Balancedness and Correlation Immunity of Symmetric Boolean Functions

Palash Sarkar & Subhamoy Maitra

Applied Statistics Unit,
Indian Statistical Institute
203, B.T. Road, Calcutta 700 108, INDIA

Abstract

New subsets of symmetric balanced and symmetric correlation immune functions are identified. The method involves interesting relations on binomial coefficients and highlights the combinatorial richness of these classes. As a consequence of our constructive techniques, we improve upon the existing lower bounds on the cardinality of the above sets. We consider higher order correlation immune functions and show how to construct n-variable, 3rd order correlation immune function for each perfect square $n \geq 9$.

Keywords : *Symmetric Boolean Function, Balancedness, Correlation Immunity.*

1 Introduction

An interesting subclass of Boolean functions is the set of symmetric functions. The study of balanced symmetric functions and correlation immune symmetric functions was made by Brüer [1], Mitchell [5] and later by Yang and Guo [11]. Independently Chor et al [2] and later Gopalakrishnan et al [4] studied symmetric functions possessing both the properties of balancedness and correlation immunity. Following Mitchell [5], we provide definitions of the relevant Boolean function properties. We will use $\oplus$ to denote addition modulo 2.

Definition 1.1 *Let $f(X_n, \ldots, X_1)$ be a Boolean function.*
C1. Balancedness. *The function f is balanced if the number of ones in its output column is equal to the number of zeros.*
C2. Nonaffinity. *The function f is affine if it can be written as $f(X_n, \ldots, X_1) = \oplus_{i=1}^{i=n} a_i X_i \oplus b$, where $a_i, b \in \{0, 1\}$. If $b = 0$, the function f is called linear. The function f is nonaffine if it is not affine.*

⋆ The full version of the paper can be found at [6].

Email address: {palash, subho}@isical.ac.in (Palash Sarkar & Subhamoy Maitra).

C3. Nondegeneracy. *The function f is degenerate on variable X_i if $f(X_n, \ldots, X_{i+1}, X_i = 0, X_{i-1}, \ldots, X_1) = f(X_n, \ldots, X_{i+1}, X_i = 1, X_{i-1}, \ldots, X_1)$. The function f is nondegenerate if it is not degenerate on any variable.*
C4. Correlation Immunity. *The function f is correlation immune (CI) if $Prob(f = X_i) = \frac{1}{2}$ for all $1 \leq i \leq n$.*
C5. Symmetry. *The function f is symmetric if $f(X_n, \ldots, X_1)$ is the same for all the vectors $(X_n, \ldots, X_1)$ of same weight.*

$A_n(i_1, \ldots, i_t)$ is the set of all n-variable Boolean functions having the properties $Ci_1, \ldots, Ci_t$.

In Sections 2 and 3 we provide construction of new functions in the sets $A_n(1,2,3,5)$ and $A_n(2,3,4,5)$ respectively. These are used to improve known lower bounds on the sizes of such sets. Our constructions explain the "sporadic" examples in $A_n(1,2,3,5)$ reported by Brüer [1] and Mitchell [5]. In Section 4 we present a method to construct 3rd order correlation immune functions and compute the algebraic degree of some of these functions.

Let $wt(s)$ denote the Hamming weight of a binary string s. For a symmetric Boolean function all input vectors with the same weight have the same output value. Based on this observation, we define $WTS(f)$ for a symmetric function f as $WTS(f) = \{i : wt(X_n \ldots X_1) = i \text{ implies } f(X_n, \ldots, X_1) = 1\}$. The weight of a Boolean function f is $wt(f) = |\{(X_n, \ldots, X_1) : f(X_n, \ldots, X_1) = 1\}|$. If f is symmetric then $wt(f) = \sum_{i \in WTS(f)} \binom{n}{i}$.

We first state some binomial coefficient identities. These will be interpreted in terms of symmetric functions in later sections to provide constructions of balanced and correlation immune symmetric functions.

