolean function f is balanced iff $W_f(0) = 0$. The nonlinearity of f is given by

$$nl(f) = 2^{n-1} - \frac{1}{2} \max_{\omega \in \{0,1\}^n} |W_f(\omega)|.$$

As the function f explained in Theorem 2 is a symmetric Boolean function, we here concentrate on the Walsh spectra of this class. The Walsh spectra of symmetric Boolean functions have very nice combinatorial properties related to Krawtchouk polynomial [24].

Krawtchouk polynomial [20, Page 151, Part I] of degree i is given by

$$K_i(x,n) = \sum_{j=0}^{i} (-1)^j \binom{x}{j}\binom{n-x}{i-j}, \; i = 0, 1, \ldots, n. \tag{3}$$

It is known that for a fixed ω, such that $wt(\omega) = k$,

$$\sum_{wt(x)=i} (-1)^{\omega\cdot x} = K_i(k,n).$$

Thus it can be checked that if $f \in B_n$ is symmetric, then for $wt(\omega) = k$,

$$W_f(\omega) = \sum_{i=0}^{n} (-1)^{re_f(i)} K_i(k,n).$$

It is also known that for a symmetric function $f \in B_n$ and $\alpha, \beta \in \{0,1\}^n$, $W_f(\alpha) = W_f(\beta)$, if $wt(\alpha) = wt(\beta)$. Thus it is enough to calculate the Walsh spectra for the inputs of $n+1$ different weights. Keeping this in mind, given a symmetric Boolean function $f \in B_n$, we denote $rw_f(i) = W_f(\omega)$, such that $wt(\omega) = i$. Thus rw_f can be seen as a mapping from $\{0,\ldots,n\}$ to $\mathbb{Z}$.

Let us now list some known results in this area [18, 20].

Proposition 4

1. *$K_0(k,n) = 1, K_1(k,n) = n - 2k$,*
2. *$(i+1)K_{i+1}(k,n) = (n-2k)K_i(k,n) - (n-i+1)K_{i-1}(k,n)$,*
3. *$K_i(k,n) = (-1)^k K_{n-i}(k,n)$ (for n even and k odd, $K_{\frac{n}{2}}(k,n) = 0$),*
4. *$\binom{n}{k} K_i(k,n) = \binom{n}{i} K_k(i,n)$,*

5. *$K_i(k,n) = (-1)^i K_i(n-k,n)$, (for n even and i odd, $K_i(\frac{n}{2},n) = 0$),*

6. *$(n-k)K_i(k+1,n) = (n-2i)K_i(k,n) - kK_i(k-1,n)$,*

7. *$(n-i+1)K_i(k,n+1) = (3n-2i-2k+1)K_i(k,n) - 2(n-k)K_i(k,n-1)$.*

Proposition 5 *For n even, $K_i(\frac{n}{2},n) = \begin{cases} 0 \text{ for odd } i. \\ (-1)^{\frac{i}{2}}\binom{\frac{n}{2}}{\frac{i}{2}} \text{ for even } i. \end{cases}$*

Proof: For odd i, it is proved in Proposition 4. Now we will prove for even i using induction on i. For the base step, i.e., $i = 0$, we have $K_0(\frac{n}{2},n) = \binom{\frac{n}{2}}{0} = 1$. We will prove inductive step. Suppose it is true for $i = l$, i.e., $K_l(\frac{n}{2},n) = (-1)^{\frac{l}{2}}\binom{\frac{n}{2}}{\frac{l}{2}}$. Now we will prove for $i = l+2$. Following Proposition 4(item 2), we have $(l+2)K_{l+2}(\frac{n}{2},n) = -(n-l)K_l(\frac{n}{2},n)$ (in the proposition, we put $l+1$ instead of i). So, $K_{l+2}(\frac{n}{2},n) = (-1)^{\frac{l}{2}+1}\frac{n-l}{l+2}\binom{\frac{n}{2}}{\frac{l}{2}} = (-1)^{\frac{l}{2}+1}\binom{\frac{n}{2}}{\frac{l}{2}+1}$. Hence proved. ■

Let us now concentrate on the Walsh spectra of the symmetric function f as explained in Construction 3.

