

Certificate complexity and symmetry of nested canalizing functions

Yuan Li¹Frank Ingram²Huaming Zhang³¹ Department of Mathematics, Winston-Salem State University, USA² Department of Mathematics, Winston-Salem State University, USA³ Computer Science Department, The University of Alabama in Huntsville, USA

received 2020-03-10, revised 2020-07-29, 2021-02-23, 2021-08-17, accepted 2021-10-07.

Boolean nested canalizing functions (NCFs) have important applications in molecular regulatory networks, engineering and computer science. In this paper, we study their certificate complexity. For both Boolean values $b \in \{0, 1\}$, we obtain a formula for b -certificate complexity and consequently, we develop a direct proof of the certificate complexity formula of an NCF. Symmetry is another interesting property of Boolean functions and we significantly simplify the proofs of some recent theorems about partial symmetry of NCFs. We also describe the algebraic normal form of s -symmetric NCFs. We obtain the general formula of the cardinality of the set of n -variable s -symmetric Boolean NCFs for $s = 1, \dots, n$. In particular, we enumerate the strongly asymmetric Boolean NCFs.

Keywords: Boolean Function, Nested Canalizing Function, Layer Structure, Sensitivity, Certificate Complexity, Symmetry, Partial Symmetry.

1 Introduction

Nested canalizing functions (NCFs) were introduced in Kauffman et al. (2003). It was shown in Jarrah et al. (2007) that they are identical to the *unate cascade* functions, which have been studied extensively in engineering and computer science. It was shown in Butler et al. (2005) that this class of functions produces binary decision diagrams with the shortest average path length. Recently, canalizing and (partially) NCFs have received a lot of attention He and Macauley (2016); Jarrah et al. (2007); Kadelka et al. (2017a,b); Layne et al. (2012); Li and Adeyeye (2019); Li et al. (2013); Morizumi (2014); Murrugarra and Laubenbacher (2012); Shmulevich and Kauffman (2004).

In Cook et al. (1986), Cook et al. introduced the notion of sensitivity as a combinatorial measure for Boolean functions. It was extended by Nisan Nisan (1989, 1991) to block sensitivity. Certificate complexity was first introduced by Nisan in 1989 Nisan (1989, 1991).

In Li et al. (2013), a complete characterization for NCFs was obtained via its unique algebraic normal form, from which explicit formulas enumerating NCFs and their average sensitivity were derived.

In Theorem 3.6 Li and Adeyeye (2019), the formula of the sensitivity of any NCF was obtained based on a characterization of NCFs from Theorem 4.2 Li et al. (2013). It was shown that block sensitivity is the same as sensitivity for NCFs.

In Morizumi (2014), the author proved sensitivity is the same as the certificate complexity for *read-once* functions, a class of functions which include the NCFs, characterized as those that can be written using the logical conjunction, logical disjunction, and negation operations, where each variable appears at most once.

In this paper, we obtain formulas of b -certificate complexity of an NCF f for $b = 0, 1$. We denote them by $C_0(f)$ and $C_1(f)$. As a byproduct, we obtain a direct proof of the certificate complexity formula which is still the same as the formula of sensitivity Li and Adeyeye (2019).

Symmetric Boolean functions have important applications in coding theory and cryptography. In Section 4, based on Theorem 4.2 in Li et al. (2013), we study the properties of symmetric NCFs. We significantly simplify the proofs of some theorems in Rosenkrantz et al. (2019). We also investigate the relationship between the number of layers of an NCF and its number of symmetry levels. For $1 \leq s \leq n$, we obtain an explicit formula of the number of n -variable s -symmetric Boolean NCFs. When $s = n$, this number is the cardinality of strongly asymmetric NCFs. Specifically, we prove that there are more than $n!2^{n-1}$ strongly asymmetric NCFs when $n \geq 4$.

2 Preliminaries

In this section, we introduce the definitions and notations. Let \mathbb{F} be the field $\mathbb{F}_2 = \{0, 1\}$ and $f: \mathbb{F}^n \rightarrow \mathbb{F}$ be a function. It is well known Lidl and Niederreiter (1977) that f can be expressed as a polynomial, called the algebraic normal form (ANF):

$$f(x_1, \dots, x_n) = \bigoplus_{\substack{0 \leq k_i \leq 1 \\ i=1, \dots, n}} a_{k_1 \dots k_n} x_1^{k_1} \dots x_n^{k_n},$$

where each $a_{k_1 \dots k_n} \in \mathbb{F}$. The symbol \bigoplus stands for addition modulo 2.

A permutation of $[n] = \{1, \dots, n\}$ is a bijection from $[n]$ to $[n]$.

Definition 2.1 (Definition 2.3 in Jarrah et al. (2007), page 168) *Let f be a Boolean function in n variables and σ a permutation of $\{1, \dots, n\}$. The function f is nested canalizing in the variable order $x_{\sigma(1)}, \dots, x_{\sigma(n)}$ with canalizing input values a_1, \dots, a_n and canalized output values b_1, \dots, b_n , if it can be represented in the form*

$$f(x_1, \dots, x_n) = \begin{cases} b_1 & x_{\sigma(1)} = a_1 \\ b_2 & x_{\sigma(1)} = \bar{a}_1, x_{\sigma(2)} = a_2 \\ b_3 & x_{\sigma(1)} = \bar{a}_1, x_{\sigma(2)} = \bar{a}_2, x_{\sigma(3)} = a_3 \\ \vdots & \\ b_n & x_{\sigma(1)} = \bar{a}_1, x_{\sigma(2)} = \bar{a}_2, \dots, x_{\sigma(n-1)} = \bar{a}_{n-1}, x_{\sigma(n)} = a_n \\ \bar{b}_n & x_{\sigma(1)} = \bar{a}_1, x_{\sigma(2)} = \bar{a}_2, \dots, x_{\sigma(n-1)} = \bar{a}_{n-1}, x_{\sigma(n)} = \bar{a}_n, \end{cases}$$

where $\bar{a} = a \oplus 1$. The function f is nested canalizing if it is nested canalizing in some variable order.

