Further Constructions of Resilient Boolean Functions With Very High Nonlinearity

Subhamoy Maitra and Enes Pasalic

***Abstract*—One well-known method of generating key stream sequences for stream ciphers is to combine the outputs of several linear-feedback shift registers (LFSR) using a combining Boolean function. Here we concentrate on the design of good combining Boolean functions. We provide resilient Boolean functions with currently best known nonlinearity. These functions were not known earlier and the issues related to their existence were posed as open questions in the literature. Some of the functions we construct here achieve the provable upper bound on nonlinearity for resilient Boolean functions. Our technique interlinks mathematical results with classical computer search.**

***Index Terms*—Boolean functions, correlation immunity, nonlinearity, resiliency, stream ciphers.**

I. Introduction

CONSTRUCTION of correlation immune and resilient (balanced correlation immune) Boolean functions has been an interesting research area from mid 1980s [20], [21], [9], [2], [19]. These functions have immediate applications in stream cipher systems. Very recently, Sarkar and Maitra [18] have provided weight divisibility results on correlation immune and resilient Boolean functions which, in turn, present nontrivial upper bounds on the nonlinearity of such functions. Similar kinds of results related to weight divisibility and upper bounds on nonlinearity of resilient and correlation immune Boolean functions have also been presented independently by Tarannikov [22] and Zheng and Zhang [25]. Currently, the weight divisibility results have been settled by Carlet [4] (see also the work by Carlet and Sarkar [5]) for resilient and correlation immune Boolean functions involving the algebraic degree as well.

These weight divisibility results have direct consequences for the upper bound on nonlinearity of these functions and a benchmark in design of such resilient Boolean functions has thus been settled. In the other direction, construction of these functions achieving the upper bound on nonlinearity strengthens the tightness of the upper bound results.

In a more practical direction, these functions have immediate applications in stream cipher cryptosystems. A standard model of stream cipher [20], [21], [7] combines the outputs of several independent linear feedback shift register (LFSR) sequences using a nonlinear Boolean function to produce the key stream. This key stream is bitwise XORed with the message bit stream to produce the cipher. The decryption machinery is identical to the encryption machinery. Getting the kind of Boolean functions which we propose here provides the best possible tradeoff among the parameters important to resist the known cryptanalytic techniques (see [21], [12], [10], [3] and the references therein).

It is now well accepted that for a Boolean function to be used in stream cipher systems, it must satisfy the properties of balancedness, high nonlinearity, high algebraic degree, and high order of correlation immunity (see Section II for definitions). All of these parameters are important in resisting different kinds of attacks. Also, it is not possible to get the best possible values for each of these parameters separately and there are certain tradeoffs involved among the above parameters. Siegenthaler showed [20] that for an n-variable balanced function of degree d and order of correlation immunity $m (1 \leq m \leq n-2)$, we have $m + d \leq n - 1$ (known as Siegenthaler's inequality in the literature). Recently, the exact nature of the tradeoff among order of correlation immunity, nonlinearity, and algebraic degree has also been investigated [18], [22], [25], [4], [5], [26]. Earlier, a series of papers (see [19], [8], [14], [17] and the references therein) have approached the construction problem by fixing the number of variables and the order of correlation immunity and then trying to design balanced Boolean functions with as high nonlinearity as possible. However, the most recent papers [18], [22], [25], [4], [5], [15], [11] concentrate on the construction of functions achieving the upper bound on nonlinearity.

It should be noted that the current results in construction of resilient functions, achieving upper bounds on nonlinearity, concentrate on high order of resiliency. In fact, for n-variable, m-resilient functions, when the value of m is in the range $m > \frac{n}{2} - 2$ all of these functions have three-valued Walsh spectra. However, the functions we consider here are of low order of resiliency and the Walsh spectra are not three-valued.

In this paper, for the first time we construct 8-variable, 1-resilient Boolean functions with nonlinearity 116. Earlier, all the 8-variable resilient functions of different orders (except order 1) with maximum possible algebraic degree and maximum possible nonlinearity (equal to the upper bound) were known. We here close the issue by proving the case for order 1 as well. Our construction uses 5-variable Boolean functions with nonlinearity 11 as input and then concatenate two such functions to construct a 6-variable unbalanced first-order correlation immune function

Manuscript received August 27, 2001; revised January 23, 2002. The material in this paper was presented in part at the Conference on Sequences and Their Applications, SETA '01, Bergen, Norway, May 13–17, 2001.

S. Maitra is with the Computer and Statistical Service Center, Indian Statistical Institute, Calcutta, Pin 700 108, India (e-mail: subho@isical.ac.in).

E. Pasalic is with the Department of Information Technology, Lund University, 221 00 Lund, Sweden (e-mail: enes@it.lth.se).

Communicated by N. I. Koblitz, Associate Editor for Complexity and Cryptography.

Publisher Item Identifier S 0018-9448(02)05165-9.

with nonlinearity 26. From such a function we provide a construction of an 8-variable unbalanced first-order correlation immune function with nonlinearity 116. We then exploit a randomized search technique using bit manipulation and linear transformation to get an 8-variable, 1-resilient function with nonlinearity 116. The construction of this function was posed as an open question in [18].

We also present a new algorithm for modifying degree nonoptimized 1-resilient functions into degree optimized ones. We use this for the construction of 1-resilient, degree $n-2$ functions on n variables, $n \geq 12$ even, which gives better nonlinearity than that presented in [17]. In particular we could construct highly nonlinear 1-resilient functions on 10 variables, which gives functions with nonlinearity 488 and algebraic degree 6 and 8. The construction of such functions has been left as open questions in [18]. We also provide a method to construct degree optimized m-resilient functions ($m > 1$) from a special class of degree nonoptimized m-resilient functions considered in [6].

We conclude with n-variable, m-resilient functions for $1 < m \leq \frac{n}{2} - 2$. An existing recursive construction [22], [15] is analyzed in detail which generates resilient functions on higher number of variables from resilient functions on lower number of variables. We identify important functions (the 1-resilient functions we find here and the functions from [17]) as initial inputs to these recursive constructions. As example, we use the $(8, 1, 6, 116)$ functions in *desired* form to construct the $(11, 3, 7, 976)$ functions. Also, using the functions of [17] as initial ones, we provide a method to construct $(n, m, n-m-1, 2^{n-1} - 2^{\frac{2n+m-3}{2}})$ functions when $2n - 3m \equiv 3 \bmod 4$. In particular, for $m = 3$, a construction of $(n, 3, n-4, 2^{n-1} - 2^{\frac{n}{2}})$ functions is provided. For this order of resiliency, the functions in this series either attain the same quality results or supersede all previous constructions in terms of nonlinearity value.

II. Preliminaries

Definition 1: A Boolean function on n variables may be viewed as a mapping from $\{0, 1\}^n$ into $\{0, 1\}$. By Ω_n we mean the set of all Boolean functions of n variables. We interpret a Boolean function $f(X_1, \ldots, X_n)$ as the output column of its *truth table* f, i.e., a binary string of length 2^n

$$f = [f(0, 0, \ldots, 0), f(1, 0, \ldots, 0), f(0, 1, \ldots, 0), \ldots, f(1, 1, \ldots, 1)].$$

We also use the notation $f[a]$, which corresponds to the ath entry in the function's truth table. To save the space we also represent this binary string of length 2^n as a string of hexadecimal digits of length $\frac{2^n}{4} = 2^{n-2}$. For binary strings S_1, S_2 of the same length λ, we denote by $\#(S_1 = S_2)$ (respectively, $\#(S_1 \neq S_2)$), the number of places where S_1 and S_2 are equal (respectively, unequal). The *Hamming distance* between S_1, S_2 is denoted by $d(S_1, S_2)$, i.e.,

$$d(S_1, S_2) = \#(S_1 \neq S_2).$$

We also define

$$wd(S_1, S_2) = \#(S_1 = S_2) - \#(S_1 \neq S_2).$$

Note that, $wd(S_1, S_2) = \lambda - 2\,d(S_1, S_2)$. Also, the *Hamming weight* or simply the weight of a binary string S is the number of ones in S. This is denoted by $\mathrm{wt}(S)$. An n-variable function f is said to be *balanced* if its output column in the truth table contains equal number of 0's and 1's (i.e., $\mathrm{wt}(f) = 2^{n-1}$).