Lemma 4 *Consider the function f on n number of variables as given in Construction 3.*

1. *For k even, $rw_f(k) = \begin{cases} K_{\frac{n}{2}}(k,n) \text{ for even } n. \\ 0 \text{ for odd } n. \end{cases}$*

2. *For k odd, $rw_f(k) = 2\sum_{i=0}^{\lfloor\frac{n-1}{2}\rfloor} K_i(k,n)$.*

3. *$rw_f(1) = 2\binom{n-1}{\lfloor\frac{n}{2}\rfloor}$.*

4. *$rw_f(n) = \begin{cases} (-1)^{\frac{n}{2}}\binom{n}{\frac{n}{2}} \text{ for even } n. \\ (-1)^{\frac{n-1}{2}}2\binom{n-1}{\frac{n-1}{2}} \text{ for odd } n. \end{cases}$*

5. *For even n, $rw_f(\frac{n}{2}) = \begin{cases} (-1)^{\frac{n}{4}}\binom{\frac{n}{2}}{\frac{n}{4}} \text{ for even } \frac{n}{2}. \\ 2\sum_{i=0}^{\frac{n-2}{4}}(-1)^i\binom{\frac{n}{2}}{i} \text{ for odd } \frac{n}{2}. \end{cases}$*

Proof: From Proposition 4(3), we have $K_i(k,n) = (-1)^k K_{n-i}(k,n)$, i.e., if k is even, $K_i(k,n) = K_{n-i}(k,n)$. Now

$$rw_f(k) = \sum_{i=0}^{n}(-1)^{re_f(i)}K_i(k,n) = \sum_{i=0}^{\lfloor\frac{n}{2}\rfloor}K_i(k,n) - \sum_{i=\lfloor\frac{n}{2}\rfloor+1}^{n}K_i(k,n),$$

as

$$\begin{aligned} re_f(i) &= 0 \quad \text{for } 0 \le i \le \lfloor\tfrac{n}{2}\rfloor \text{ and} \\ &= 1 \quad \text{for } \lfloor\tfrac{n}{2}\rfloor < i \le n. \end{aligned}$$

Moreover, $\sum_{i=\lfloor \frac{n}{2}\rfloor+1}^{n} K_i(k,n) = \sum_{j=0}^{\lceil \frac{n}{2}\rceil-1} K_{j+\lfloor \frac{n}{2}\rfloor+1}(k,n) = \sum_{j=0}^{\lceil \frac{n}{2}\rceil-1} K_{n-j}(k,n) = \sum_{i=0}^{\lceil \frac{n}{2}\rceil-1} K_i(k,n) = \sum_{i=0}^{\lfloor \frac{n-1}{2}\rfloor} K_i(k,n)$. Hence, $rw_f(k) = K_{\frac{n}{2}}(k,n)$ for even n and $rw_f(k) = 0$ for odd n. This proves the first item.

For the second item, note that $K_i(k,n) = -K_{n-i}(k,n)$ as k is odd. Following the proof of item 1, we get $rw_f(k) = 2\sum_{i=0}^{\lfloor\frac{n-1}{2}\rfloor} K_i(k,n)$ (the even n and odd k case is handled under the same formula as $K_{\frac{n}{2}(k,n)} = 0$). So, we prove the second item.

For the third item, note that, $K_i(1,n) = \binom{n-1}{i} - \binom{n-1}{i-1}$. Thus, following item 2, $rw_f(1) = 2\sum_{i=0}^{\lceil\frac{n}{2}\rceil-1}(\binom{n-1}{i} - \binom{n-1}{i-1}) = 2\binom{n-1}{\lceil\frac{n}{2}\rceil-1}$. So, for odd n, $rw_f(1) = 2\binom{n-1}{\frac{n-1}{2}}$ and for even n, $rw_f(1) = 2\binom{n-1}{\frac{n}{2}-1} = 2\binom{n-1}{\frac{n}{2}}$. Therefore for any n, $rw_f(1) = 2\binom{n-1}{\lfloor\frac{n}{2}\rfloor}$.

For the fourth item, note that, $K_i(n,n) = (-1)^i K_i(0,n) = (-1)^i\binom{n}{i}$. For n even, following item 1, $rw_f(n) = K_{\frac{n}{2}}(n,n) = (-1)^{\frac{n}{2}} K_{\frac{n}{2}}(0,n) = (-1)^{\frac{n}{2}}\binom{n}{\frac{n}{2}}$. For odd n, following item 2, $rw_f(n) = 2\sum_{i=0}^{\frac{n-1}{2}}(-1)^i\binom{n}{i} = 2\sum_{i=0}^{\frac{n-1}{2}}(-1)^i(\binom{n-1}{i}+\binom{n-1}{i-1}) = \pm 2\binom{n-1}{\frac{n-1}{2}}$ (positive when $n = 1 \bmod 4$, negative when $n = 3 \bmod 4$).

For fifth item, following item 1 of this lemma and Proposition 5 the case $\frac{n}{2}$ even is proved. Similarly, following item 2 of this lemma and Proposition 5 the case $\frac{n}{2}$ odd is proved. ■

Lemma 5 *For* $1 \le k \le \lfloor\frac{n-1}{2}\rfloor$ *and* $0 \le i \le \lfloor\frac{n-1}{2}\rfloor$, $K_i(1,n) \ge |K_i(k,n)|$.