Theorem 2.1 (Theorem 4.2 in Li et al. (2013), page 28) *Let $n \geq 2$. Then $f(x_1, \dots, x_n)$ is nested canalizing iff it can be uniquely written as*

$$f(x_1, \dots, x_n) = M_1(M_2(\dots(M_{r-1}(M_r \oplus 1) \oplus 1) \dots) \oplus 1) \oplus b, \quad (1)$$

where $M_i = \prod_{j=1}^{k_i} (x_{i_j} \oplus a_{i_j})$, $i = 1, \dots, r$, $k_i \geq 1$ for $i = 1, \dots, r-1$, $k_r \geq 2$, $k_1 + \dots + k_r = n$, $a_{i_j} \in \mathbb{F}_2$, $\{i_j \mid j = 1, \dots, k_i, i = 1, \dots, r\} = \{1, \dots, n\}$.

Because each NCF can be uniquely written as (1) and the number r is uniquely determined by f , we can define the following.

Definition 2.2 For $i = 1, \dots, r$, each M_i of an NCF f in (1) is defined as the i -th layer of f , where r is the number of layers. The vector $\langle k_1, \dots, k_r \rangle$ is called the layer structure, where $k_i \geq 1$ for $i = 1, \dots, r-1$, $k_r \geq 2$, $k_1 + \dots + k_r = n$. Each k_i is the size of M_i .

The i -th layer M_i is a product of variables and their negations. Such a product is called *extended monomial* in Li et al. (2013) or *psedomonomial* in Curto et al. (2013).

Note that we always have $k_r \geq 2$ by Theorem 2.1. Throughout this paper, all NCFs will be assumed to be on n variables, with layer structure $\langle k_1, \dots, k_r \rangle$.

3 Certificate Complexity of NCFs

Let $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{F}^n$. For any subset S of $[n]$, we form \mathbf{x}^S by negating the bits in \mathbf{x} indexed by elements of S . We denote $\mathbf{x}^{\{i\}}$ by \mathbf{x}^i .

Definition 3.1 (Definition 2.1 in Kenyon and Kutin (2004), page 45; Definition 1 in Rubinfeld (1995), page 297) The sensitivity of f at \mathbf{x} , denoted as $s(f, \mathbf{x})$, is the number of indices i such that $f(\mathbf{x}) \neq f(\mathbf{x}^i)$. The sensitivity of f is $s(f) = \max_{\mathbf{x} \in \{0,1\}^n} s(f, \mathbf{x})$.

Certificate complexity was first introduced by Nisan (1989, 1991), and was initially called sensitive complexity. In the following, we will slightly modify (actually, simplify) the definition of certificate, but the definition of certificate complexity will remain the same.

Definition 3.2 Let $f(x_1, \dots, x_n)$ be a Boolean function and $\alpha = (a_1, \dots, a_n) \in \mathbb{F}^n$ a word. If $\{i_1, \dots, i_k\} \subset [n]$ and the restriction $f(x_1, \dots, x_n)|_{x_{i_1}=a_{i_1}, \dots, x_{i_k}=a_{i_k}}$ is a constant function, where its constant value is $f(\alpha)$, then we call the subset $\{i_1, \dots, i_k\}$ a certificate of f on α .

Definition 3.3 The certificate complexity $C(f, \alpha)$ of f on α is defined as the smallest cardinality of a certificate of f on α . The certificate complexity $C(f)$ of f is defined as $\max\{C(f, y) \mid y \in \mathbb{F}^n\}$. The b -certificate complexity $C_b(f)$ of f , $b \in \mathbb{F}$, is defined as $\max\{C(f, y) \mid y \in \mathbb{F}^n, f(y) = b\}$.

Obviously, $C(f) = \max\{C_0(f), C_1(f)\}$.

Example 3.4 Let $f(x_1, x_2, x_3) = x_1 x_2 x_3 \oplus x_1 x_2 \oplus x_3$ and $g(x_1, x_2, x_3) = x_1 x_2 x_3$. We list the certificate complexity of f on every word in Table 1.

It is easy to check $C(g, (1, 1, 1)) = 3$ and $C(g, \alpha) = 1$, where $\alpha \neq (1, 1, 1)$. Hence, $C(g) = 3$.

Lemma 3.5 Let $f(x_1, \dots, x_n)$ be a Boolean function, σ be a permutation on $[n]$, and $\beta = (b_1, \dots, b_n) \in \mathbb{F}^n$. If $g = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ and $h = f(x_1 \oplus b_1, \dots, x_n \oplus b_n)$, then the certificate complexities of f , $f \oplus 1$, g , and h are the same.

α	$f(\alpha)$	$C(f, \alpha)$	Minimal certificates
(0,0,0)	0	2	$\{1,3\}, \{2,3\}$
(0,0,1)	1	1	$\{3\}$
(0,1,0)	0	2	$\{1,3\}$
(0,1,1)	1	1	$\{3\}$
(1,0,0)	0	2	$\{2,3\}$
(1,0,1)	1	1	$\{3\}$
(1,1,0)	1	2	$\{1,2\}$
(1,1,1)	1	1	$\{3\}$

Tab. 1: The certificate complexity for $f(x_1, x_2, x_3) = x_1x_2x_3 \oplus x_1x_2 \oplus x_3$ is 2.

Proof: Note that $f(x_1, \dots, x_n)|_{x_{i_1}=a_{i_1}, \dots, x_{i_k}=a_{i_k}}$ is a constant function if and only if

$$f(x_{\sigma(1)}, \dots, x_{\sigma(n)})|_{x_{\sigma(i_1)}=a_{i_1}, \dots, x_{\sigma(i_k)}=a_{i_k}}$$

is a constant function. Hence, $C(f, \alpha) = C(g, \alpha)$ for any $\alpha = (a_1, \dots, a_n) \in \mathbb{F}^n$, and thus $C(f) = C(g)$.