Definition 2: The addition operator over GF(2) is denoted by $+$. An n-variable Boolean function $f(X_1, \ldots, X_n)$ can be considered to be a multivariate polynomial over GF(2). This polynomial can be expressed as a sum of kth-order products ($0 \leq k \leq n$) of distinct variables. More precisely, $f(X_1, \ldots, X_n)$ can be written as

$$a_0 + \sum_{i=1}^{i=n} a_i X_i + \sum_{1 \leq i < j \leq n} a_{ij} X_i X_j + \cdots + a_{12\ldots n} X_1 X_2 \cdots X_n$$

where the coefficients $a_0, a_{ij}, \ldots, a_{12\ldots n} \in \{0, 1\}$. This representation of f is called the *algebraic normal form* (ANF) of f. The number of variables in the highest order product term with nonzero coefficient is called the *algebraic degree*, or simply the degree of f.

In this paper, we will use concatenation of Boolean functions. Consider $f_1, f_2 \in \Omega_{n-1}$, and $f \in \Omega_n$. Then by concatenation of f_1 and f_2, we mean that the output columns of the truth tables of f_1, f_2 will be concatenated to provide the output column of the truth table of an n-variable function. We denote the concatenation of f_1, f_2 by $f_1 \| f_2$. Thus, $f = f_1 \| f_2$ means that in algebraic normal form,

$$f(X_1, \ldots, X_n) = (1 + X_n) f_1(X_1, \ldots, X_{n-1}) + X_n f_2(X_1, \ldots, X_{n-1}).$$

Also, for the complement function of f we use the notation $\overline{f}$, i.e., $\overline{f} = 1 + f$.

Definition 3: Functions of degree at most one are called *affine* functions. An affine function with constant term equal to zero is called a *linear* function. The set of all n-variable affine (respectively, linear) functions is denoted by $A(n)$ (respectively, $L(n)$). The nonlinearity of an n-variable function f is

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

i.e., the distance from the set of all n-variable affine functions.

Definition 4: Let $X = (X_1, \ldots, X_n)$ and $\omega = (\omega_1, \ldots, \omega_n)$ both belong to $\{0, 1\}^n$ and

$$X \cdot \omega = X_1 \omega_1 + \cdots + X_n \omega_n.$$

Let $f(X)$ be a Boolean function on n variables. Then the *Walsh transform* of $f(X)$ is a real-valued function over $\{0, 1\}^n$ that can be defined as

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

Note that $W_f(\omega) = wd(f, l_\omega)$, where l_ω denotes the linear function on n variables given by $l_\omega(X) = \omega \cdot X$. For a Boolean function f, we define

$$NZ(f) = \{\omega | W_f(\omega) \neq 0\}$$

where W_f is the Walsh transform of f.

Definition 5 [9]: A function $f(X_1, \ldots, X_n)$ is mth-order *correlation immune* (CI) iff its Walsh transform satisfies $W_f(\omega) = 0$, for $1 \leq wt(\omega) \leq m$. Note that f is balanced iff $W_f(0) = 0$. Balanced mth-order correlation immune functions are called *m-resilient* functions. Thus, a function $f(X_1, \ldots, X_n)$ is m-resilient iff its Walsh transform satisfies $W_f(\omega) = 0$, for $0 \leq wt(\omega) \leq m$.

At this point, we recall two important properties of the Walsh spectrum.

1) The first one is referred to as Parseval's equality [7], and it states that for any Boolean function $f \in \Omega_n$

$$\sum_{\omega \in \{0,1\}^n} (W_f(\omega))^2 = 2^{2n}.$$

2) The second one is a more recent result [18] concerning the weight divisibility of the Walsh spectrum. According to this result, for any m-resilient (respectively, m-CI) function $f \in \Omega_n$ we have $W_f(\omega) \equiv 0 \bmod 2^{m+2}$ (respectively, 2^{m+1}), for any ω from $\{0, 1\}^n$. Moreover, involving the algebraic degree the result is as follows [4], [5]. For any m-resilient (respectively, m-CI) degree d function $f \in \Omega_n$ we have

$$W_f(\omega) \equiv 0 \bmod 2^{m+2+\lfloor \frac{n-m-2}{d} \rfloor}$$

(respectively, $2^{m+1+\lfloor \frac{n-m-1}{d} \rfloor}$), for any ω from $\{0, 1\}^n$.

By an (n, m, d, x) function we denote an n-variable, m-resilient function with degree d and nonlinearity x. By $(n, 0, d, x)$ function we mean a balanced n-variable function with degree d and nonlinearity x. By $[n, m, d, x]$ function we denote an n-variable unbalanced correlation immune function of order m, nonlinearity x, and degree d. In the above notation, a component is replaced by a "$-$" if it is not specified, e.g., $(n, m, -, x)$ if the degree is not specified.

Let us now clearly clarify the exact upper bounds on the nonlinearity of resilient Boolean functions from the weight divisibility results. We use the term $\mathrm{nlmax}(n)$ to denote the maximum nonlinearity of an n-variable Boolean function. It is known that for n even, $\mathrm{nlmax}(n) = 2^{n-1} - 2^{\frac{n}{2}-1}$ [16] and the functions which attain this nonlinearity are called bent functions. However, the problem remains open for odd n. It is clear that the bent functions cannot be correlation immune. For the n-odd case, to write the upper bound on the nonlinearity of resilient functions, we assume here that the functions attaining the maximum possible nonlinearity $\mathrm{nlmax}(n)$ may have the correlation immunity property. As example, for $n = 5, 7$, the maximum possible nonlinearities are 12, 56, respectively, [1], [13] and we get resilient functions, e.g., $(5, 1, 3, 12)$, $(7, 1, 5, 56)$, $(7, 2, 4, 56)$ [17], [15], at those nonlinearities. We first consider the case of (n, m, d, x) functions.

1) If n is even, and $m + \lfloor \frac{n-m-2}{d} \rfloor > \frac{n}{2} - 2$, then

$$x \leq 2^{n-1} - 2^{m+1+\lfloor \frac{n-m-2}{d} \rfloor}.$$

2) If n is even, and $m + \lfloor \frac{n-m-2}{d} \rfloor \leq \frac{n}{2} - 2$, then

$$x \leq 2^{n-1} - 2^{\frac{n}{2}-1} - 2^{m+1+\lfloor \frac{n-m-2}{d} \rfloor}.$$

3) If n is odd, and

$$\mathrm{nlmax}(n) \geq 2^{n-1} - 2^{m+1+\lfloor \frac{n-m-2}{d} \rfloor}$$

then

$$x \leq 2^{n-1} - 2^{m+1+\lfloor \frac{n-m-2}{d} \rfloor}.$$

4) If n is odd, and

$$\mathrm{nlmax}(n) < 2^{n-1} - 2^{m-1-\lfloor \frac{n-m-2}{d} \rfloor}$$

then x is the highest multiple of $2^{m-1-\lfloor \frac{n-m-2}{d} \rfloor}$ which is $\leq \mathrm{nlmax}(n)$.

To simplify the scenario, let us now explain the result for $(n, m, n-m-1, x)$ functions. That is the case when we consider the maximum possible algebraic degree (the functions optimizing the Siegenthaler's inequality).

1) If n is even, and $m > \frac{n}{2} - 2$, then $x \leq 2^{n-1} - 2^{m+1}$.
2) If n is even, and $m \leq \frac{n}{2} - 2$, then $x \leq 2^{n-1} - 2^{\frac{n}{2}-1} - 2^{m+1}$.
3) If n is odd, and $\mathrm{nlmax}(n) \geq 2^{n-1} - 2^{m+1}$, then $x \leq 2^{n-1} - 2^{m+1}$.
4) If n is odd, and $\mathrm{nlmax}(n) < 2^{n-1} - 2^{m+1}$, then x is the highest multiple of 2^{m+1} which is $\leq \mathrm{nlmax}(n)$.

The above upper bounds will help in comparing the quality of our results. Next we present a standard technique of constructing correlation immune functions [14]. The method is as follows. *We will refer to this method as the LT-method throughout this paper.*

Given a function $f \in \Omega_n$, we define

$$S_f = \{\omega \in \{0, 1\}^n | W_f(\omega) = 0\}$$

where W_f is the Walsh transform of f. If there exists n linearly independent vectors in S_f, then one can construct a nonsingular $n \times n$ matrix B_f whose rows are linearly independent vectors from S_f. Let, $C_f = {B_f}^{-1}$. Now if we construct a function $f'(X) = f(C_f X)$, then both f' and f *have the same nonlinearity and algebraic degree*. Moreover, $W_{f'}(\omega) = 0$ for $\mathrm{wt}(\omega) = 1$, where $W_{f'}$ is the Walsh transform of f'. This ensures that f' is first-order correlation immune. Also, if f is balanced then f' is balanced and hence 1-resilient.