Proof: Note that, $K_i(1,n) = \binom{n-1}{i} - \binom{n-1}{i-1} \ge 0$ for $0 \le i \le \lfloor\frac{n-1}{2}\rfloor$ and that implies $|K_i(1,n)| = K_i(1,n)$ in $0 \le i \le \lfloor\frac{n-1}{2}\rfloor$.

First, we will prove it for $i \ge k$ using induction on k. In this direction for the base step we need to show $K_i(1,n) \ge |K_i(1,n)|$ (which is obvious) and $K_i(1,n) \ge |K_i(2,n)|$. Now $K_i(2,n) = \binom{n-2}{i} - 2\binom{n-2}{i-1} + \binom{n-2}{i-2}$ and $K_i(1,n) = \binom{n-1}{i} - \binom{n-1}{i-1} = \binom{n-2}{i} + \binom{n-2}{i-1} - \binom{n-2}{i-1} - \binom{n-2}{i-2} = \binom{n-2}{i} - \binom{n-2}{i-2}$. If $K_i(2,n) \ge 0$ then $K_i(1,n) - K_i(2,n) = 2(\binom{n-2}{i-1} - \binom{n-2}{i-2}) \ge 0$ as $(i-1) \le \lfloor\frac{n-2}{2}\rfloor$. If $K_i(2,n) \le 0$ then $K_i(1,n) + K_i(2,n) = 2(\binom{n-2}{i} - \binom{n-2}{i-1}) \ge 0$ for $i \le \lfloor\frac{n-2}{2}\rfloor$. Note that, $\lfloor\frac{n-1}{2}\rfloor = \lfloor\frac{n-2}{2}\rfloor$ when n is even and $\binom{n-2}{i} - \binom{n-2}{i-1} = 0$ for $i = \lfloor\frac{n-1}{2}\rfloor$ when n is odd. Therefore, $|K_i(1,n)| \ge |K_i(2,n)|$, i.e., $K_i(1,n) \ge |K_i(2,n)|$. Thus the base steps are proved.

Suppose for some $1 \le k < \lfloor\frac{n-1}{2}\rfloor$, $K_i(1,n) \ge |K_i(j,n)|$ for all j, $1 \le j \le k$. Now we will prove $K_i(1,n) \ge |K_i(k+1,n)|$. From Proposition 4(6), we have
$(n-k)K_i(k+1,n) = (n-2i)K_i(k,n) - kK_i(k-1,n)$,
i.e., $(n-k)|K_i(k+1,n)| \le (n-2i)|K_i(k,n)| + k|K_i(k-1,n)|$,
i.e., $(n-k)|K_i(k+1,n)| \le (n-2i)K_i(1,n) + kK_i(1,n)$,
i.e., $|K_i(k+1,n)| \le \frac{n-2i+k}{n-k}K_i(1,n)$,
i.e., $|K_i(k+1,n)| \le K_i(1,n)$, since $\frac{n-2i+k}{n-k} \le 1$ for $i \ge k$. So, the proof is completed for $j = k+1$. Hence, $K_i(1,n) \ge |K_i(k,n)|$ for $0 \le i \le \lfloor\frac{n-1}{2}\rfloor$, $1 \le k \le \lfloor\frac{n-1}{2}\rfloor$ and $i \ge k$.

Now we will prove for $0 \leq i < k \leq \lfloor \frac{n-1}{2} \rfloor$. Since $k > i$, following the above proof, we have $K_k(1,n) \geq |K_k(i,n)|$ by interchanging the role of k and i. Thus, $\binom{n}{i} K_k(1,n) \geq \binom{n}{i} |K_k(i,n)|$. Now following Proposition 4(4), we have $\binom{n}{i} K_k(1,n) \geq \binom{n}{k} |K_i(k,n)|$, i.e.,

$$\frac{\binom{n}{i}}{\binom{n}{k}} K_k(1,n) \geq |K_i(k,n)|. \tag{4}$$

Further, following Proposition 4(4), we have $K_k(1,n) = \frac{\binom{n}{k}}{\binom{n}{1}} K_1(k,n) = \frac{\binom{n}{k}}{n}(n-2k)$ and $K_i(1,n) = \frac{\binom{n}{i}}{n}(n-2i)$. So, $\frac{K_k(1,n)}{K_i(1,n)} = \frac{\binom{n}{k}(n-2k)}{\binom{n}{i}(n-2i)}$, i.e., $K_k(1,n) = \frac{\binom{n}{k}(n-2k)}{\binom{n}{i}(n-2i)} K_i(1,n)$. Now putting the value of $K_k(1,n)$ in Equation 4, we have $\frac{n-2k}{n-2i} K_i(1,n) \geq |K_i(k,n)|$, i.e., $K_i(1,n) \geq |K_i(k,n)|$, since $\frac{n-2k}{n-2i} < 1$ as $i < k$. Hence the proof. ■

In the next corollary we extend the range of i and k.