The function $f(x_1, \dots, x_n)|_{x_{i_1}=a_{i_1}, \dots, x_{i_k}=a_{i_k}}$ is a constant function if and only if

$$h = f(x_1 \oplus b_1, \dots, x_n \oplus b_n)|_{x_{i_1}=a_{i_1} \oplus b_{i_1}, \dots, x_{i_k}=a_{i_k} \oplus b_{i_k}}$$

is a constant. Hence, $C(f, \alpha) = C(h, \alpha + \beta)$ for any α and given β . Thus $C(f) = C(h)$ since $\alpha \mapsto \alpha \oplus \beta$ is a bijection of \mathbb{F}^n .

The function f is constant if and only if $f \oplus 1$ is constant, thus $C(f) = C(f \oplus 1)$. Specifically, $C_0(f) = C_1(f \oplus 1)$ and $C_1(f) = C_0(f \oplus 1)$. \square

In the following, let

$$f(x_1, \dots, x_n) = f_r = M_1(M_2(\dots(M_{r-1}(M_r \oplus 1) \oplus 1) \dots) \oplus 1) \quad (2)$$

be an NCF with r layers with monomials $M_1 = x_1 \cdots x_{k_1}$, $M_2 = x_{k_1+1} \cdots x_{k_1+k_2}$, \dots , $M_r = x_{k_1+\dots+k_{r-1}+1} \cdots x_n$.

With a straightforward calculation, we rewrite Equation (2) as

$$f(x_1, \dots, x_n) = f_r = M_1M_2 \cdots M_r \oplus M_1M_2 \cdots M_{r-1} \oplus \cdots \oplus M_1M_2 \oplus M_1. \quad (3)$$

Lemma 3.6 *If $f(x_1, \dots, x_n) = x_1 \cdots x_n$, then $C_0(f) = 1$ and $C_1(f) = n$. Hence, $C(f) = n$.*

Proof: It is clear that $C(f, (1, \dots, 1)) = n$, $f(1, \dots, 1) = 1$ and $C(f, \alpha) = 1$, $f(\alpha) = 0$ with $\alpha \neq (1, \dots, 1)$. \square Lemma 3.6 provides the certificate complexity of an NCF f_r with $r = 1$ layer. We are

ready to prove the following theorem.

Theorem 3.7 *If $f(x_1, \dots, x_n) = f_r = M_1(M_2(\dots(M_{r-1}(M_r \oplus 1) \oplus 1) \dots) \oplus 1)$ and $M_1 = x_1 \cdots x_{k_1}$, $M_2 = x_{k_1+1} \cdots x_{k_1+k_2}$, \dots , $M_r = x_{k_1+\dots+k_{r-1}+1} \cdots x_n$, $r \geq 2$, then*

$$C_0(f_r) = \begin{cases} k_2 + k_4 + \cdots + k_{r-1} + 1, & 2 \nmid r \\ k_2 + k_4 + \cdots + k_r, & 2 \mid r, \end{cases}$$

$$C_1(f_r) = \begin{cases} k_1 + k_3 + \cdots + k_r, & 2 \nmid r \\ k_1 + k_3 + \cdots + k_{r-1} + 1, & 2 \mid r, \end{cases}$$

Proof: We use induction on r to prove the formula of $C_0(f_r)$, and the proof of $C_1(f_r)$ is similar.

If $r = 2$, then $f_r = f_2 = M_1 M_2 + M_1 = M_1(M_2 \oplus 1)$. We will calculate $C(f_2, \alpha)$ for every α such that $f(\alpha) = 0$. Since $f(\alpha) = M_1(M_2 \oplus 1)(\alpha) = 0$ if and only if $M_1(\alpha) = 0$ or $M_1(\alpha) = M_2(\alpha) = 1$, we divide all such α into two disjoint groups. In the following, we simply write $M_1(\alpha) = 0$ as $M_1 = 0$, $M_1(\alpha) = 1$ as $M_1 = 1$ and so on.

Group 1: $M_1 = 0$.

In this case, at least one component of α corresponding to a variable in the first layer must be 0. Obviously, for such α , $C(f_2, \alpha) = 1$.

Group 2: $M_1 = 1$ and $M_2 = 1$.

In this case, there is only one possibility, namely, $\alpha = (1, \dots, 1)$. It is easy to check that $C(f_2, (1, \dots, 1)) = k_2$, the number of variables in M_2 .

Take the maximal value, we have $C_0(f_2) = k_2$.

If $r = 3$, then $f_3 = M_1(M_2(M_3 \oplus 1) \oplus 1) = 0 \iff M_1 = 0$ or $M_1 = M_2 = M_3 \oplus 1 = 1$. There are two disjoint groups.

Group A: $M_1 = 0$.

In this group, the certificate complexity for each word is 1.

Group B: $M_1 = 1$, $M_2 = 1$ and $M_3 = 0$.

In this group, $\alpha = (\overbrace{1, \dots, 1}^{k_1}, \overbrace{1, \dots, 1}^{k_2}, \overbrace{*, \dots, *, 0, *, \dots, *}^{k_3})$. First of all, if we just assign the values of the variables in M_1 and M_2 (all of those variables in α are 1s), since $f_3 = M_1 M_2 M_3 \oplus M_1 M_2 \oplus M_1$, the variables in M_3 never disappear (which means the function is not constant). So, we must assign one 0 to its corresponding variable in M_3 and reduce f_3 to $M_1(M_2 \oplus 1)$. Obviously, in order to make f_3 zero, it is necessary and sufficient to choose all the components of α corresponding to the variables in M_2 to assign. So, in this group, for any α , we have $C(f_3, \alpha) = k_2 + 1$.

In summary, taking the maximal value, yields $C_0(f_3) = k_2 + 1$.