III. Construction of $(8, 1, 6, 116)$ Functions

In this section, we present the construction method of an $(8, 1, 6, 116)$ function. For this purpose we first consider how to construct a $[6, 1, 5, 26]$ function with weight 30. Next we provide the construction of an $[8, 1, 5, 116]$ function having weight 124. Using these $[8, 1, 5, 116]$ functions, we present the construction of $(8, 1, 6, 116)$ functions. Note that we use a similar kind of idea for the constructions of a $[6, 1, 5, 26]$ function (see [15]) and an $[8, 1, 5, 116]$ function (see [11]). However, we choose the weight of the functions with more care in this initiative.

A. Construction of $[6, 1, 5, 26]$ and $[8, 1, 5, 116]$ Functions

Consider a 6-variable function h with algebraic degree 5 and nonlinearity 26. Note that by proper permutation of the input variables, the function h can always be written as $(1 \mid X_6)h_1 \mid X_6 h_2$ where h_1, h_2 are 5-variable functions with algebraic degree 5. Concentrate on a specific degree 5 term in the ANF of h.

Clearly, there is one variable X_i which is not present in that degree 5 term. Let us consider some permutation of the input variables such that X_i goes to X_6. Then , by fixing $X_6 = 0$, we get the function h_1 and putting $X_6 = 1$ we get the function h_2. It is clear that the algebraic degree of h_1, h_2 must be 5. It is also important to note that the nonlinearities of both h_1, h_2 must be 11. The functions h_1, h_2 are of degree 5 and hence of odd weight. This redsults in the nonlinearity of h_1, h_2 being odd. Now if the nonlinearity of either h_1 or h_2 is less than 11, then the function h cannot have nonlinearity 26.

We initially concentrate on [1, Table I], where there are just four out of 49 equivalence classes with nonlinearity 11. These functions are also of weight 11. Consider such a function h_1'. Now, there are linear functions l of 5-variables, such that $d(h_1', l) = 15$ and hence $wt(h_1' + l) = 15$. In this manner, we can select four 5-variable functions from four different representative classes [1] which are of nonlinearity 11 and weight 15. *The choice of weight* 15 *function is from the motivation that they are closed to balancedness, and its importance will be clearer in the next subsection.* Let the representative functions be as follows, which we present as output column of the truth tables in hexadecimal format:

$$h_1 = 565656A8 \qquad h_2 = 6633550E,$$
$$h_3 = 635F7240 \qquad h_4 = 724E6351.$$

Thus, concatenating each representative of the ith class, $i = 1, \ldots, 4$, with all nonsingular affine transformations applied to a representant of the jth class, $j = i, \ldots, 4$, we will get a lot of 6-variable functions h with weight 30. We will choose those with nonlinearity 26. Now, if the function h has $wd(.)$ value zero with six independent linear functions, then we can use the LT-method as described in Section II to get [6, 1, 5, 26] functions with weight 30. At this point, let us explain the following technical result.

Proposition 1: Let h be an n-variable Boolean function with algebraic degree $d \geq 2$ and nonlinearity x. Consider the function

$$g(X_1, \ldots, X_{n+2}) = X_{n+2}X_{n+1} + h(X_1, \ldots, X_n)$$

i.e., the truth table of g is of the form $h\|h\|h\|\overline{h}$. Then g has degree d and nonlinearity $2^n + 2x$.

Proof: Note that for any affine function $\lambda \in A(n+2)$, we can write λ in any one of the forms $l\|l\|l\|l, l\|l\|l\|\overline{l}, l\|l\|\overline{l}\|l, l\|l\|\overline{l}\|\overline{l}$, $l\|\overline{l}\|l\|l$, $l\|\overline{l}\|\overline{l}\|\overline{l}$, where $l \in A(n)$. Now consider $\lambda = l\|l\|l\|l$. Then,

$$\begin{aligned} d(g, \lambda) &= d(h\|h\|h\|\overline{h}, l\|l\|l\|l) \\ &= d(h, l) + d(h, l) + d(h, l) + d(\overline{h}, l) \\ &= 2d(h, l) + d(h, l) + d(\overline{h}, l) \\ &= 2x + 2^n. \end{aligned}$$

The result is similar for λ of other forms as well. This gives the nonlinearity result.

Since g is of the form $g(X_1, \ldots, X_{n+2}) = X_{n+2}X_{n+1} + h(X_1, \ldots, X_n)$, the algebraic degree of g is the same as the algebraic degree of h. □

Now we concentrate on [8, 1, 5, 116] functions. We recall the construction method provided in [11]. Let h be an $[n, 1, d, x]$ function, where n is even. Consider the function

$$g(X_1, \ldots, X_{n+2}) = X_{n+2}X_{n+1} + h(X_1, \ldots, X_n)$$

i.e., the truth table of g is of the form $h\|h\|h\|\overline{h}$. From Proposition 1, the nonlinearity of g is $2^n + 2x$. Now concentrate on the following $(n+2)\times(n+2)$ matrix C_g in terms of the LT-method as given in Section II

$$C_g = \begin{bmatrix} 1 & 0 & 0 & . & . & 0 & 0 & 1 \\ 1 & 0 & 0 & . & . & 0 & 1 & 0 \\ 0 & 0 & 0 & . & . & 1 & 0 & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & 1 & . & . & 0 & 0 & 0 \\ 0 & 1 & 0 & . & . & 0 & 0 & 0 \\ 1 & 0 & 0 & . & . & 0 & 0 & 0 \end{bmatrix}.$$

Consider $X = (X_1, \ldots, X_{n+2})$ and we interpret it as a column vector here. From [11], the function $g'(X) = g(C_g X)$ is an $[n-2, 1, d, 2^n - 2x]$ function. Also, if the weight of h is y, then the weight of g is $2^n + 2y$, and g' has also the same weight.

Hence, if we start from a [6, 1, 5, 26] function h with weight 30, it is possible to construct an [8, 1, 5, 116] function g' with weight 124. To get an (8, 1, 6, 116) function we will modify these [8, 1, 5, 116] functions.

B. An (8, 1, 6, 116) *Function*

Let f be an [8, 1, 5, 116] function with weight 124. Select any four bits i_1, i_2, i_3, i_4 of f such that $f[i_1] = f[i_2] = f[i_3] = f[i_4] = 0$. Then construct a function g such that

$$g[i_1] = g[i_2] = g[i_3] = g[i_4] = 1$$

and g is equal to f for all other bits. Note that g is balanced. Now check the nonlinearity of f. If it stays at 116, then we try to find out a set of right independent nonzero linear functions of 8-variables and if such a set exists, we use the LT-method used in Section II to get an (8, 1, −, 116) function. We know [4], [5] that the maximum possible nonlinearity of an n-variable (n even), m-resilient ($m \leq \frac{n}{2} - 2$), degree d function is

$$2^{n-1} - 2^{\frac{n}{2}-1} - 2^{m+1+\lfloor\frac{n-m-2}{d}\rfloor}.$$

Thus, the degree of an (8, 1, −, 116) function must be 6, since if the degree is less than 6, then the nonlinearity will be at most 112. Hence any (8, 1, −, 116) function must be an (8, 1, 6, 116) function.

Let us start with the [6, 1, 5, 26] function

$$h = \texttt{3A1A6C17E8D1582D}$$

which has weight 30. From this we get the [8, 1, 5, 116] function of weight 124

$$\texttt{1E111E11EEEEE11EEEE1111EE111E1E1}$$
$$\texttt{877787888878887778877778888888777}.$$

Then we use this [8, 1, 5, 116] function to get one (8, 1, 6, 116) function

$$\texttt{6F4FC635EE280B7135159C4BB472512B}$$
$$\texttt{CA8A932DD2E4A84D90D0C977CABEF217}.$$

Let us now highlight why we are more careful about the choice of the weight of correlation immune functions [6, 1, 5, 26], [8, 1, 5, 116]. Note that the weight of the [8, 1, 5, 116] function is 124 and we need to change (0 to 1)

TABLE I
NONLINEARITY RESULTS FOR 8-VARIABLE RESILIENT BOOLEAN FUNCTIONS

m	1	2	3	4	5	6
x	116	112	112	96	64	0

only four positions in its truth table to get a balanced function. If the weight of the [8, 1, 5, 116] function is chosen further away from 128 (the weight for balancedness), then we need to change more bit positions to attain balancedness and, in that case, preserving the nonlinearity at the same value 116 would have been less probable. Thus, in this construction method the weight 124 of the [8, 1, 5, 116] function plays a crucial role. This in turn justifies the choice of weight 30 for the [6, 1, 5, 26] function.