Corollary 1

1. *For odd n, $|K_i(1,n)| \geq |K_i(k,n)|$ where $0 \leq i \leq n$ and $1 \leq k \leq n-1$.*
2. *For even n, $|K_i(1,n)| \geq |K_i(k,n)|$ where $0 \leq i \leq n$ and $1 \leq k \leq n-1$ except $i = \frac{n}{2}$ or $k = \frac{n}{2}$.*

Proof: The proof for $0 \leq i \leq \lfloor \frac{n-1}{2} \rfloor$ and $1 \leq k \leq \lfloor \frac{n-1}{2} \rfloor$ is done in Lemma 5. The remaining part can be proved using the symmetry relations $K_i(k,n) = (-1)^k K_{n-i}(k,n)$ and $K_i(k,n) = (-1)^i K_i(n-k,n)$ in Proposition 4(item 3 and item 5). ■

When n is even the relation proved above is not true for $i = \frac{n}{2}$ and even k, since $K_{\frac{n}{2}}(1,n) = 0$ and $K_{\frac{n}{2}}(k,n)$ is a non zero number for even k.

Theorem 3 *Consider the functions $f \in B_n$, as explained in Construction 3. Then $nl(f) = 2^{n-1} - \binom{n-1}{\lfloor \frac{n}{2} \rfloor}$.*

Proof: First we prove that $rw_f(1)$ is maximum among all $rw_f(k)$ in $0 \leq k \leq n$.

Case 1. Let n be odd. First we show that $|rw_f(k)| \leq rw_f(1)$ for all k in the range $1 \leq k \leq n-1$. We know, $|rw_f(k)| = |2\sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} K_i(k,n)| \leq 2\sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} |K_i(k,n)|$. From Lemma 5 we have, $K_i(1,n) \geq |K_i(k,n)|$ for $1 \leq k \leq n-1$, and $0 \leq i \leq \lfloor \frac{n-1}{2} \rfloor$. This gives, $|rw_f(k)| \leq 2\sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} K_i(1,n) = rw_f(1)$. Again from Lemma 4 we have, $rw_f(1) = |rw_f(n)|$. Finally $rw_f(0) = 0$. Hence $rw_f(1) \geq |rw_f(k)|$ for $0 \leq k \leq n$.

Case 2. Let n be even. Let us first consider that k is odd and in $1 \leq k \leq n-1$ except $k = \frac{n}{2}$. From Lemma 4 we get that $|rw_f(k)| = |2\sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} K_i(k,n)|$. So following the same argument used in the previous case, we get $|rw_f(k)| \leq rw_f(1)$. For $k = \frac{n}{2}$ odd, from Lemma 4(item 5) we have $|rw_f(\frac{n}{2})| = |2\sum_{i=0}^{\frac{n-2}{4}} (-1)^i \binom{\frac{n}{2}}{i}| \leq 2\sum_{i=0}^{\frac{n-2}{4}} \binom{\frac{n}{2}}{i} = 2^{\frac{n}{2}}$. By induction on n it can be proved that $2^{\frac{n}{2}} \leq 2\binom{n-1}{\frac{n}{2}} = rw_f(1)$. So, for k odd and $1 \leq k \leq n-1$, the proof is done. When k even and $2 \leq k \leq n-2$, we have from Lemma 4 that $rw_f(k) = K_{\frac{n}{2}}(k,n)$. Now $K_{\frac{n}{2}}(k,n) = \sum_{j=0}^{\frac{n}{2}}(-1)^j \binom{k}{j}\binom{n-k}{\frac{n}{2}-j} \leq \sum_{j=0}^{\frac{n}{2}} \binom{k}{j}\binom{n-k}{\frac{n}{2}-j} = \binom{n}{\frac{n}{2}} = rw_f(1)$. Further, since $K_{\frac{n}{2}}(0,n) = \binom{n}{\frac{n}{2}} = |K_{\frac{n}{2}}(n,n)|$, we get, $rw_f(1) = rw_f(0) = |rw_f(n)|$. Thus $|rw_f(k)| \leq rw_f(1)$ for all k in $0 \leq k \leq n$.