Now we assume that the formula of $C_0(f_r)$ is true for any NCF with no more than $r - 1$ layers. Let us consider

$$f(x_1, \dots, x_n) = f_r = M_1(M_2(\cdots(M_{r-1}(M_r \oplus 1) \oplus 1) \cdots) \oplus 1)$$

$$= M_1 M_2 \cdots M_r \oplus M_1 M_2 \cdots M_{r-1} \oplus \cdots \oplus M_1 M_2 \oplus M_1.$$

If $g(x_{k_1+k_2+1}, \dots, x_n) = M_3 \cdots M_r \oplus M_3 \cdots M_{r-1} \oplus \cdots \oplus M_3 M_4 \oplus M_3$, we get $f_r = M_1(M_2(g \oplus 1) \oplus 1) = M_1 M_2 g \oplus M_1 M_2 \oplus M_1$. It is clear that $f_r = 0 \iff M_1 = 0$ or $M_1 = M_2 = g \oplus 1 = 1$. Next, we will evaluate $C(f_r, \alpha)$ for all $\alpha \in \mathbb{F}$ with $f(\alpha) = 0$.

Case 1: $M_1 = 0$.

In this case, the certificate complexity of the word is 1.

Case 2: $M_1 = 1$, $M_2 = 1$ and $g = 0$.

In this case, $\alpha = (\overbrace{1, \dots, 1}^{k_1}, \overbrace{1, \dots, 1}^{k_2}, \alpha')$, where α' is a word with length $n - k_1 - k_2$. Obviously, we have $f_r(\alpha) = 0$ if and only if $g(\alpha') = 0$.

For a fixed α' (equivalently, a fixed α), we try to reduce $f_r = M_1 M_2 g \oplus M_1 M_2 \oplus M_1$ to zero by assigning values of α to the variables of f_r . Since $M_1 M_2$ will never be zero, we must try to reduce g to zero first. Once g is zero, we get $f_r = M_1(M_2 \oplus 1)$. Hence, we have $C(f_r, \alpha) = k_2 + C(g, \alpha')$, and

$$\max\{C(f_r, \alpha) \mid \alpha, f_r(\alpha) = 0\} = k_2 + \max\{C(g, \alpha') \mid \alpha', g(\alpha') = 0\} = k_2 + C_0(g).$$

Since g is an NCF with $r - 2$ layers (the first layer is M_3 , the second layer is M_4 and so on), by the induction hypothesis, we have

$$C_0(g) = \begin{cases} k_4 + k_6 + \dots + k_{r-1} + 1, & 2 \nmid (r-2) \\ k_4 + k_6 + \dots + k_r, & 2 \mid (r-2). \end{cases}$$

Hence, $\max\{C(f_r, \alpha) \mid \alpha, f_r(\alpha) = 0\} = k_2 + C_0(g)$ is

$$k_2 + \begin{cases} k_4 + k_6 + \dots + k_{r-1} + 1, & 2 \nmid (r-2) \\ k_4 + k_6 + \dots + k_r, & 2 \mid (r-2) \end{cases} = \begin{cases} k_2 + k_4 + \dots + k_{r-1} + 1, & 2 \nmid r \\ k_2 + k_4 + \dots + k_r, & 2 \mid r. \end{cases}$$

For any word in Case 1, the certificate complexity is only 1. In summary, we have

$$C_0(f_r) = \begin{cases} k_2 + k_4 + \dots + k_{r-1} + 1, & 2 \nmid r \\ k_2 + k_4 + \dots + k_r, & 2 \mid r. \end{cases}$$

□

Because of Lemma 3.5, we have the following.

Corollary 3.8 *If any NCF is written as the one in Theorem 2.1, then*

$$C(f_r) = \begin{cases} \max\{k_1 + k_3 + \dots + k_r, k_2 + k_4 + \dots + k_{r-1} + 1\}, & 2 \nmid r \\ \max\{k_1 + k_3 + \dots + k_{r-1} + 1, k_2 + k_4 + \dots + k_r\}, & 2 \mid r. \end{cases}$$

Hence, the certificate complexity of NCF is uniquely determined by the layer structure (k_1, \dots, k_r) .

The above formula is the same as the sensitivity formula $s(f_r)$ in Theorem 3.6 Li and Adeyeye (2019).

4 Symmetric Properties of NCFs

In 1938, Shannon Shannon (1938) recognized that symmetric functions have efficient switch network implementations. Since then, a lot of research has been done on symmetric or partially symmetric Boolean functions. Symmetry detection is important in logic synthesis, technology mapping, binary decision diagram minimization, and testing Arnold and Harrison (1963); Das and Sheng (1971); Mishchenko (2003). In Rosenkrantz et al. (2019), the authors investigated the symmetric and partial symmetric properties of Boolean NCFs. They also presented an algorithm for testing whether a given partial symmetric function is an NCF. In this section, we use a formula in Li et al. (2013) to give simple proofs for several theorems in Rosenkrantz et al. (2019). We also study the relationship between the number of layers r and the number of symmetry levels s (the function is s -symmetric) of NCFs. Furthermore, we obtain the formula of the number of n -variable s -symmetric NCFs. In particular, we obtain the formula of the number of strongly asymmetric NCFs. We start this section by providing some basic definitions and notations.

It is well known that a permutation can be written as the product of disjoint cycles. A t -cycle $(i_1 \cdots i_t)$ sends i_k to i_{k+1} for $k = 1, \dots, t-1$ and sends i_t to i_1 . Namely, $i_1 \mapsto i_2 \mapsto \cdots \mapsto i_t \mapsto i_1$. A 2-cycle is called a transposition. Any permutation can be written as a product of transpositions. For example, $(12 \cdots n) = ((n-1)n) \cdots (2n)(1n)$, where cycles are read right-to-left, as in function composition.

Definition 4.1 Let f be a Boolean function and $\sigma = (ij)$ a 2-cycle. We say that variable x_i is equivalent to x_j if $f(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ (namely, $f(\dots, x_i, \dots, x_j, \dots) = f(\dots, x_j, \dots, x_i, \dots)$). We denote this by $i \sim_f j$.