This construction completely solves the maximum nonlinearity issue of resilient functions on eight variables. Table I shows the maximum nonlinearity x corresponding to each order of resiliency m. Note that, for 8-variable functions, it is possible to construct resilient functions achieving the upper bound on nonlinearity for each order. Also, it is very clear from [22], [4], [5] that all these functions possess the maximum possible algebraic degree $(7 - m)$ where m is the order of resiliency.

IV. DEGREE OPTIMIZATION OF RESILIENT FUNCTIONS

In this section, we discuss a construction of degree optimized resilient functions from degree nonoptimized resilient functions with minimum decrease in nonlinearity. We start with the following technical result concerning the construction of 1-resilient degree optimized functions.

Theorem 1: For any 1-resilient degree nonoptimized function $f(X)$ with nonlinearity $nl(f)$, it is possible to obtain a new degree optimized 1-resilient function $f^*(X)$ having nonlinearity $nl(f^*) \in \{nl(f), nl(f) \pm 4\}$ by complementing four suitably chosen bit positions in the truth table of f.

Proof: Without loss of generality, we suppose that the function $f(X)$ depends on the variable X_n.

Let us assume that there exist $\alpha, \beta \in \{0, 1\}^{n-1}$, such that

$$f(\alpha, 0) \neq f(\alpha, 1)$$
$$f(\beta, 0) \neq f(\beta, 1)$$

and

$$f(\alpha, 0) \neq f(\beta, 0).$$

That is, $f(\alpha, 0) = f(\beta, 1)$ and $f(\alpha, 1) = f(\beta, 0)$. Note that by $f(\alpha, 0)$ we mean $f(\alpha, X_n = 0)$. The existence of such α, β is left to the end of the proof.

For convenience, we denote

$$K = \{(\alpha, 0), (\alpha, 1), (\beta, 0), (\beta, 1)\}.$$

Let us construct a function $f^*(X)$ defined by

$$f^*(X) = \begin{cases} f(X), & X \notin K \\ 1 + f(X), & X \in K. \end{cases}$$

Under the assumptions above, we prove that f^* is a 1-resilient degree optimized function.

Clearly, f^* is balanced since f is balanced. Also, it can be checked that the functional value of the four inputs in K has contribution 0 in the calculation of both $W_{f^*}(\omega)_{|wt(\omega)=1}$ and $W_f(\omega)_{|wt(\omega)=1}$. Thus,

$$W_{f^*}(\omega)_{|wt(\omega)=1} = W_f(\omega)_{|wt(\omega)=1} = 0.$$

Hence, f^* is also first-order correlation immune. Thus, f^* is 1-resilient.

Combining the weight divisibility results [18] with the fact that we changed 4 bits in the truth table of f, it can be deduced that

$$nl(f^*) \in \{nl(f), nl(f) \pm 4\}.$$

Next, we prove that the function $f^*(X)$ is a degree optimized function, i.e., $\deg(f^*) = n - 2$. The ANF of the function f can be expanded as

$$\begin{aligned} f(X) = {} & f(0, \ldots, 0)(1 + X_1) \cdots (1 + X_n) \\ & + f(1, \ldots, 0)X_1 \cdots (1 + X_n) + \cdots + f(\alpha, 0) \\ & \cdot \underbrace{(1 + X_1 + \alpha_1) \cdots (1 + X_{n-1} + \alpha_{n-1})}_{g(X_1, \ldots, X_{n-1})} \\ & \cdot (1 + X_n) + \cdots + f(\alpha, 1) \\ & \cdot \underbrace{(1 + X_1 + \alpha_1) \cdots (1 + X_{n-1} + \alpha_{n-1})}_{g(X_1, \ldots, X_{n-1})} X_n \\ & + \cdots + f(\beta, 0) \\ & \cdot \underbrace{(1 + X_1 + \beta_1) \cdots (1 + X_{n-1} + \beta_{n-1})}_{h(X_1, \ldots, X_{n-1})} \\ & \cdot (1 + X_n) + \cdots + f(\beta, 1) \\ & \cdot \underbrace{(1 + X_1 + \beta_1) \cdots (1 + X_{n-1} + \beta_{n-1})}_{h(X_1, \ldots, X_{n-1})} X_n \\ & + \cdots + f(1, \ldots, 1)X_1 \cdots X_n. \end{aligned}$$

The ANF of $f^*(X)$ will be given by

$$\begin{aligned} f^*(X) &= f(X) + g(X_1, \ldots, X_{n-1})(1 + X_n) \\ &\quad + g(X_1, \ldots, X_{n-1})X_n \\ &\quad + h(X_1, \ldots, X_{n-1})(1 + X_n) \\ &\quad + h(X_1, \ldots, X_{n-1})X_n \\ &= f(X) + g(X_1, \ldots, X_{n-1}) + h(X_1, \ldots, X_{n-1}). \end{aligned}$$

Since $\deg(f) < n - 2$, it is sufficient to prove that $\deg(g + h) = n - 2$. We first note that the terms of order $n - 1$ cancel out each other, i.e., $\deg(g + h) < n - 1$. Since $\alpha \neq \beta$, we have $\alpha_i \neq \beta_i$ for some i, $i = 1, \ldots, n-1$. This ensures the presence of the term $X_1 \cdots X_{i-1}X_{i+1} \cdots X_{n-1}$ in the ANF of $f^*(X)$. Thus, $\deg(f^*) = n - 2$ as claimed.

At last, we prove the existence of α, β satisfying the conditions above. Let us denote

$$S_{i,i} = \{X' : f(X', 0) = f(X', 1) = i\}, \qquad \text{for } i = 0, 1.$$

Furthermore, let

$$S_{0,1} = \{X' : f(X', 0) = 0,\ f(X', 1) = 1\}$$

and

$$S_{1,0} = \{X' : f(X', 0) = 1,\ f(X', 1) = 0\}.$$

Thus, we want to prove that both $S_{0,1}$ and $S_{1,0}$ are nonempty.

Suppose, to the contrary, that $f(X', 0) = f(X', 1)$ for all $X' \in \{0, 1\}^{n-1}$. Then f does not depend on x_n, which is a contradiction to the assumption. Thus, there exists a vector, say α, such that $f(\alpha, 0) \neq f(\alpha, 1)$. Without loss of generality, we assume $f(\alpha, 0) = 0$, i.e., $\alpha \in S_{0,1}$. We now prove that $S_{1,0}$ is also nonempty. Suppose that $S_{1,0} = \emptyset$. As f is balanced, $|S_{0,0}| = |S_{1,1}|$. Furthermore, since f is a 1-CI function, we also must have $|S_{0,0}| + |S_{0,1}| = |S_{1,1}| + |S_{1,0}|$. Thus, combining these two conditions, we obtain $|S_{0,1}| = |S_{1,0}|$ and $S_{1,0}$ is nonempty. This concludes the proof. □

The importance of this result lies in the fact that we do not put any requirement on degree nonoptimized function f. Thus, it enables us to use a construction that attains the highest nonlinearity value for 1-resilient functions not taking into account the algebraic degree of such functions. Applying Theorem 1 on such a function we obtain a degree optimized function paying the price of decreasing the nonlinearity by a constant value of 1 in the worst case.

Note that in certain cases the degree optimization algorithm may be extended to higher order resilient functions, i.e., $m > 1$. In this case, the number of points to be complemented is 2^{m+1}. Similarly as in Theorem 1, it can be shown that there exists a set of cardinality 2^{m+1} such that complementing the function's values in the points belonging to this set would yield a degree optimized function. However, in general, the order of resiliency is not preserved. On the positive side, a good example of applying the degree optimization algorithm, when constructing degree optimized functions of higher resiliency order, is the linear concatenation method in [6]. Here, an m-resilient, n-variable function f is constructed by concatenating 2^{n-k} distinct m-resilient linear functions in k variables. In this construction, the nonlinearity value is directly related to the parameter k, and given by $nl(f) = 2^{n-1} - 2^{k-1}$, where k is to be minimized to achieve a high nonlinearity value. For convenience, we denote these linear m-resilient functions by l_i, $i = 0, \ldots, 2^{n-k} - 1$.

