Highly Nonlinear Resilient Functions Optimizing Siegenthaler's Inequality

Subhamoy Maitra[1] and Palash Sarkar[2]

[1] Computer and Statistical Service Center, Indian Statistical Institute
203, B T Road, Calcutta 700 035, India
subho@isical.ac.in
[2] Applied Statistics Unit, Indian Statistical Institute
203, B T Road, Calcutta 700 035, India
palash@isical.ac.in

Abstract. Siegenthaler proved that an n input 1 output, m-resilient (balanced mth order correlation immune) Boolean function with algebraic degree d satisfies the inequality : $m + d \leq n - 1$. We provide a new construction method using a small set of recursive operations for a large class of highly nonlinear, resilient Boolean functions optimizing Siegenthaler's inequality $m + d = n - 1$. Comparisons to previous constructions show that better nonlinearity can be obtained by our method. In particular, we show that as n increases, for almost all m, the nonlinearity obtained by our method is better than that provided by Seberry et al in Eurocrypt'93. For small values of n, the functions constructed by our method is better than or at least comparable to those constructed using the methods provided in papers by Filiol et al and Millan et al in Eurocrypt'98. Our technique can be used to construct functions on large number of input variables with simple hardware implementation.
Keywords: *Stream Cipher, Boolean Function, Algebraic Degree, Correlation Immunity, Nonlinearity, Balancedness.*

1 Introduction

In stream cipher cryptography, the message is considered to be a stream of bits. The cipher is obtained by bitwise XORing (addition over GF(2)) the message with a sequence of bits called the key stream. In most common models of stream ciphers the key stream is produced by using a Boolean function to combine the output sequences of several Linear Feedback Shift Registers (LFSRs). If the combining Boolean function is not properly chosen, then the system becomes susceptible to several kinds of cryptanalytic attacks. An important class of *divide-and-conquer* attacks on such systems was proposed by Siegenthaler [18]. Moreover, Siegenthaler [17] himself introduced a class of Boolean functions, the set of correlation immune functions, which can resist such attacks. However, it is not sufficient to use functions with only correlation immunity, since certain types of correlation immune functions are susceptible to other kinds of attacks. For example, it is well known that the linear functions are correlation immune

Michael Wiener (Ed.): CRYPTO'99, LNCS 1666, pp. 198–215, 1999.

but not suitable for use in cryptography. There are two measures for nonlinearity of Boolean functions. The algebraic degree is the degree of the algebraic normal form of a Boolean function. Having a high algebraic degree ensures a high linear complexity of the produced key stream and hence better immunity against the Berlekamp Massey shift register synthesis algorithm [9]. A second measure of nonlinearity is the distance from the set of affine functions. A high value of this parameter ensures that the best affine approximation [4] attack will fail. Siegenthaler in [17] proved a fundamental inequality relating the number of variables n, order of correlation immunity m and algebraic degree d of a Boolean function: $m + d \leq n$. Moreover, if the function is balanced then $m + d \leq n - 1$. Also, a balanced mth order correlation immune function is said to be m-resilient. Since it is natural to use balanced functions in stream cipher systems we concentrate only on resilient functions. A resilient Boolean function is said to be optimized if $m + d = n - 1$. The maximum possible nonlinearity (distance from the set of linear functions) for this class of functions is not known. Here we provide construction methods for optimized functions having high nonlinearities. The functions are built using a small set of recursive operations and hence functions on large number of variables are easy to implement using nominal hardware.

Construction procedures for correlation immune (CI) functions were first described by Siegenthaler in [17]. The methods described in [17] are recursive, where a function of $(n + 1)$ variables is built from two functions of n variables. Siegenthaler considered two different kinds of constructions, one where the order of correlation immunity remains constant and the other where the order of correlation immunity increases by one at each step. An important spectral characterization of correlation immunity, based on *Walsh transform* of a Boolean function, was given in [6].

Further attempts at construction was made by Camion et al. in [1], where construction procedure for a certain subset of correlation immune functions were described. In [2], the construction procedure for bent functions is modified to get correlation immune functions. Seberry et al. [16], also provided a method of constructing the same subset as in [1] of correlation immune functions. They also separately considered the algebraic degree, nonlinearity and propagation characteristics of their construction method. The functions constructed in [16] has good nonlinearity for non optimized functions. However, for optimized functions the nonlinearity of [16] decreases. We interpret the *direct construction* method proposed in [16] in a simpler manner (see Section 5) as a concatenation of linear functions. This interpretation simplifies the proofs related to correlation immunity and nonlinearity.