So for any n, $nl(f) = 2^{n-1} - \frac{1}{2}|rw_f(1)| = 2^{n-1} - \binom{n-1}{\lfloor \frac{n}{2} \rfloor}$. ■

Now we would like to present a few observations.

1. We have checked for odd n up to $n = 11$, the function we have constructed in Construction 3, is the only function with maximum possible annihilator immunity among the symmetric functions. There is no other symmetric Boolean function on odd number of variables that are of annihilator immunity $\lceil \frac{n}{2} \rceil$ as far as we have experimented. This is an important open question to be proved or disproved.

2. For even n, we have found that there are symmetric functions with full annihilator immunity other than what we have presented in Construction 3. In fact so far we have experimented, up to $n = 12$, we found functions with full annihilator immunity $\frac{n}{2}$ and nonlinearity greater than that of the function constructed in Construction 3. In Table 2, we present the maximum nonlinearity available for symmetric Boolean functions on even number of variables having maximum possible annihilator immunity. This we found by computer search by writing computer program. It will be interesting to characterize the symmetric functions on even number of variables with maximum possible nonlinearity and maximum possible annihilator immunity $\frac{n}{2}$.

n	4	6	8	10	12
nonlinearity of Construction 3	5	22	93	386	1586
maximum nonlinearity (by exhaustive search)	6	26	94	394	1630

Table 1: Nonlinearity of symmetric Boolean functions on even number of variables by Construction 3 and maximum nonlinearity by exhaustive search.

4 Results comparing that of ϕ_{2k} in [15]

We have proved that the nonlinearity of these functions are same as the weight. Most interestingly they are also same with what observed (not proved) for the function ϕ_{2k} in [15] for $k = 1, \ldots, 8$. However, our functions can not always be linear transformation of ϕ_{2k} as the algebraic degree of our functions are different from that of ϕ_{2k} as available in Table 2.

Let us now concentrate on construction of balanced f with maximum possible annihilator immunity for even n. Refer to the general form of f as given in Construction 2. If b is so chosen that out of $\binom{n}{\frac{n}{2}}$ inputs, half of the corresponding outputs are 1 and the other half

$n=2k$	2	4	6	8	10	12	14	16
$\deg(\phi_{2k})$	2	4	5	8	9	11	13	16
$\deg(f)$	2	4	4	8	8	8	8	16

Table 2: Comparison of algebraic degree.

are 0, then f will be balanced. To formalize it, consider two sets $S_n, T_n \subset \{x | wt(x) = \frac{n}{2}\}$, $S_n \cap T_n = \emptyset$, $|S_n| = |T_n| = \frac{1}{2}\binom{n}{\frac{n}{2}}$. Note that there are $\binom{\binom{n}{\frac{n}{2}}}{\frac{1}{2}\binom{n}{\frac{n}{2}}} = \binom{\binom{n}{\frac{n}{2}}}{\binom{n-1}{\frac{n}{2}}}$ many different options to choose any S_n and correspondingly a T_n.

Now we have the following result.

Proposition 6 *Let F be an n-variable balanced function (n even) as follows.*

$$\begin{aligned} F(x_1,\ldots,x_n) &= 0 \textit{ for } wt(x_1,\ldots,x_n) < \frac{n}{2}, \\ &= 1 \textit{ for } wt(x_1,\ldots,x_n) > \frac{n}{2}, \\ &= 0 \textit{ for } (x_1,\ldots,x_n) \in S_n, \\ &= 1 \textit{ for } (x_1,\ldots,x_n) \in T_n. \end{aligned}$$

Then $nl(F) \geq 2^{n-1} - \binom{n}{\frac{n}{2}}$.

Proof: Consider the function f in Construction 3. It is clear that $\frac{1}{2}\binom{n}{\frac{n}{2}}$ many output points in the truth table of f need to be toggled to get the function F. Thus $nl(F) \geq nl(f) - \frac{1}{2}\binom{n}{\frac{n}{2}}$. From Theorem 3, $nl(f) = 2^{n-1} - \binom{n-1}{\frac{n}{2}}$. Thus $nl(F) \geq 2^{n-1} - \binom{n-1}{\frac{n}{2}} - \frac{1}{2}\binom{n}{\frac{n}{2}} = 2^{n-1} - \frac{1}{2}\binom{n}{\frac{n}{2}} - \frac{1}{2}\binom{n}{\frac{n}{2}} = 2^{n-1} - \binom{n}{\frac{n}{2}}$. ■

However, we now show a heuristic construction with which we can really get much better value of nonlinearity of the balanced functions. Note that we do not present any theoretical proof here, but only list the experimental results.