Theorem 2 [6]: For given n and m, $n > m + 2$, it is possible to construct a nonlinear $(n, m, d, 2^{n-1} - 2^{k-1})$ function f, where $k = \min k_p$ is the minimum integer satisfying

$$\binom{k_p}{m+1} + \binom{k_p}{m+2} + \cdots + \binom{k_p}{k_p} \geq 2^{n-k_p}. \quad (1)$$

Furthermore, the maximum algebraic degree of f depends on k and, in general, $d \leq n - m - 1$. The ANF of function f is given by

$$f(X, Y) = \sum_{\tau \in \{0,1\}^{n-k}} (Y_1 + \tau_1) \cdots (Y_{n-k} + \tau_{n-k}) l_{\lfloor \tau \rfloor}(X) \quad (2)$$

where $\lfloor \tau \rfloor$ is a decimal representation of the vector τ, and $X \in \{0, 1\}^k$, $Y \in \{0, 1\}^{n-k}$.

Without loss of generality, we assume that given a degree nonoptimized m-resilient function f (as constructed by Theorem 2) there exists $\alpha \in \{0, 1\}^{n-m-1}$, such that

$$f(X) = \begin{cases} X_1 + \cdots + X_m, & X = (X_1, \ldots, X_m, 0, \alpha) \\ 1 + X_1 + \cdots + X_m, & X = (X_1, \ldots, X_m, 1, \alpha). \end{cases}$$

Define a new function f^* derived from f as

$$f^*(X) = \begin{cases} f(X), & \lambda \notin \{(z, \alpha); z \in \{0, 1\}^{m+1}\} \\ 1 + f(X), & \lambda \in \{(z, \alpha); z \in \{0, 1\}^{m+1}\}. \end{cases}$$

Theorem 3: Let f be an $(n, m, d, 2^{n-1} - 2^{k-1})$ function constructed by means of Theorem 2. We assume that $d < n - m - 1$. Let $l_0 \in L(k)$ used in the construction of f be given by $l_0(X) = X_1 + \cdots + X_{m+1}$, whereas $l_1, \ldots, l_{2^{n-k}-1}$ are arbitrarily chosen distinct m-resilient linear functions on $L(k)$. Then the function f^*, constructed above from f, is an $(n, m, n - m - 1, x)$ function. Here, $x \in \{nl(f), nl(f) \pm 2^{m+1}\}$.

Proof: For convenience, let

$$\lambda(\lambda_1, \ldots, \lambda_{m+1}) = \lambda_1 + \cdots + \lambda_{m+1}.$$

Then

$$l_0 = \underbrace{\lambda \| \lambda \| \cdots \| \lambda \| \lambda}_{2^{k-m-1}}$$

or written in ANF,

$$l_0(X_1, \ldots, X_k) = \lambda(X_1, \ldots, X_{m+1}).$$

Let $\alpha = (0, \ldots, 0)$ is an all-zero vector of length $n - m - 1$. Now, complementing 2^{m-1} bits in f at positions $\{(z, \alpha), z \in \{0, 1\}^{m+1}\}$ corresponds to the use of a new component function $\hat{l}_0 = \overline{\lambda} \| \lambda \| \cdots \| \lambda \| \lambda$ in f^*. Thus, $f^* = \hat{l}_0 l_1 \cdots l_{2^{n-k}-1}$. Since λ is m-resilient, $\overline{\lambda}$ is m-resilient as well. Hence, $\hat{l}_0$ is m-resilient and f^* is m-resilient. Using weight divisibility results of resilient functions [18], we conclude that

$$nl(f^*) \in \{nl(f) nl(f) \pm 2^{m+1}\}.$$

It remains to prove that $\deg(f^*) = n - m - 1$. This is equivalent to proving that $\deg(\hat{l}_0) = k - m - 1$ for $m + 2 \leq k \leq n - 1$. Then, $\deg(f^*) = n - k + k - m - 1 = n - m - 1$. We prove this by induction on k. For convenience, we write $k = i + m + 1$, for $i \in \{1, 2, \ldots, n - m - 2\}$. Let $\hat{l}_0^i$ denote a function on Ω_{i+m+1} defined by

$$\hat{l}_0^i = \underbrace{\overline{\lambda} \| \lambda \| \cdots \| \lambda \| \lambda}_{2^i}.$$

Thus, we need to prove that $\deg(\hat{l}_0^i) = i$.

1) Base case of induction, $i = 1$. Then $\hat{l}_0^1 = \overline{\lambda} \| \lambda$. Clearly, $\hat{l}_0^1 = 1 + X_1 + \cdots + X_{m+1} + X_{m+2}$, i.e., $\hat{l}_0^1 \in A(m + 2)$ and $\deg(\hat{l}_0^1) = 1$.
2) We assume that the statement is shown for $i = j - 1$, i.e., $\deg(\hat{l}_0^{j-1}) = j - 1$.
3) We prove now that $\deg(\hat{l}_0^j) = j$ for $\hat{l}_0^j \in \Omega_{j+m+1}$. We write $\hat{l}_0^j = \hat{l}_0^{j-1} l_0$ for $l_0 \in L(m + j)$. Since $\deg(\hat{l}_0^{j-1}) = j - 1$ and $\deg(l_0) = 1$, we conclude that $\deg(\hat{l}_0^j) = j$. □

Example 1: We use the extended degree optimization algorithm to construct a $(7, 2, 4, 48)$ function starting from a $(7, 2, 3, 48)$ function provided in [6]. For $n = 7$, $m = 2$, and the minimum value for k is $k = 5$. According to the earlier discussion, let us choose $2^{n-k} = 4$ distinct 2-resilient linear functions, say $l_0, \ldots, l_3$, given by $l_0 = X_1 + X_2 + X_3$, $l_1 = X_1 + X_2 + X_4$, $l_2 = X_1 + X_4 + X_5$, $l_3 = X_3 + X_4 + X_5$. Define $f: \{0, 1\}^7 \to \{0, 1\}$ as

$$f(X, Y) = \sum_{\tau \in \{0,1\}^2} (Y_1 + \tau_1)(Y_2 + \tau_2) l_{[\tau]}(X) \qquad (3)$$

where $[\tau]$ is a decimal representation of vector τ. It is easy to verify that f is a $(7, 2, 3, 48)$ function. Let $\alpha = (0, 0, 0, 0)$. By complementing the values of f in points (y, α) for $y \in \{0, 1\}^3$, we construct a $(7, 2, 4, x)$ function f^*, where $x \in \{40, 48, 56\}$. It can be checked that in this case $x = 48$ yields a $(7, 2, 4, 48)$ function.

Note that the maximum possible nonlinearity of a 7-variable, 2-resilient, degree optimized function is 56. Such $(7, 2, 4, 56)$ functions were constructed in [15] using a computer search.

Example 2: Now we consider the construction of degree optimized 2-resilient 11-variable functions. The best known nonlinearity for this type of function is 984, i.e., a function $(11, 2, 8, 984)$ has been reported in [17]. Here, as in the previous example, we start with an $(11, 2, 6, 992)$ function g constructed by means of method in [6]. Applying the extended degree optimization algorithm on g, that is, complementing $2^{m-1} = 8$ bits satisfying the conditions above, we obtain an $(11, 2, 8, x)$ function g^*, where $x \in \{984, 992, 1000\}$. Thus, even if the worst case is considered, i.e., $x = 984$, we obtain the same quality result as in [17].

V. 1-Resilient Functions on Even Number of Variables

Now we concentrate on 1-resilient Boolean functions on higher number of variables. First we present a technical result.

Proposition 2: Let h be an $(n, 1, d, x)$ function, where $d \geq 2$. Consider the function

$$g(X_1, \ldots, X_{n+2}) = X_{n+2}X_{n+1} + h(X_1, \ldots, X_n)$$

i.e., the truth table of g is of the form $h\|h\|h\|\overline{h}$. Then g is an $(n + 2, 1, d, 2^n + 2x)$ function.

Proof: The nonlinearity and algebraic degree results follow from Proposition 1. Since h is 1-resilient, $\overline{h}$ is also 1–resilient. Thus, it is easy to see that $g = h\|h\|h\|\overline{h}$ is also 1-resilient. □

Initially, we start with degree nonoptimized functions and then provide a modification to get degree optimized resilient functions. We already have $(8, 1, 6, 2^{8-1} - 2^{\frac{8}{2}} + 2^{\frac{8}{2}-2})$ functions. Thus, using such a function as the initial one we get the following result.