It is clear that $i \sim_f j$ is an equivalence relation over $[n]$. We call $\tilde{i} = \{j \mid j \sim_f i\}$ a symmetric class of f . If $[n]/\sim_f = \{\tilde{i} \mid i \in [n]\}$ and $s = |[n]/\sim_f|$ is the cardinality of $[n]/\sim_f$, we call $f(x_1, \dots, x_n)$ s -symmetric.

The definition of s -symmetric in this paper is equivalent to the concept of properly s -symmetric in Rosenkrantz et al. (2019).

Example 4.2 Let $f(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = x_1x_2x_3x_4 \oplus x_5x_6 \oplus x_7$. Then $\tilde{1} = \tilde{2} = \tilde{3} = \tilde{4} = \{1, 2, 3, 4\}$, $\tilde{5} = \tilde{6} = \{5, 6\}$, $\tilde{7} = \{7\}$. This function is 3-symmetric.

Definition 4.3 If there is an index i such that $|\tilde{i}| \geq 2$, i.e., $s = |[n]/\sim_f| \leq n-1$, then we call f partially symmetric. If $s = 1$, we call f totally symmetric or symmetric.

Obviously, a function is not partially symmetric if and only if it is n -symmetric.

For applications of 1-symmetric (totally symmetric) Boolean functions to cryptography, see Canteaut and Videau (2005) from 2005. More results on (totally) symmetric Boolean functions can be found in Cai et al. (1996); Castro et al. (2018); Cusick and Li (2005); Cusick et al. (2008); Li and Qi (2006); Li and Xiang (2007); Maitra and Sarker (2002); Mitchell (1990); Savicky (1994).

Definition 4.4 (Rosenkrantz et al. (2019), page 3) A Boolean function $f(x_1, \dots, x_n)$ is strongly asymmetric if $f(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ implies σ is the identity.

Obviously, if a Boolean function is strongly asymmetric then it is n -symmetric.

Let

$$f(x_1, x_2, x_3, x_4, x_5, x_6) = x_1x_2 \oplus x_2x_3 \oplus x_3x_4 \oplus x_4x_5 \oplus x_5x_1 \oplus x_6.$$

It is easy to check that f is 6-symmetric (not partially symmetric) but not strongly asymmetric since

$$f(x_1, x_2, x_3, x_4, x_5, x_6) = f(x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)}, x_{\sigma(4)}, x_{\sigma(5)}, x_{\sigma(6)}) \text{ for } \sigma = (12345).$$

In the following, we frequently use Equation (1). Recall that a_{i_j} is called the canalizing input of the variable x_{i_j} .

Proposition 4.5 (Theorem 3.1 in Rosenkrantz et al. (2019)) All variables in the same symmetric class of an NCF must be in the same layer and have the same canalizing input.

Proof: This follows immediately from the uniqueness of Equation (1). □

Remark 4.6 In each layer M_j , for $j = 1, \dots, r$, there are either one or two symmetric classes. If there are two symmetric classes, then one has canalizing input 0, and the other has canalizing input 1.

Proposition 4.7 Let $n \geq 2$ and $\langle k_1, \dots, k_r \rangle$ be the layer structure of an NCF f . If $k_j \geq 3$ for some j , then f is partially symmetric. Moreover, if f is s -symmetric, then $\lceil \frac{s}{2} \rceil \leq r \leq \min\{n-1, s\}$.

Proof: If $k_j \geq 3$ for some j , then at least two variables have the same canalizing inputs by Remark 4.6. Hence, this layer has a symmetric class with at least 2 variables and f is partially symmetric. From Equation (1), the last layer has at least two variables, so $r \leq n - 1$. We have $r \leq s$ since all variables from different layers must belong to different symmetric classes. Finally, because each layer contributes at most two symmetric classes, we obtain $s \leq 2r$ which means $\lceil \frac{s}{2} \rceil \leq r$. \square

Proposition 4.8 *Let f be an s -symmetric NCF with r layers. Then $r \leq s \leq \min\{2r, n\}$.*

Proof: It follows from the proof of the previous property. \square

Proposition 4.9 *(Theorem 3.2 in Rosenkrantz et al. (2019)) If an NCF contains r_1 layers with only one canalizing input, and r_2 layers with two distinct canalizing inputs, then it is $(r_1 + 2r_2)$ -symmetric.*

Proof: This is a straightforward application of the uniqueness of Equation (1). \square

Next, we will provide a new and shorter proof for the following proposition.

Proposition 4.10 *(Theorem 3.7 in Rosenkrantz et al. (2019)) An n -variable NCF is strongly asymmetric iff it is n -symmetric.*

Proof: We already know that strong asymmetry implies n -symmetry.

If an NCF f is n -symmetric, i.e., not partially symmetric, then each layer has one or two variables with different canalizing inputs by Proposition 4.7. If there is a permutation σ such that $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n)$, then, for any i , because of the uniqueness of Equation (1), we know $x_{\sigma(i)}$ and x_i must be in the same layer of $f(x_1, \dots, x_n)$. If this layer has only one variable, then $\sigma(i) = i$. If this layer has two variables x_i and x_j with $i \neq j$, then this layer must be $M = x_i(x_j \oplus 1)$ or $M = (x_i \oplus 1)x_j$. Without loss of the generality, we assume $M = x_i(x_j \oplus 1)$, if $\sigma(i) = j$, then $\sigma(j) = i$ since $x_{\sigma(i)}$ and x_i must be in the same layer. Because $x_{\sigma(i)}(x_{\sigma(j)} \oplus 1) = x_j(x_i \oplus 1) \neq M$, which is contrary to the uniqueness of Equation (1). Hence, we still have $\sigma(i) = i$. In summary, we always have $\sigma(i) = i$ for any i . Therefore, σ is the identity and f is strongly asymmetric. \square

Strongly asymmetric NCFs were studied in Rosenkrantz et al. (2019), and in Theorem 3.8, the authors enumerated those that have exactly $n - 1$ layers, which is the maximal possible number because $k_r \geq 2$. Though they used this assumption in their proof, they apparently omitted it from the theorem statement. We will state the correct version below, and refer the reader to Rosenkrantz et al. (2019) (Theorem 3.8) for the proof.