For that we first refer to Maiorana-McFarland type of bent functions. The Maiorana-McFarland class of bent function is as follows [16]. Consider n-variable Boolean functions on (X, Y), where $X, Y \in \{0,1\}^{\frac{n}{2}}$ of the form $f(X, Y) = X \cdot \pi(Y) + g(Y)$ where π is a permutation on $\{0,1\}^{\frac{n}{2}}$ and g is any Boolean function on $\frac{n}{2}$ variables. The function f can be seen as concatenation of $2^{\frac{n}{2}}$ distinct (up to complementation) affine function on $\frac{n}{2}$ variables. For our purpose we consider π as an identity permutation, g as a constant zero function and refer to this function on n variables as $b(x_1,\ldots,x_n)$, for n even. Now we construct an n-variable function G as follows.

$$\begin{aligned} G(x_1,\ldots,x_n) &= 0 \text{ for } wt(x_1,\ldots,x_n) < \frac{n}{2}, \\ &= 1 \text{ for } wt(x_1,\ldots,x_n) > \frac{n}{2}, \\ &= b(x_1,\ldots,x_n) \text{ for } wt(x_1,\ldots,x_n) = \frac{n}{2}. \end{aligned}$$

Experimentally we observe that $nl(G) = nl(f)$, for even n up to 16, where f is the function as described in Construction 3. Note that G is much closer to balancedness than the function f.

1. If $wt(G) < 2^{n-1}$, then we choose $2^{n-1} - wt(G)$ points randomly from the inputs having weight $\frac{n}{2}$ and output 0 of G and toggle those outputs to 1.

2. If $wt(G) > 2^{n-1}$, then we choose $wt(G) - 2^{n-1}$ points randomly from the inputs having weight $\frac{n}{2}$ and output 1 of G and toggle those outputs to 0.

After this change G will become balanced. Experimentally we get the following result for the function G in Table 3. We execute 100 runs for each n and take the best result among the runs in terms of nonlinearity. We also observe that algebraic degree of the reported functions is the maximum possible, i.e., $n - 1$.

$n = 2k$	4	6	8	10	12	14	16
$2^{n-1} - \binom{n-1}{\frac{n}{2}}$	5	22	93	386	1586	6476	26333
$nl(G)$	4	22	92	384	1582	6468	26316
$4(2^{n-3} - \binom{n-3}{\frac{n-2}{2}})$	4	20	88	372	1544	6344	25904

Table 3: Comparison of nonlinearities.

We have also checked that G is always the maximum possible, i.e., $n-1$, for a balanced function.

As by itself the function ϕ_{2k} was not balanced, the construction of balanced function that has been mentioned in [15] with full annihilator immunity is basically $x_1 + x_2 + \phi_{2k-2}$, where ϕ_{2k-2} was on the variables $x_3, \ldots, x_{2k}$. The nonlinearity of this function is $4\,nl(\phi_{2k-2}) = 4(2^{n-3} - \binom{n-3}{\frac{n-2}{2}})$. That is also presented in the last line of Table 3. Clearly our heuristic construction presents better nonlinearity than the balanced functions presented in [15].

5 Computing Walsh spectra of Symmetric Boolean Functions

Here we present an algorithm to calculate rw_f from re_f for a symmetric function $f \in B_n$. Note that in [21, Page 33] it has been mentioned that calculating the Walsh spectra for an n-variable symmetric function requires $O(n^3)$ time and $O(n^2)$ space. In that case $\sum_{wt(x)=i}(-1)^{\omega \cdot x} = K_i(k, n)$, has been stored in the (i, k)-th location of an $(n + 1) \times (n + 1)$ integer matrix ($O(n^2)$ space) and getting the value of each location required $O(n)$ operations. Thus $O(n^3)$ time is spent. Then for each weight k the value of $rw_f(k)$ is calculated in $O(n)$ time and this is again done for $(n + 1)$ different weights $[0, \ldots, n]$. This takes additional $O(n^2)$ steps. However, we here show that using the properties of

Krawtchouk polynomial [18,20] this can be done in $O(n^2)$ time and $O(n)$ space. The basic idea is as follows:

1. (a) At the same step, once we get $K_i(k,n)$ we can calculate $K_{n-i}(k,n)$ using Proposition 4(3). Thus in the calculation of $rw_f(k)$, we can add these two values at the same time, i.e., we get $(-1)^{re_f(i)}K_i(k,n) + (-1)^{re_f(n-i)}K_{n-i}(k,n)$. To get the complete value of $rw_f(k)$, we need to apply this for $i = 0$ to $\frac{n-1}{2}$ for n odd. If n is even, one more step is required where the value $(-1)^{re_f(\frac{n}{2})}K_{\frac{n}{2}}(k,n)$ will also be added.