Theorem 4: For even $n \geq 8$, it is possible to construct $(n, 1, 6, 2^{n-1} - 2^{\frac{n}{2}} + 2^{\frac{n}{2}-2})$.

Proof: Given an $(m, 1, 6, 2^{m-1} - 2^{\frac{m}{2}} + 2^{\frac{m}{2}-2})$ function we can use Proposition 2 to get an

$$(m+2, 1, 6, 2^m + 2(2^{m-1} - 2^{\frac{m}{2}} + 2^{\frac{m}{2}-2}) = 2^{m+1} - 2^{\frac{m+2}{2}} + 2^{\frac{m+2}{2}-2})$$

function, which is the inductive step. The initial condition is proved using the $(8, 1, 6, 2^{8-1} - 2^{\frac{8}{2}} + 2^{\frac{8}{2}-2} = 116)$ function. □

Putting $n = 10$, from Theorem 4 we get a $(10, 1, 6, 488)$ function, which has also been placed as an open problem in [18], [15]. Also, this function has an important significance in terms of weight divisibility results. From the weight divisibility results [4], [5], it is known that the maximum possible nonlinearity of an n-variable, m-resilient, degree d function is

$$2^{n-1} - 2^{\frac{n}{2}-1} - 2^{m+1+\lfloor \frac{n-m-2}{d} \rfloor}.$$

For the $(10, 1, 6, -)$ function, $n = 10$, $m = 1$, $d = 6$, and hence the maximum possible nonlinearity is $2^9 - 2^4 - 2^3 = 488$. Thus, the nonlinearity 488 is the maximum possible nonlinearity of any $(10, 1, 6, \quad)$ function. Hence, this result in turn shows the tightness of the upper bound on nonlinearity obtained from weight divisibility results [4], [5].

Note that the construction of Theorem 4 provides functions with algebraic degree 6. We modify the construction in Theorem 4 to construct 1-resilient n-variable $(n \geq 10)$ functions with maximum possible algebraic degree $(n - 2)$. The function $(8, 1, 6, 116)$ is already optimized with respect to algebraic degree.

Next, we turn our attention back to the construction of 1-resilient degree optimized functions.

Theorem 5: For even $n \geq 10$, it is possible to construct $(n, 1, n-2, 2^{n-1} - 2^{\frac{n}{2}} + 2^{\frac{n}{2}-2} - 4)$ functions.

Proof: The proof follows from Theorems 1 and 4. Here we consider the minimum possible nonlinearity. □

Remark 1: Using Theorem 5, it is possible to get a $(10, 1, 8, 484)$ function. A result with the same quality is available in [17, Theorem 9]. Now it is important to note that, in the proof of Theorem 5, we mentioned that this is the least possible value of nonlinearity in our construction. In fact, we exhausted all possibilities by changing two pairs of valid positions as in the proof of Theorem 1 for a $(10, 1, 6, 488)$ function and found that the nonlinearity stays unchanged at 488 in some cases. Thus, we could construct $(10, 1, 8, 488)$ functions, one of which is completely described as follows:

```
6F4FC675EE280B7135159C4BB472512B
6F4FC635EE280B7135159C4BB472512B
6F4FC635EE280B7135159C4BB472512B
90B0398A11D7F48ECAEA63B44B8DAED4
CA8A932DD2E4A84D90D0C977C8BEF217
CA8A932DD2E4A84D90D0C977CABEF217
CA8A932DD2E4A84D90D0C977CABEF217
35756CD22D1B57B26F2F368837410DE8.
```

This is a better result than [17, Theorem 9]. In general, there is a possibility of getting $(n, 1, n-2, 2^{n-1} - 2^{\frac{n}{2}} + 2^{\frac{n}{2}-2})$ or $(n, 1, n-2, 2^{n-1} - 2^{\frac{n}{2}} + 2^{\frac{n}{2}-2} + 4)$ functions, if one can select proper positions.

Note that in [17], 1-resilient functions with nonlinearity $2^{n-1} - 2^{\frac{n}{2}} + 2^{\frac{n}{2}-2}$ could be achieved for $n \geq 12$. However, the maximum possible algebraic degree for that construction is $\frac{n}{2} + 2$ [17, Theorem 7], which is much less than the optimized degree $(n-2)$. Here we concentrate on functions with optimized degree only.

For even $n \geq 12$, the currently best known nonlinearity achieved by 1-resilient functions [17, Theorem 8] with maximum possible algebraic degree $(n-2)$ is $2^{n-1} - 2^{\frac{n}{2}} + y$, where y is the maximum possible nonlinearity of an $(\frac{n}{2} - 1)$-variable 1-resilient function with algebraic degree $(\frac{n}{2} - 3)$. We estimate y as $2^{\frac{n}{2}-2} - 2^{\frac{n}{4}-2} - 4$ [18], the upper bound of nonlinearity for an $(\frac{n}{2} - 1)$-variable function which is 1-resilient. So, the currently best known nonlinearity achieved by 1-resilient functions [17] with maximum possible algebraic degree $(n-2)$ is $2^{n-1} - 2^{\frac{n}{2}} + 2^{\frac{n}{2}-2} - 2^{\frac{n}{4}-2} - 4$.

We here achieve the nonlinearity $2^{n-1} - 2^{\frac{n}{2}} + 2^{\frac{n}{2}-2} - 4$ for 1-resilient functions with maximum possible algebraic degree $(n-2)$.

Hence, considering the functions with maximum possible algebraic degree, we find that for even $n \geq 12$, the nonlinearity achieved in this paper for 1-resilient Boolean functions (algebraic degree $n-2$) is $2^{n-1} - 2^{\frac{n}{2}} - 2^{\frac{n}{2}-2} - 4$, which is strictly greater than the nonlinearity $2^{n-1} - 2^{\frac{n}{2}} + 2^{\frac{n}{2}-2} - 2^{\frac{n}{4}-2} - 4$ achieved in [17] for 1-resilient Boolean functions (algebraic degree $n-2$). The constructions $(8, 1, 6, 116)$ and $(10, 1, 8, 488)$ provides better nonlinearity than [17] for the cases $n = 8, 10$ and also these are placed as important open questions in [18].

VI. Functions With Order of Resiliency $m > 1$

Here we analyze an existing construction technique [22], [15] for generating resilient functions on a certain number of variables from functions on a lower number of variables. Note that the construction of [22] has been further extended in [23], [24]. So far, the basic construction [22], [15] has been utilized in construction of resilient functions with high orders $(m > \frac{n}{2} - 2)$. Here, we utilize the construction for low order of resiliency as well.

This construction was first proposed in [22] and then modified in [15]. The construction provides $(n+3)$-variable, $(m+2)$-resilient functions from n-variable, m-resilient initial functions. In [22], the requirement was two initial functions with some specific properties and this was later modified in [15], where the requirement was only one initial function. This initial function [15] has to be in *desired* form. Let us now start with the definition and basic construction method [15].

Definition 6: An $(n, m, d, -)$ function f is in *desired* form if it is of the form $f = (1 + X_n)f_1 + X_n f_2$, where f_1, f_2 are $(n-1, m, d-1, -)$ functions.

Construction 1 [15]: Let f be an (n, m, d, x) function in *desired* form, where f_1, f_2 are both $(n-1, m, d-1, -)$ functions. Let

$$F = f \| \overline{f} \| \overline{f} \| f$$

or written in ANF

$$F = X_{n+2} + X_{n+1} + f.$$

Let

$$G = g \| h \| \overline{h} \| \overline{g}, \qquad \text{where } g = f_1 \| \overline{f}_1 \text{ and } \quad h = f_2 \| \overline{f}_2.$$

In ANF, the function G is given by

$$G = (1 + X_{n+2} + X_{n+1})f_1 + (X_{n+2} + X_{n+1})f_2 + X_{n+2} + X_n.$$

In the language of [22], the function G above is said to depend quasi-linearly on the pair of variables (X_{n+2}, X_{n+1}). We construct a function H in $n+3$ variables in the following way:

$$H = (1 + X_{n+3})F + X_{n+3}G.$$

Then the function H constructed from f is an $(n+3, m+2, d+1, 2^{n+1} + 4x)$ function in the *desired* form.