Theorem 4.11 *There are $n!2^{n-1}$ strongly asymmetric NCFs on n variables with exactly $n - 1$ layers.*

In the remainder of this section, we will enumerate the s -symmetric NCFs on n variables. As a corollary, we will derive a formula for the number of strongly asymmetric NCFs.

Let $N(n, s)$ be the cardinality of the set of n -variable s -symmetric Boolean NCFs.

Proposition 4.12 *(Proposition 3.9 in Rosenkrantz et al. (2019)) If $n \geq 2$, then $N(n, 1) = 4$.*

Proof: Since f is 1-symmetric, i.e., totally symmetric, then there is only one layer, and all canalizing inputs must be the same. So, f must be one of the following functions: $x_1 \cdots x_n$, $x_1 \cdots x_n \oplus 1$, $(x_1 \oplus 1) \cdots (x_n \oplus 1)$ or $(x_1 \oplus 1) \cdots (x_n \oplus 1) \oplus 1$. \square

Theorem 4.13 For $n \geq 2$, the number of strongly asymmetric NCFs is

$$N(n, n) = \frac{n!}{\sqrt{2}}((1 + \sqrt{2})^{n-1} - (1 - \sqrt{2})^{n-1}).$$

Proof: By Theorem 2.1, we have

$$f(x_1, \dots, x_n) = M_1(M_2(\dots(M_{r-1}(M_r \oplus 1) \oplus 1) \dots) \oplus 1) \oplus b.$$

1. It is clear that b has two choices.
2. By Proposition 4.7, we have $\lceil \frac{n}{2} \rceil \leq r \leq n - 1$.
3. For each layer structure $\langle k_1, \dots, k_r \rangle$, since f is strongly asymmetric (not partially symmetric), we have $1 \leq k_i \leq 2$ by Proposition 4.7 and thus $k_r = 2$ due to $k_r \geq 2$ always. There are

$$\binom{n}{k_1} \binom{n - k_1}{k_2} \binom{n - k_1 - k_2}{k_3} \dots \binom{n - k_1 - \dots - k_{r-1}}{k_r} = \frac{n!}{k_1! k_2! \dots k_r!}$$

ways to distribute the n variables to the layers.

4. Each layer M_j is either $x_i \oplus a$ or $(x_k \oplus a)(x_l \oplus a \oplus 1)$. In any case, there are two choices. Hence, totally, there are 2^r choices.

Combining the information above, we obtain

$$N(n, n) = 2 \sum_{\lceil \frac{n}{2} \rceil \leq r \leq n-1} \sum_{\substack{k_1 + \dots + k_r = n \\ 1 \leq k_i \leq 2, k_r = 2}} \frac{n!}{k_1! k_2! \dots k_r!} 2^r.$$

If $n \geq 3$, then it can be written as

$$N(n, n) = \sum_{\lceil \frac{n}{2} \rceil \leq r \leq n-1} \sum_{\substack{k_1 + \dots + k_{r-1} = n-2 \\ 1 \leq k_i \leq 2}} \frac{n!}{k_1! k_2! \dots k_{r-1}!} 2^r.$$

Suppose that exactly j elements of the set $\{k_1, \dots, k_{r-1}\}$ are equal to 2. We obtain $2j + r - 1 - j = n - 2$ since $k_1 + \dots + k_{r-1} = n - 2$. This implies $j = n - r - 1$. Hence,

$$N(n, n) = \sum_{\lceil \frac{n}{2} \rceil \leq r \leq n-1} \binom{r-1}{n-r-1} \frac{n!}{2^{n-r-1}} 2^r = 2n! \sum_{\lceil \frac{n}{2} \rceil \leq r \leq n-1} \binom{r-1}{n-r-1} 2^{2r-n}.$$

Let $k = n - r - 1$, and so $r = n - k - 1$. It is clear that $\lceil \frac{n}{2} \rceil \leq r \leq n - 1 \Leftrightarrow 0 \leq k \leq \lfloor \frac{n}{2} \rfloor - 1$. We have

$$N(n, n) = 2n! \sum_{0 \leq k \leq \lfloor \frac{n}{2} \rfloor - 1} \binom{n-2-k}{k} 2^{n-2-2k}.$$

Since $\binom{n-2-k}{k} = 0$ if $k \geq \lfloor \frac{n}{2} \rfloor$, we have

$$N(n, n) = 2n! \sum_{k=0}^{n-2} \binom{n-2-k}{k} 2^{n-2-2k}.$$

We assumed that $n \geq 3$ in the above proof. A direct calculation shows that the formula is still true for $n = 2$.

Let

$$p_n(t) = 2^{n-2}t^{n-2}\left(1 + \frac{t}{2}\right)^{n-2} + 2^{n-3}t^{n-3}\left(1 + \frac{t}{2}\right)^{n-3} + \dots + 1 = \frac{2^{n-1}t^{n-1}\left(1 + \frac{t}{2}\right)^{n-1} - 1}{2t\left(1 + \frac{t}{2}\right) - 1}.$$

A direct computation shows that the sum $\sum_{k=0}^{n-2} \binom{n-2-k}{k} 2^{n-2-2k}$ is the coefficient of t^{n-2} in the polynomial $p_n(t)$. We rewrite $p_n(t)$ as a sum of two rational expressions:

$$p_n(t) = t^{n-1} \frac{(2+t)^{n-1}}{t^2 + 2t - 1} + \frac{-1}{t^2 + 2t - 1}.$$

If we write these two rational expressions as power series, it is clear that the smallest order of the terms in the first rational expression is $n - 1$. So, the sum $\sum_{k=0}^{n-2} \binom{n-2-k}{k} 2^{n-2-2k}$ is the coefficient of t^{n-2} in the power series of $\frac{-1}{t^2 + 2t - 1}$. We have

$$\frac{-1}{t^2 + 2t - 1} = \frac{-1}{2\sqrt{2}(-1 - \sqrt{2} - t)} + \frac{1}{2\sqrt{2}(-1 + \sqrt{2} - t)}.$$

By the formula of geometric series, we obtain

$$\frac{-1}{t^2 + 2t - 1} = \frac{1}{2\sqrt{2}} \sum_{k=0}^{\infty} (-(1 - \sqrt{2})^{k+1} + (\sqrt{2} + 1)^{k+1}) t^k.$$

Therefore, the coefficient of t^{n-2} is $\frac{(\sqrt{2}+1)^{n-1} - (1-\sqrt{2})^{n-1}}{2\sqrt{2}}$. Consequently, we obtain

$$N(n, n) = \frac{n!}{\sqrt{2}} ((1 + \sqrt{2})^{n-1} - (1 - \sqrt{2})^{n-1}).$$

□

When $n = 2, 3, 4$, we have $N(2, 2) = 4$ and $N(3, 3) = 24$ and $N(4, 4) = 240$.