Proposition 1.1 *Let $n > 0$ and $1 \leq r \leq n$ be positive integers. Then (1) $3r = n + 1$ if and only if $2\binom{n}{r-1} = \binom{n}{r}$; (2) $(n - 2r)^2 = n + 2$ if and only if $2\binom{n}{r} = \binom{n}{r+1} + \binom{n}{r-1}$ [4]; (3) $(n - 2r - 1)^2 = n + 3$ if and only if $\binom{n}{r-1} + \binom{n}{r+2} = \binom{n}{r} + \binom{n}{r+1}$.*

2 Balancedness

In [1; 5], the problem of enumerating $A_n(1,5)$ is discussed, where a lower bound on the number of balanced symmetric functions is obtained. A simple way to obtain balanced symmetric functions is provided in [5]. Let f, g be symmetric functions such that $WTS(f) = \{i : i \text{ even}\}$ and $WTS(g) = \{i : i \text{ odd}\}$. From properties of binomial coefficients both f and g are balanced. Also these are the two nondegenerate n-variable affine functions.

Further, if n is odd, one can form additional balanced functions in the following

way. Since n is odd, for $1 \leq i \leq n$, we have that i is odd if and only if $n-i$ is even. Let $P_i = \{i, n-i\}$. We form a set S by choosing exactly one element from each P_i. Clearly $\sum_{i \in S} \binom{n}{i} = \sum_{i \notin S} \binom{n}{n-i} = 2^{n-1}$. Thus the function f such that $WTS(f) = S$ is balanced. From the construction it is clear that there are $2^{\frac{n+1}{2}}$ such possible functions which also includes the two nondegenerate affine functions. We will call these ways of partitioning to be trivial. These partitionings immediately give rise to the lower bound $\mid A_n(1,5) \mid \geq 2^{\frac{n+1}{2}}$ if n is odd, and ≥ 2 if n is even.

The inequality is strict when some nontrivial partitioning is found. Brüer [1] tabulates $\mid A_n(1,5) \mid$ for odd $n \leq 17$ and obtains $\mid A_n(1,5) \mid = 2^{\frac{n+1}{2}}$ except for $\mid A_{13}(1,5) \mid = 144$. Mitchell [5] has also shown that $\mid A_8(1,5) \mid > 2$ and termed these as "sporadic" examples. We show that these are not sporadic and there exist infinitely many integer values of n for which we get strict inequality.

Theorem 2.1 *1. Let $n \equiv 2 \bmod 6$. Then it is possible to construct $f \in A_n(1,2,3,5)$. Consequently, $\mid A_n(1,5) \mid > 2$.*
2. Let $n \geq 14$ be an even integer such that $n+2$ is a perfect square. Then it is possible to construct functions in $A_n(1,2,3,5)$. Consequently, $|A_n(1,5)| > 2$.
3. Let $n \geq 13$ be odd and $(n+3)$ a perfect square. Then $\mid A_n(1,5) \mid \geq 2^{\frac{n+1}{2}} + 2^{\frac{n+1}{2}-3}$.

Theorem 2.1 explains the sporadic examples obtained by Mitchell [5] for $n = 8$. For $n = 13$, Theorem 2.1 provides $|A_{13}(1,5)| \geq 144$. In fact, $|A_{13}(1,5)| = 144$ as observed by Brüer [1].

3 Correlation Immunity

Here we consider the construction problem for the set of symmetric correlation immune functions. The following is a characterization of correlation immunity for symmetric functions.

Theorem 3.1 *Let $f \in A_n(5)$ with $WTS(f) = \{i_1, \ldots, i_r\}$. Then f is CI iff $\binom{n-1}{i_1} + \ldots + \binom{n-1}{i_r} = \binom{n-1}{i_1-1} + \ldots + \binom{n-1}{i_r-1}$.*

A consequence of Theorem 3.1 is the following fact: *Let f and f' be such that $k, n-k \notin WTS(f)$ and $WTS(f') = WTS(f) \cup \{k, n-k\}$. Then f is CI if and only if f' is CI.* A Boolean function f is said to be *palindromic* if for each n-bit vector $(b_n, \ldots, b_n)$, we have $f(b_n, \ldots, b_1) = f(1 \oplus b_n, \ldots, 1 \oplus b_1)$.