For the functions F, G, we have $NZ(F) \cap NZ(G) = \emptyset$, i.e., the set of nonzero Walsh coefficients of the constituent functions of H are disjoint. This has been clearly mentioned in [15]. However, it is not very clear what is the relationship among $NZ(f_1)$, $NZ(f_2)$, where f_1, f_2 are the constituent functions of the initial function f. We will look into this problem in the following.

Proposition 3: Let $m > \frac{n}{2} - 2$ and f be an $(n, m, n-m-1, 2^{n-1} - 2^{m+1})$ function in *desired* form. Also, $f = (1 + X_n)f_1 + X_n f_2$, where f_1, f_2 are $(n-1, m, n-m-2, -)$ functions. Then $NZ(f_1) \cap NZ(f_2) = \emptyset$.

Proof: Note that,

$$nl(f_1) \geq nl(f) \quad 2^{n-2} = 2^{n-2} \quad 2^{m+1}.$$

Given the upper bound on nonlinearity of a resilient function [18], $nl(f_1) \leq 2^{n-2} - 2^{m+1}$. Thus, $nl(f_1) = 2^{n-2} - 2^{m+1}$. The nonlinearity for f_2 is the same. Also from [18] we know that the Walsh spectra of f, f_1, f_2 are three-valued $0, \pm 2^{m+2}$.

Now consider $NZ(f_1) \cap NZ(f_2) \neq \emptyset$. Then there exists an $(n-1)$-variable linear function l (may be degenerate) such that $wd(f_1, l) \neq 0$ and also $wd(f_2, l) \neq 0$. If both $wd(f_1, l)$, $wd(f_2, l)$ are of the same sign, then the Walsh spectra of f will contain the value $+2^{m+3}$ or -2^{m+3}, which is a contradiction. If $wd(f_1, l)$, $wd(f_2, l)$ are of different sign, then without loss of generality, consider $wd(f_1, l) = 2^{m+2}$ and $wd(f_2, l) = 2^{m+2}$. Now

$$wd(f, X_n + l) = wd(f_1, l) - wd(f_2, l) = 2^{m+3}$$

which is again not possible. □

The above proposition guarantees that if we get an n-variable, m-resilient $(m > \frac{n}{2} - 2)$ Boolean function f with maximum possible nonlinearity in *desired* from, then the constituent functions f_1, f_2 (which are $(n-1)$-variable, m-resilient functions with maximum possible nonlinearity) will satisfy the condition $NZ(f_1) \cap NZ(f_2) = \emptyset$.

However, Proposition 3 does not always hold for the functions where the Walsh spectrum is not three-valued. This happens in the following two cases.

1) Consider the range $m > \frac{n}{2} - 2$, where the functions do not possess the maximum possible nonlinearity. Then the Walsh spectrum is not three-valued.

2) Consider the range $m \leq \frac{n}{2} - 2$, and the functions may or may not possess the maximum possible nonlinearity. Here

the Walsh spectrum is not three-valued irrespective of the value of nonlinearity, whether it is maximum possible or not.

There is also a fundamental difference in the resulting functions out of this construction for different order of resiliency.

1) For $m > \frac{n}{2} - 2$, if one can start with a *desired* function with maximum possible nonlinearity, then the construction, when repeatedly used, will generate an infinite sequence of functions which are of maximum possible nonlinearity.

2) For $m \leq \frac{n}{2} - 2$, even if one can start with a *desired* function with maximum possible nonlinearity, then the construction, when repeatedly used, will generate an infinite sequence of functions which are not guaranteed to be of maximum possible nonlinearity.

We also like to clarify the issue of the nonlinearity of f_1, f_2, the constituent functions of f. Given the (n, m, d, x) function f, we get, both $nl(f_1), nl(f_2) \geq x - 2^{n-2}$. Now we have the following cases.

1) If the function f achieves the maximum possible nonlinearity and $m > \frac{n}{2} - 2$, then we have $nl(f_1), nl(f_2) = x - 2^{n-2}$.
2) For $m > \frac{n}{2} - 2$, if the function f does not possess maximum possible nonlinearity, then it is possible that either of f_1, f_2 may have nonlinearity strictly greater than $x - 2^{n-2}$.
3) For $m \leq \frac{n}{2} - 2$, it is possible that either f_1 or f_2 may have nonlinearity strictly greater than $x - 2^{n-2}$ irrespective of whether it achieves the maximum possible nonlinearity or not.

Next we discuss the application of Construction 1 in different cases.

A. Use of Desired Form

To use Construction 1 on $(8, 1, 6, 116)$ functions, we need a function in *desired form*. One way is to get an $(8, 1, 6, 116)$ function as the concatenation of two $(7, 1, 5, 52)$ functions f_1, f_2. However, this is not automatically guaranteed. For example, the $(8, 1, 6, 116)$ function presented in Section III-B is the concatenation of two $(7, 0, 5, 52)$ functions, not $(7, 1, 5, 52)$ functions.

We searched the database of $(8, 1, 6, 116)$ functions and generate functions with permutations of input variables. Note that these functions are also $(8, 1, 6, 116)$ functions. We check whether these new functions are in *desired* form. The following $(8, 1, 6, 116)$ function is in *desired* form with nonintersecting Walsh spectrum of its constituent $(7, 1, 5, 52)$ functions. The truth table of the function in hexadecimal format is

A0BE6E1D8191B45FFA8289477E76D205
6C72A6D1B2A3876C364E658B4DC5E136.

This is the way, we found $(8, 1, 6, 116)$ functions in desired form. Thus, using Construction 1, we get an $(11, 3, 7, 976)$ function. Note that the previously best known nonlinearity for an $(11, 3, 7, -)$ function is 960 [17, Sec. 6].

It is of interest to mention the following $(8, 1, 6, 116)$ function f we have found. Here, $NZ(f_1) \cap NZ(f_2) \neq \emptyset$, f_1 is a $(7, 1, 5, 52)$ function, and f_2 is $(7, 1, 5, 56)$. This $(8, 1, 6, 116)$ function is in *desired* form with intersecting Walsh spectrum of its constituent functions. The function is

CF6730989E329A3A4C795C59190CE6F3
AC592DB1F90C871B309FD8E165CA724B.

We list the functions in the series provided by repeated use of Construction 1. These functions are $(8, 1, 6, 116)$, $(11, 3, 7, 976)$, $(14, 5, 8, 8000)$, $(17, 7, 9, 64768)$, ..., etc. Although these functions do not attain maximum possible nonlinearity, the nonlinearity values are superior to all previous constructions.

Next we concentrate on the functions provided in [17, Theorem 10b] and Construction 1 [22], [15] given above to get highly nonlinear resilient functions with odd order. This is the first time where Construction 1 is used for the functions with $m \leq \frac{n}{2} - 2$ in a generalized manner.

Theorem 6: It is possible to construct $(n, m, n - m - 1, 2^{n-1} - 2^{\frac{3n-m-3}{4}})$ functions when $2n - 3m \equiv 3 \bmod 4$.

Proof: We start with a p-variable (p odd) function f provided in [17, Theorem 10b] with order of resiliency 1 and nonlinearity $2^{p-1} - 2^{\frac{p-1}{2}}$. It is clear that the functions provided in [17, Theorem 10b] are in *desired* form as each of f_1, f_2 are also 1-resilient functions. Also, the algebraic degree of the function f is $p - 2$.

Thus, given some odd $m > 1$, we will calculate the value $t = \frac{m-1}{2}$. As Construction 1 increases the order of resiliency by two in each step and the number of variables by three in each step, we have to start with a function on $p = n - 3t$ variables and order of resiliency 1. We will then recursively use Construction 1 t times.

Now we will show that this recursive series of constructions always yield $(n_i, m_i, n_i - m_i - 1, 2^{n_i-1} - 2^{\frac{3n_i+m_i-3}{4}})$ functions. For $i = 0$, we take $n_0 = p$, $m_0 = 1$. This satisfies the nonlinearity

$$2^{n_i - 1} - 2^{\frac{2n_i - m_i - 3}{4}} = 2^{p-1} - 2^{\frac{p-1}{2}}.$$

This is the base case of the induction. Now assume that this is true for all i less than or equal to j. Thus, we can construct$(n_j, m_j, n_j - m_j - 1, 2^{n_j-1} - 2^{\frac{3n_j+m_j-3}{4}})$ functions. By applying Construction 1 on one such function we get a

$$\left(n_j - 3, m_j + 2, n_j - m_j, 2^{n_j+2} + 4\left(2^{n_j-1} - 2^{\frac{3n_j+m_j-3}{4}}\right)\right)$$

function. Note that $n_{j+1} = n_j + 3$ and $m_{j+1} = m_j + 2$. Now

$$\begin{aligned} 2^{n_j+1} + 4\left(2^{n_j-1} - 2^{\frac{3n_j+m_j-3}{4}}\right) &= 2^{n_j+2} - 2^{\frac{3n_j+m_j-3}{4}+2} \\ &= 2^{n_j+2} - 2^{\frac{3(n_j+3)-(m_j+2)-3}{4}} \\ &= 2^{n_{j+1}-1} - 2^{\frac{3n_{j+1}-m_{j+1}-3}{4}}. \end{aligned}$$

This proves the inductive step.