   (b) At the same step, once we get $K_i(k,n)$ we can calculate $K_i(n-k,n)$ using Proposition 4(5) and then $K_{n-i}(n-k,n)$ using Proposition 4(3). Thus in the calculation of $rw_f(n-k)$, we can add these two values at the same time, i.e., we get $(-1)^{re_f(i)}K_i(n-k,n) + (-1)^{re_f(n-i)}K_{n-i}(n-k,n)$. To get the complete value of $rw_f(n-k)$, we need to apply this for $i = 0$ to $\frac{n-1}{2}$ for n odd. If n is even, one more step is required where the value $(-1)^{re_f(\frac{n}{2})}K_{\frac{n}{2}}(n-k,n)$ will also be added.

   Thus at the same time $rw_f(k), rw_f(n-k)$ are calculated for $0 \leq k \leq \frac{n-1}{2}$, when n is odd. If n is even, we need to calculate $rw_f(\frac{n}{2})$ separately. Thus if $K_i(k,n)$ values are available in constant time (see below), then calculation of complete Walsh spectra requires $O(n^2)$ time.

2. From Proposition 4(1), we get $K_0(k,n) = 1, K_1(k,n) = n - 2k$ as the initial values. Then given $K_{i-1}(k,n)$ and $K_i(k,n)$, it is possible to get $K_{i+1}(k,n)$ by Proposition 4(2). Thus, just by storing two old values and keeping one temporary variable, it is possible to get $K_i(k,n)$ for each i in constant time.

   Moreover, it is clear that apart from storing $(n+1)$ Walsh spectra value, the number of other variables to be used are constant. Thus the space complexity is $O(n)$.

The exact C program like algorithm (Algorithm 1) is presented below.

Algorithm 1 *Algorithm to calculate the Walsh spectra of a Symmetric Boolean function.*

```
input: number of variables $n$, symmetric function $re_f$;
output: Walsh spectra $rw_f$;

for ($k = 0$ to $\lfloor \frac{n-1}{2} \rfloor$){
    $v_1 = (-1)^{re_f(0)} + (-1)^{re_f(n)+k}$;
    $v_2 = (-1)^{re_f(0)} + (-1)^{re_f(n)+n-k}$;
    $p = n - 2k, q = 1$;
    for ($i = 1$ to $\lfloor \frac{n-1}{2} \rfloor$){
        $v_1 = v_1 + ((-1)^{re_f(i)} + (-1)^{re_f(n-i)+k})p$;
        $v_2 = v_2 + ((-1)^{re_f(i)+i} + (-1)^{re_f(n-i)+i+n-k})p$;
        $r = \frac{(n-2k)p-(n-i+1)q}{i+1}$;
        $q = p, p = r$;
    }
    if $n$ is even{
        $i = \frac{n}{2}$;
        $v_1 = v_1 + (-1)^{re_f(i)}p$;
        $v_2 = v_2 + (-1)^{re_f(i)+i}p$;
    }
    $rw_f(k) = v_1, rw_f(n-k) = v_2$;
}
if ($n$ is even){
    $k = \frac{n}{2}$;
    $v_1 = (-1)^{re_f(0)} + (-1)^{re_f(n)+k}$;
    $p = n - 2k, q = 1$;
    for ($i = 1$ to $\lfloor \frac{n-1}{2} \rfloor$){
        $v_1 = v_1 + ((-1)^{re_f(i)} + (-1)^{re_f(n-i)+k})p$;
        $r = \frac{(n-2k)p-(n-i+1)q}{i+1}$;
        $q = p, p = r$;
    }
    $i = \frac{n}{2}$;
    $v_1 = v_1 + (-1)^{re_f(i)}p$;
    $rw_f(k) = v_1$;
}
```

6 Conclusion

In this paper we could identify the basic theory towards the construction of Boolean functions with full annihilator immunity. Based on the theory we present some concrete construction ideas. Further we could study the other cryptographic properties like nonlinearity and algebraic degree theoretically. Our work compares favourably than what has been presented in a recent paper [15].

Examples are now available [14, Section 4.1] that there exist Boolean functions having optimum parameters in terms of different cryptographic properties such as balancedness, nonlinearity, algebraic degree, annihilator immunity and correlation immunity. However, there is no such constructions yet in that direction. The existing constructions, that achieve optimization in terms of the parameters balancedness, nonlinearity, algebraic degree, and correlation immunity, do not provide maximum possible annihilator immunity. This is an important open area of research.

References

[1] F. Armknecht. Improving Fast Algebraic Attacks. In *FSE 2004*, number 3017 in Lecture Notes in Computer Science, pages 65–82. Springer Verlag, 2004.