From the above proof, if $n \geq 4$, then

$$N(n, n) = 2n! \sum_{k=0}^{n-2} \binom{n-2-k}{k} 2^{n-2-2k} = 2n!(2^{n-2} + (n-3)2^{n-4} + \dots) > 2n!2^{n-2} = n!2^{n-1}.$$

We have obtained the formulas of $N(n, 1)$ and $N(n, n)$. In the following, we derive the formula $N(n, s)$ for $n \geq 3$ and $2 \leq s \leq n - 1$.

Theorem 4.14 *Let $n \geq 3$ and $2 \leq s \leq n - 1$. Then $N(n, s)$, the number of n -variable s -symmetric NCFs, is*

$$2 \sum_{\substack{[\frac{s}{2}] \leq r \leq s \\ k_1 + \dots + k_r = n \\ k_i \geq 1, k_r \geq 2}} \sum_{k_1! k_2! \dots k_r!} \frac{n!}{k_1! k_2! \dots k_r!} \sum_{\substack{t_1 + \dots + t_r = s \\ 1 \leq t_i \leq \min\{2, k_i\}}} \prod_{1 \leq i \leq r} ((t_i - 1)(2^{k_i} - 2) + 1 - (-1)^{t_i}).$$

Proof: By Theorem 2.1, we have

$$f(x_1, \dots, x_n) = M_1(M_2(\dots(M_{r-1}(M_r \oplus 1) \oplus 1) \dots) \oplus 1) \oplus b.$$

1. It is clear that b has two choices.
2. By Proposition 4.7, we get $\lceil \frac{s}{2} \rceil \leq r \leq s$.
3. For each layer structure $\langle k_1, \dots, k_r \rangle$, there are

$$\frac{n!}{k_1!k_2! \dots k_r!}$$

ways to distribute the n variables.

4. Each layer M_i contributes t_i symmetry classes, where $1 \leq t_i \leq \min\{2, k_i\}$ and $t_1 + \dots + t_r = s$ since f is s -symmetric.

5. For each fixed layer $M_i = \prod_{j=1}^{k_i} (x_{i_j} \oplus a_{i_j})$, there are 2^{k_i} choices for M_i . Two of them contribute one symmetric class (all canalizing inputs a_{i_j} are equal) and $2^{k_i} - 2$ of them contribute two symmetric classes. Since

$$(t_i - 1)(2^{k_i} - 2) + 1 - (-1)^{t_i} = \begin{cases} 2, & t_i = 1 \\ 2^{k_i} - 2, & t_i = 2, \end{cases}$$

there are $(t_i - 1)(2^{k_i} - 2) + 1 - (-1)^{t_i}$ choices of M_i contributing t_i symmetric classes for $t_i = 1, 2$. Combining the information above, we obtain the formula of $N(n, s)$. □

We have

$$\sum_{j=1}^n N(n, j) = 2^{n+1} \sum_{r=1}^{n-1} \sum_{\substack{k_1 + \dots + k_r = n \\ k_i \geq 1, k_r \geq 2}} \frac{n!}{k_1!k_2! \dots k_r!}.$$

The right side is the cardinality of the set of n -variable Boolean NCFs according to Li et al. (2013).

When $n \geq 2$, it is clear that $N(n, s) \geq 1$. Consequently, for any s , there exists NCFs which are not s -symmetric. In particular, there exists n -variable NCFs that are not $(n - 1)$ -symmetric (Corollary 3.3 in Rosenkrantz et al. (2019)).

From Corollary 4.9 in Li et al. (2013), the number of NCFs with r layers is

$$2^{n+1} \sum_{\substack{k_1 + \dots + k_r = n \\ k_i \geq 1, k_r \geq 2}} \frac{n!}{k_1!k_2! \dots k_r!}. \quad (4)$$

When r is the maximal value $n - 1$, the above number can be simplified as $n!2^n$.

5 Conclusion

In this paper, we obtained the formulas of the b -certificate complexity of any NCF for $b = 0, 1$. We extended some results from Rosenkrantz et al. (2019) on symmetric and partially symmetric NCFs and we studied the relationship between the number of layers and the number of symmetry levels. We derived the formulas of the cardinality of all n -variable s -symmetric Boolean NCFs. As a special case, we obtained the number of n -variable strongly asymmetric Boolean NCFs.

Acknowledgements

We greatly appreciate the referees for their patience and insightful comments. In particular, we are very grateful to a referee for his/her constructive suggestions to significantly simplify our original formula and receive a much better formula in Theorem 4.13.