Evolutionary techniques are applied in [11] to design first order correlation immune balanced functions with high nonlinearity. The technique considers the output column of the function as a string and applies genetic algorithm to manipulate this string. Therefore this technique is difficult to apply to construct functions on n variables for even moderate values of n. Moreover, it is not clear whether these functions optimize the Siegenthaler's inequality. To be precise, by relaxing the optimization criterion of the Siegenthaler's inequality, we can

achieve better nonlinearity than in [11]. Favorable results can also be found using construction procedure in [16].

In another approach to the problem, Filiol and Fontaine [5, Section 5] describe a method to construct functions which achieve a good trade-off between nonlinearity, balancedness, degree and correlation immunity. They identify a 7 variable function f with nonlinearity 56 and degree 6. Using f, in [5, Section 5], they construct a balanced 9 variable function g with nonlinearity 224, correlation immunity of order 2 and degree 6, where they use a technique which was first introduced in [17, Section VI], and later in [1, Corollary 4.1]. The function g is optimized with respect to Siegenthaler's inequality. *The function g is so far the best known optimized function on 9 variables with correlation immunity of order 2.* The key of this construction is the existence of f.

We use concatenation techniques and introduce generic construction functions (see Definition 5), which recursively build a correlation immune function of $(n+1)$ variables from two correlation immune functions of n variables. We initially start with bent functions which are modified a little to get optimized algebraic degree. A sequence of such constructors is applied to build correlation immune functions of desired orders from non correlation immune balanced Boolean functions with high nonlinearity. The degree of the resulting function is same as that of the initial function. The method can easily be extended to design functions with moderate to large number of input variables using a special representation of the constructed Boolean functions (see Definition 6). The actual trade-off between nonlinearity and correlation immunity is explicit (see Theorem 11). Also Theorem 11 provides a lower bound on the nonlinearity of a function optimized with respect to Siegenthaler's inequality [17].

Both our technique as well as the technique of [16] can be used to construct highly nonlinear, balanced, n variable, mth order correlation immune (m resilient) functions having algebraic degree $n-m-1$. We show that for all m such that $m + 2\log_2(m+3) + 3 < n$, the nonlinearity obtained by our method for **optimized functions** is better than that of [16]. Thus as n increases, for almost all m, we obtain a better nonlinearity. Conversely, if we fix an m, then there exists an N, such that for all $n \geq N$, the nonlinearity obtained by our method is better. As examples, using our techniques one can construct
1. 10 variable balanced functions with degree 8, order of correlation immunity 1 and nonlinearity 476 and
2. 50 variable balanced functions with degree 20, order of correlation immunity 29 and nonlinearity $2^{49} - 2^{39} - 2^{30}$.
None of the currently known methods can be used to construct such optimized functions. Moreover, there are widely different functions in the constructed class (see Example 1 in Section 6).

Next we provide a list of notations.
1. For strings S_1, S_2 of 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 as $D(S_1, S_2)$, i.e. $D(S_1, S_2) = \#(S_1 \neq S_2)$. The *Walsh Distance* is defined as, $wd(S_1, S_2) =$

$\#(S_1 = S_2) - \#(S_1 \neq S_2)$. Note that, $wd(S_1, S_2) = \lambda - 2\,D(S_1, S_2)$. Also the *Hamming weight* or simply the weight (number of 1s in S) of S is denoted as $wt(S)$.

2. By Ω_n, we denote the set of all Boolean functions on n variables, i.e., the set of all binary strings of length 2^n. If $f, g \in \Omega_{n-1}$, then $F = fg$ is a function in Ω_n whose output column is the concatenation of the output columns of f and g. Given the truth table of a function f of n input variables $\{X_1, X_2, \ldots, X_n\}$, we also interpret f as a binary string of length 2^n, the output column of the truth table. The first half of the string f is denoted as f^u and the second half is denoted as f^l. If $f \in \Omega_n$, then $f^u, f^l \in \Omega_{n-1}$ and are given by $f^u(X_{n-1}, \ldots, X_1) = f(0, X_{n-1}, \ldots, X_1)$ and $f^l(X_{n-1}, \ldots, X_1) = f(1, X_{n-1}, \ldots, X_1)$. In the truth table the column corresponding to an input variable X_j occurs to the left of the column corresponding to the input variable X_i, if $j > i$. Note that a function $f \in \Omega_n$ may be a non degenerate function of i variables for $i < n$.