Proposition 3.1 *A symmetric function f is palindromic if and only if for each i, $WTS(f)$ contains either both i and $n-i$ or none of them.*

The importance of Proposition 3.1 lies in the fact that any palindromic Boolean function is CI [5]. The number of symmetric palindromic functions is clearly $2^{\lfloor \frac{n}{2} \rfloor + 1}$ (see Theorem 8 of [11]). Thus it is of interest to find nonpalindromic CI functions. We provide such constructions in this section.

Theorem 3.2 *1. Take n, r, i such that $2\binom{n-1}{r} = \binom{n-1}{r-i} + \binom{n-1}{r+i}$, $i \geq 1$. Then one can construct nonpalindromic $f \in A_n(4,5)$.*
2. Let $n+1$ be a perfect square and $n \geq 8$. Then $\mid A_n(2,3,4,5) \mid \geq 2^{\lfloor \frac{n}{2} \rfloor + 1} + 2^{\lceil \frac{n-1}{2} \rceil} - 2$.
3. Let $n+2$ be a perfect square and $n \geq 14$. Then $\mid A_n(2,3,4,5) \mid \geq 2^{\lfloor \frac{n}{2} \rfloor + 1} + 2^{\lceil \frac{n-1}{2} \rceil} - 2$.
4. Let $n+3$ be a perfect square and $n \geq 13$. Then $\mid A_n(2,3,4,5) \mid \geq 2^{\lfloor \frac{n}{2} \rfloor + 1} + 2^{\lceil \frac{n-5}{2} \rceil} - 2$.
5. Take n, r, i such that $2\binom{n-1}{r} = \binom{n-1}{r-i-1} + \binom{n-1}{r+i}$, $i \geq 1$. Then there exists nonpalindromic $f \in A_n(4,5)$.

It is interesting to note that $\mid A_n(4,5) \mid = 2^{\lfloor \frac{n}{2} \rfloor + 1} + 2(n \bmod 2)$ for $n = 4, 5, 10, 11, 17, 28$. However, n does not appear to follow any obvious pattern for this exact equality condition.

4 Higher Order Correlation Immunity

The class of correlation immune functions was introduced by Siegenthaler [8]. In the introduction we mentioned only the special case of first order CI functions as considered in Mitchell [5]. In this section we consider the general class of CI functions and present new constructions of 3rd order CI functions.

Definition 4.1 *A Boolean function $f(X_n, \ldots, X_1)$ is said to be correlation immune of order m (m-CI for short), if*
$Prob(f(X_n, \ldots, X_1) = 1 \mid Y_t = c_t, \ldots, Y_1 = c_1) = Prob(f(X_n, \ldots, X_1) = 1)$, where the variables $Y_t, \ldots, Y_1$ are chosen from $\{X_n, \ldots, X_1\}$, $c_t, \ldots, c_1 \in \{0,1\}$ and $1 \leq t \leq m$. A balanced m-CI function is called m-resilient.

Construction of 1-resilient and 2-resilient symmetric functions were presented in [4]. In section 3, we presented new constructions of 1-CI symmetric functions. Here we present a new construction of 3-CI functions.

A consequence of [7, Theorem 3.1] is the following result.

Theorem 4.1 *A symmetric function $f(X_n, \ldots, X_1)$ is m-CI if and only if for each t, $1 \leq t \leq m$, $wt(f_0) = \ldots = wt(f_{2^t-1})$ where for $0 \leq k \leq 2^t - 1$, $f_k(X_{n-t}, \ldots, X_1) = f(X_n = k_t, \ldots, X_{n-t+1} = k_1, X_{n-t}, \ldots, X_1)$ and $k_t \ldots k_1$ is the t-bit binary representation of k.*

For $0 \leq k \leq 2^t - 1$, by $wt(k)$ we will denote the weight of the t-bit binary representation of k.