References

- R. F. Arnold and M. A. Harrison. *Algebraic properties of symmetric and partially symmetric Boolean functions*, iee transactions of electronic computers, vol: ec-12 issue: 3, pp. 244-251 edition, 1963.
- J. T. Butler, T. Sasao, and M. Matsuura. *Average path length of binary decision diagrams*, iee transactions on computers, 54, pp. 1041-1053 edition, 2005.
- J.-Y. Cai, F. Green, and T. Thierauf. *On the correlation of symmetric functions*, math. syst. theory, vol. 29, no. 3, pp. 245-258 edition, 1996.
- A. Canteaut and M. Videau. *Symmetric Boolean functions*, iee transactions on information theory, vol. 51, no. 8, pp. 2791-2811. edition, 2005.
- F. N. Castro, O. E. González, and L. A. Medina. *Diophantine equations with binomial coefficients and perturbation of symmetric Boolean functions*, iee transactions on information theory, vol 64, no.2, pp.1347-1360 edition, 2018.
- S. A. Cook, C. Dwork, and R. Reischuk. *Upper and lower time bounds for parallel random access machines without simultaneous writes*, siam j. comput, 15, pp. 87-89 edition, 1986.
- C. Curto, V. I. Veliz-Cuba, and N. Youngs. *The Neural Ring: An Algebraic Tool for Analyzing the Intrinsic Structure of Neural Codes*, bull math biol, 75 (9) edition, 2013.
- T. W. Cusick and Y. Li. *k-th order symmetric SAC Boolean functions and bisecting binomial coefficients*, discrete applied mathematics, 149 (2005) 73-86 edition, 2005.
- T. W. Cusick, Y. Li, and P. Štaničá. *Balanced symmetric functions over $GF(p)$* , iee transactions on information theory, vol 54, pp.1304-1307 edition, 2008.
- S. R. Das and C. L. Sheng. *On detecting total or partial symmetry of switching function*, iee transactions on computers, vol: c-20 , issue: 3, march 1971, pp. 352-355 edition, 1971.
- Q. He and M. Macauley. *Stratification and enumeration of Boolean functions by canalizing depth*, physica d 314 (2016), pp. 1-8. edition, 2016.
- A. Jarrah, B. Raposa, and R. Laubenbacher. *Nested canalizing, unate cascade, and polynomial*, physica d 233 (2007), pp. 167-174 edition, 2007.
- C. Kadelka, J. Kuipers, and R. Laubenbacher. *The influence of canalization on the robustness of Boolean networks*, physica d, 353-354 (2017), 39-47. edition, 2017a.
- C. Kadelka, Y. Li, J. Kuipers, J. O. Adeyeye, and R. Laubenbacher. *Multistate nested canalizing functions and their networks*, theoretical computer science, 675,2 (2017), 1-14 edition, 2017b.

- S. A. Kauffman, C. Peterson, B. Samueleson, and C. Troein. *Random Boolean network models and the yeast transcription network*, *proc. natl. acad. sci* 100 (25) (2003), pp. 14796-14799. edition, 2003.
- C. Kenyon and S. Kutin. *Sensitivity, block sensitivity, and l -block sensitivity of Boolean functions*, *information and computation*, 189 (2004), pp. 43-53 edition, 2004.
- L. Layne, E. Dimitrova, and M. Macauley. *Nested canalizing depth and network stability*, *bull math biol*, 74 (2012), pp. 422-433. edition, 2012.
- N. Li and W.-F. Qi. *Symmetric Boolean functions depending on an odd number of variables with maximum algebraic immunity*, *ieee transactions on information theory*, vol 52, no.5, may, pp.2271-2273 edition, 2006.
- Y. Li and J. O. Adeyeye. *Maximal sensitivity of Boolean nested canalizing functions*, *theoretical computer science*, 791, (2019), 116-122 edition, 2019.
- Y. Li and Z.-H. Xiang. *To determine symmetric $PC(k)$ Boolean functions by its definition*, *sichuan daxue xuebao*, 44 (2007) no.2, 209-212 edition, 2007.
- Y. Li, J. O. Adeyeye, D. Murrugarra, B. Aguilar, and R. Laubenbacher. *Boolean nested canalizing functions: A comprehensive analysis*, *theoretical computer science*, 481, (2013), 24-36 edition, 2013.
- R. Lidl and H. Niederreiter. *Finite Fields*, *cambridge university press*, new york edition, 1977.
- S. Maitra and P. Sarker. *Maximum nonlinearity of symmetric Boolean functions on odd number of variable*, *ieee transactions on information theory*, vol 48, no. 9, pp.2626-2630, edition, 2002.
- A. Mishchenko. *Fast computation of symmetries in Boolean function*, *ieee transactions on computer-aided design of integrated circuits and system*, vol. 22, no. 11, november 2003 pp. 1588-1593. edition, 2003.
- C. J. Mitchell. *Enumerating Boolean functions of cryptographic significance*, *j. crypto.*, vol. 2, no 3, pp. 155-170 edition, 1990.
- H. Morizumi. *Sensitivity, block sensitivity, and certificate complexity of unate functions and read-once functions*, in: diaz j., lanese i., sangiorgi d. (eds) *theoretical computer science. tcs 2014. lecture notes in computer science*, vol 8705. *springer*, berlin, heidelberg. edition, 2014.
- D. Murrugarra and R. Laubenbacher. *The number of multistate nested canalizing functions*, *physica d: nonlinear phenomena*, 241, 929-938 edition, 2012.
- N. Nisan. *CREW PRAMs and decision trees*, *proc. 21th acm stoc* (1989), pp. 327-335 edition, 1989.
- N. Nisan. *CREW PRAMs and decision trees*, *siam j. comput*, 20 (6) (1991), pp. 999-1007 edition, 1991.
- D. J. Rosenkrantz, M. V. Marathe, S. S. Ravi, and R. E. Stearns. *Symmetric properties of nested canalizing functions*, *discrete mathematics and theoretical computer science*, dmtcs vol. 21:4, 2019, #19 edition, 2019.
- D. Rubinstein. *Sensitivity VS. block sensitivity of Boolean functions*, *combinatorica* 15 (2) (1995), pp. 297-299 edition, 1995.

- P. Savicky. *On the bent Boolean functions that are symmetric*, europ. j. combin, vol. 15, pp. 407-410 edition, 1994.
- C. Shannon. *A symbolic analysis of relay and switching circuits*, aiee trans, 57: 713-723 edition, 1938.
- I. Shmulevich and S. A. Kauffman. *Activities and sensitivities in Boolean network models*, physical review letters vol 93, no. 4, 23 july 2004, 048701. edition, 2004.