[2] L. M. Batten. Algebraic Attacks over GF(q). In *Progress in Cryptology - INDOCRYPT 2004*, pages 84–91, number 3348, Lecture Notes in Computer Science, Springer-Verlag.

[3] A. Botev. On algebraic immunity of some recursively given sequence of correlation immune functions. In Proceedings of *XV international workshop on Synthesis and complexity of control systems*, Novosibirsk, October 18-23, 2004, pages 8-12 (in Russian).

[4] A. Botev. On algebraic immunity of new constructions of filters with high nonlinearity. In Proceedings of *VI international conference on Discrete models in the theory of control systems*, Moscow, December 7-11, 2004, pages 227-230 (in Russian).

[5] A. Botev and Y. Tarannikov. Lower bounds on algebraic immunity for recursive constructions of nonlinear filters. Preprint 2004.

[6] A. Canteaut. Open problems related to algebraic attacks on stream ciphers. In *WCC 2005*, pages 1–10, invited talk.

[7] C. Carlet. Improving the algebraic immunity of resilient and nonlinear functions and constructing bent functions. IACR ePrint server, http://eprint.iacr.org, 2004/276.

[8] J. H. Cheon and D. H. Lee. Resistance of S-boxes against Algebraic Attacks. In *FSE 2004*, number 3017 in Lecture Notes in Computer Science, pages 83–94. Springer Verlag, 2004.

[9] J. Y. Cho and J. Pieprzyk. Algebraic Attacks on SOBER-t32 and SOBER-128. In *FSE 2004*, number 3017 in Lecture Notes in Computer Science, pages 49–64. Springer Verlag, 2004.

[10] G. M. Constantine. Combinatorial Theory and Statistical Design. John Wiley & Sons, 1987.

[11] N. Courtois and J. Pieprzyk. Cryptanalysis of block ciphers with overdefined systems of equations. In *Advances in Cryptology - ASIACRYPT 2002*, number 2501 in Lecture Notes in Computer Science, pages 267–287. Springer Verlag, 2002.

[12] N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. In *Advances in Cryptology - EUROCRYPT 2003*, number 2656 in Lecture Notes in Computer Science, pages 345–359. Springer Verlag, 2003.

[13] N. Courtois. Fast algebraic attacks on stream ciphers with linear feedback. In *Advances in Cryptology - CRYPTO 2003*, number 2729 in Lecture Notes in Computer Science, pages 176–194. Springer Verlag, 2003.

[14] D. K. Dalai, K. C. Gupta and S. Maitra. Results on Algebraic Immunity for Cryptographically Significant Boolean Functions. In *INDOCRYPT 2004*, pages 92–106, number 3348, Lecture Notes in Computer Science, Springer-Verlag.

[15] D. K. Dalai, K. C. Gupta and S. Maitra. Cryptographically Significant Boolean functions: Construction and Analysis in terms of Algebraic Immunity. In *FSE 2005*. To be published in Lecture Notes in Computer Science, Springer-Verlag.

[16] J. F. Dillon. Elementary Hadamard Difference sets. PhD Thesis, University of Maryland, 1974.

[17] C. Ding, G. Xiao, and W. Shan. *The Stability Theory of Stream Ciphers*. Number 561 in Lecture Notes in Computer Science. Springer-Verlag, 1991.

[18] I. Krasikov. On integral zeros of Krawtchouk polynomials. *Journal of Combinatorial Theory, Series A*, 74:71–99, 1996.

[19] D. H. Lee, J. Kim, J. Hong, J. W. Han and D. Moon. Algebraic Attacks on Summation Generators. In *FSE 2004*, number 3017 in Lecture Notes in Computer Science, pages 34–48. Springer Verlag, 2004.

[20] F. J. MacWillams and N. J. A. Sloane. *The Theory of Error Correcting Codes*. North Holland, 1977.

[21] S. Maitra. Boolean functions with important cryptographic properties. PhD Thesis, Indian Statistical Institute, 2000.

[22] S. Maitra and P. Sarkar. Maximum Nonlinearity of Symmetric Boolean Functions on Odd Number of Variables. *IEEE Transactions on Information Theory*, 48(9):2626–2630, September 2002.

[23] W. Meier, E. Pasalic and C. Carlet. Algebraic attacks and decomposition of Boolean functions. In *Advances in Cryptology - EUROCRYPT 2004*, number 3027 in Lecture Notes in Computer Science, pages 474–491. Springer Verlag, 2004.

[24] P. Savicky. On the bent Boolean functions that are symmetric. *European Journal of Combinatorics*, 15:407–410, 1994.