Lemma 4.1 *Let f be an n-variable symmetric function and $1 \leq t \leq n-1$. Define $f_k(X_{n-t}, \ldots, X_1)$ as in Theorem 4.1. Then $WTS(f_k) = \{i - wt(k) : i \in WTS(f), 0 \leq i - wt(k) \leq n - t\}$.*

Lemma 4.2 *Let $f(X_n, \ldots, X_1)$ be a symmetric function with $WTS(f) = \{r, n-r\}$. For $0 \leq k \leq 2^t - 1$, define f_k as in Theorem 4.1. For $0 \leq i, j \leq 2^t - 1$, if $wt(i) + wt(j) = t$, then $wt(f_i) = wt(f_j)$.*

Above proof follows from Lemma 4.1. We use Theorem 4.1 and Lemma 4.2 to obtain the following result on 3-CI functions.

Theorem 4.2 *Let n and r be such that $(n-2r)^2 = n$. Then $f \in A_n(5)$ having $WTS(f) = \{r, n-r\}$ is 3-CI.*

Siegenthaler [8] showed that the maximum possible degree of an n-variable, m-resilient function is $n - m$. Next we compute the degrees of the functions described in Theorem 4.2. We present them in the form (n, deg) which are $(9, 6), (16, 11), (25, 14), (36, 29), (49, 30), (64, 55), (81, 62), (100, 61), (121, 118)$.

The maximum degree is obtained only for $n = 9, 121$. In fact, we were unable to find any other n, such that the function of Theorem 4.2 has degree $n - 3$. Also it is interesting to note that the degree of the function for $n = 100$ is less than the degree of the function for $n = 81$. Though this is a rare phenomenon, this also happens for other values of n. A good explanation of the behaviour of the degree seems elusive.

If a function f is m-CI, then a function g obtained by setting any input of f to constant is $(m-1)$-CI. Thus Theorem 4.2 also shows the existence of 2-CI functions. Earlier existence of 2-resilient functions were shown in [4]. In our computer experiments we did not find any 4-CI function for $6 \leq n \leq 20$. Further all the 3-CI functions obtained were palindromic. We give a few examples of 3-CI functions not covered by Theorem 4.2. These are written in the form $(n, WTS(f))$ as $(8, \{2, 3, 5, 6\}), (10, \{1, 3, 4, 6, 7, 9\}), (14, \{2, 3, 5, 6, 8, 9, 11, 12\}), (15, \{5, 6, 9, 10\}), (15, \{3, 6, 9, 12\}), (16, \{1, 3, 5, 6, 7, 8, 9, 10, 11, 13, 14\}), (16, \{1, 3, 4, 7, 9, 12, 13, 15\})$, and $(16, \{1, 3, 4, 6, 7, 9, 10, 12, 13, 15\})$. Some of the examples can perhaps be explained along the lines of Theorem 4.2. These form tasks of future research problems.

References

[1] J. O. Brüer. On pseudorandom sequences as crypto generators. In *International Zurich Seminar on Digital Communications*, pages 157–161. IEEE,

New York, 1984.

[2] B. Chor, O. Goldreich, J. Hastad, J. Friedman, S. Rudich, and R. Smolensky. The bit extraction problem or t-resilient functions. In *26th IEEE Symposium on Foundations of Computer Science*, pages 396–407, 1985.

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

[4] K. Gopalakrisnan, D. G. Hoffman, and D. R. Stinson. A note on a conjecture concerning symmetric resilient functions. *Information Processing Letters*, 47(3):139–143, 1993.

[5] C. J. Mitchell. Enumerating Boolean functions of cryptographic significance. *Journal of Cryptology*, 2(3):155–170, 1990.

[6] P. Sarkar and S. Maitra. Balancedness and correlation immunity of symmetric Boolean functions. http://www.isical.ac.in/~crg, Technical Report No. CRG/2002/009.

[7] P. Sarkar. A note on the spectral characterization of correlation immune Boolean functions. *Information Processing Letters*, 74:191–195, 2000.

[8] T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information Theory*, IT-30(5):776–780, September 1984.

[9] D. R. Stinson. Resilient functions and large sets of orthogonal arrays. *Congressus Numerantium*, 92(1993), 105-110.

[10] S. Wolfram. *The Mathematica Book, Mathematica Version 3.* Wolfram Media/Cambridge University Press, 1996.

[11] Y. X. Yang and B. Guo. Further enumerating Boolean functions of cryptographic significance. *Journal of Cryptology*, 8(3):115–122, 1995.