Now we have $t = \frac{m-1}{2}$ and $p = n - 3t$. Note that p is odd. Thus, $n - \frac{3m-3}{2}$ is odd, which gives $2n - 3m + 3 \equiv 2 \bmod 4$, i.e., $2n - 3m \equiv 3 \bmod 4$. □

As an immediate corollary we get the following result putting $m = 3$.

Corollary 1: For even $n \geq 8$, it is possible to construct $(n, 3, n-4, 2^{n-1} - 2^{\frac{n}{2}})$ functions which are in *desired* form.

Note that in [17, Theorem 10d], $(n, 2, n-3, 2^{n-1} - 2^{\frac{n}{2}})$ functions have been reported. Our result provides the same nonlinearity with order of resiliency 3. In [18], constructions of $(8, 3, 4, 112)$ and $(10, 3, 6, 480)$ functions have been given. These are special cases of Corollary 1. For $n = 12$, we get a $(12, 3, 8, 1984)$ function. This is better than the $(12, 3, 8, 1968)$ function given in [17, Sec. 6].

It is important to see that we can go on repeating Construction 2 and we will get a sequence of *desired* functions. These functions will initially start with the functions with $m \leq \frac{n}{2} - 2$ and then in the latter part of the sequence they land into $m > \frac{n}{2} - 2$. As an example, consider the sequence $(10, 3, 6, 480)$, $(13, 5, 7, 3968)$, $(16, 7, 8, 32\,256)$, $(19, 9, 9, 260\,096)$, ..., where though the nonlinearities are not achieving the upper bound, they provide a very interesting construction of suboptimal functions with very high nonlinearity, which were not known earlier. Also these functions do not have three-valued Walsh spectra.

VII. Conclusion

In this contribution, we present different construction ideas for resilient Boolean functions with very high nonlinearity. We mainly concentrate here on low order of resiliency. As specific examples for small numbers of input variables, we are able to construct $(8, 1, 6, 116)$, $(10, 1, 6, 488)$, $(10, 1, 8, 488)$, $(11, 3, 7, 976)$, and $(12, 3, 8, 1984)$ functions which were previously not known. We also provide some generalized construction methods for functions on higher number of variables. Our results either supersede or solve important open questions posed in the literature. Moreover, these Boolean functions have immediate application in the design of stream cipher systems.

References

[1] E. R. Berlekamp and L. R. Welch, "Weight distributions of the cosets of the (32, 6) Reed–Muller code," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 203–207, Jan. 1972.

[2] P. Camion, C. Carlet, P. Charpin, and N. Sendrier, "On correlation immune functions," in *Advances in Cryptology—CRYPTO'91 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1992, vol. 576, pp. 86–100.

[3] A. Canteaut and M. Trabbia, "Improved fast correlation attacks using parity-check equations of weight 4 and 5," in *Advances in Cryptology—EUROCRYPT 2000 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer Verlag, 2000, vol. 1807, pp. 573–588.

[4] C. Carlet, "On the coset weight divisibility and nonlinearity of resilient and correlation immune functions," in *Sequences and Their Applications—SETA 2001 (Discrete Mathematics and Theoretical Computer Science)*. Berlin, Germany: Springer Verlag, 2001, pp. 131–144.

[5] C. Carlet and P. Sarkar, "Spectral domain analysis of correlation immune and resilient Boolean functions," *Finite Fields Its Applic.*, to be published.

[6] S. Chee, S. Lee, D. Lee, and S. H. Sung, "On the correlation immune functions and their nonlinearity," in *Advances in Cryptology—ASIACRYPT '96 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1996, vol. 1163, pp. 232–243.

[7] C. Ding, G. Xiao, and W. Shan, *The Stability Theory of Stream Ciphers (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1991, vol. 561.

[8] E. Filiol and C. Fontaine, "Highly nonlinear balanced Boolean functions with a good correlation-immunity," in *Advances in Cryptology—EUROCRYPT'98 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1998, vol. 1403, pp. 475–488.

[9] X. Guo-Zhen and J. Massey, "A spectral characterization of correlation immune combining functions," *IEEE Trans. Inform. Theory*, vol. 34, pp. 569–571, May 1988.

[10] T. Johansson and F. Jonsson, "Fast correlation attacks through reconstruction of linear polynomials," in *Advances in Cryptology—CRYPTO 2000 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer Verlag, 2000, vol. 1880, pp. 300–315.

[11] S. Maitra. (2000, Oct.) Correlation immune Boolean functions with very high nonlinearity. *Cryptology ePrint Archive* [Online]. Available: eprint.iacr.org

[12] W. Meier and O. Staffelbach, "Fast correlation attack on stream ciphers," in *Advances in Cryptology—EUROCRYPT'88 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, May 1988, vol. 330, pp. 301–314.

[13] J. J. Mykkeltveit, "The covering radius of the (128, 8) Reed–Muller code is 56," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 358–362, May 1983.

[14] E. Pasalic and T. Johansson, "Further results on the relation between nonlinearity and resiliency of Boolean functions," in *Proc. IMA Conf. Cryptography and Coding (Lecture Notes in Computer Science)*. New York: Springer-Verlag, 1999, vol. 1746, pp. 35–45.

[15] E. Pasalic, S. Maitra, T. Johansson, and P. Sarkar, "New constructions of resilient and correlation immune Boolean functions achieving upper bounds on nonlinearity," in *Workshop on Coding and Cryptography—WCC 2001*, Paris, France, Jan. 8–12, 2001. Published in *Electronic Notes in Discrete Mathematics*. Amsterdam, The Netherlands: Elsevier Science, vol. 6, 2001.

[16] O. S. Rothaus, "On bent functions," *J. Comb. Theory, Ser. A*, vol. 20, pp. 300–305, 1976.

[17] P. Sarkar and S. Maitra, "Construction of nonlinear Boolean functions with important cryptographic properties," in *Advances in Cryptology—EUROCRYPT 2000 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer Verlag, 2000, vol. 1807, pp. 485–506.

[18] ——, "Nonlinearity bounds and constructions of resilient Boolean functions," in *Advances in Cryptology—CRYPTO 2000 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer Verlag, 2000, vol. 1880, pp. 515–532.

[19] J. Seberry, X. M. Zhang, and Y. Zheng, "On constructions and nonlinearity of correlation immune Boolean functions," in *Advances in Cryptology—EUROCRYPT'93 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1994, vol. 765, pp. 181–199.

[20] T. Siegenthaler, "Correlation-immunity of nonlinear combining functions for cryptographic applications," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 776–780, Sept. 1984.

[21] ——, "Decrypting a class of stream ciphers using ciphertext only," *IEEE Trans. Comput.*, vol. C-34, pp. 81–85, Jan. 1985.

[22] Y. V. Tarannikov, "On resilient Boolean functions with maximum possible nonlinearity," in *Progress in Cryptology—INDOCRYPT 2000 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer Verlag, 2000, vol. 1977, pp. 19–30.

[23] ——, "New constructions of resilient Boolean functions with maximal nonlinearity," in *Fast Software Encryption—FSE 2001 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2001, preproceedings, pp. 70–81, to be published.

[24] M. Fedorova and Y. V. Tarannikov, "On the constructing of highly nonlinear resilient Boolean functions by means of special matrices," in *Progress in Cryptology—INDOCRYPT 2001 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2001, vol. 2247 , pp. 254–266.

[25] Y. Zheng and X. M. Zhang, "Improved upper bound on the nonlinearity of high order correlation immune functions," in *Selected Areas in Cryptography—SAC 2000 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2000, vol. 2012 , pp. 264–274.

[26] ——, "New results on correlation immune functions," in *Proc. Int. Conf. Information Security and Cryptology—ICISC 2000 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2001, vol. 2015, pp. 49–63.