3. The reverse of the string S is denoted by S^r. The bitwise complement of a string S is denoted as S^c. If f is a Boolean function, then f^r, the function obtained by reversing the output column of the truth table is given by $f^r(X_n, \ldots, X_1) = f(1 \oplus X_n, \ldots, 1 \oplus X_1)$, where $\oplus$ denotes the XOR operation. Similarly, the function f^c obtained by complementing each bit of the output column of f is given by $f^c(X_n, \ldots, X_1) = 1 \oplus f(X_n, \ldots, X_1)$.

We next define the important cryptographic properties of Boolean functions for stream cipher applications. These also appear in [5,16,12,17].

Definition 1. *A Boolean function f of n variables is said to be linear/affine if f can be expressed as $f = \bigoplus_{i=1}^{n} a_i X_i \oplus b$, where $a_i, b \in \{0, 1\}$ for all i. The set of linear/affine functions of n variables is denoted as $L(n)$. A Boolean function f of n variables is said to be nonlinear if f is not linear/affine. We denote the measure of nonlinearity of an n variable function f as $nl(f) = min_{g \in L(n)}(D(f, g))$.*

Note that $L(n) = \{H \mid H = hh \text{ or } hh^c, h \in L(n-1)\}$. Let $h \in L(n)$ be a non degenerate function of $m\,(1 \leq m \leq n)$ variables. If m is even then $h^r = h$ else if m is odd, $h^r = h^c$. The linear function $h \in L(n)$ is degenerate if $m < n$. A high nonlinearity ensures that the best affine approximation cryptanalytic attack will fail. (See [4] for a description of this method). It is known [13] that for even n, the maximum nonlinearity achievable by a Boolean function is $nl(f) = 2^{n-1} - 2^{\frac{n}{2}-1}$. Such functions are called bent functions and their combinatorial properties have been studied [3,4,13]. A simple construction method for bent functions from [13] is $h(X_1, \ldots, X_p, Y_1, \ldots, Y_p) = \bigoplus_{i=1}^{p} X_i Y_i \oplus g(Y_1, \ldots, Y_p)$ where $g \in \Omega_p$ is arbitrary. For odd n, the corresponding class of functions have not been characterized. (See [15] for some best known examples). Moreover, bent functions are known to be unbalanced and are not correlation immune. Meier and Staffelbach [10] have described a procedure to construct balanced nonlinear functions from bent functions. So if one is looking for functions which optimize Siegenthaler's inequality, one cannot hope to attain the maximum value of $nl(f)$.

Another important criterion is algebraic degree, since it determines the linear complexity of the output sequence of the function (see [4]). The relationship of algebraic degree to the order of correlation immunity was studied in [17,6].

Definition 2. *The algebraic degree or simply the degree of $f \in \Omega_n$, denoted by $deg(f)$, is defined to be the degree of the algebraic normal form of f. See [17] for definition of algebraic normal form and its degree.*

Siegenthaler [17] was the first to define correlation immune functions from information theoretic point of view using the concept of mutual information.
A function $f(X_1, X_2, \ldots, X_n)$ is mth order correlation immune [17] if the mutual information $I(X_{i_1}, X_{i_2}, \ldots, X_{i_m}; Z) = 0$ for all possible choices of m distinct variables $X_{i_1}, X_{i_2}, \ldots, X_{i_m} \in \{X_1, X_2, \ldots, X_n\}$, with $1 \leq m \leq n-1$. From [7], this is equivalent to $Prob(Z = 1 \mid X_{i_1} = C_{i_1}, X_{i_2} = C_{i_2}, \ldots, X_{i_m} = C_{i_m}) = Prob(Z = 1)$ for each of the combinations $C_{i_1}, C_{i_2}, \ldots, C_{i_m} \in \{0, 1\}$.
A characterization of correlation immunity based on Walsh transform of Boolean functions was obtained in [6]. We first provide the definition of Walsh transform.

Definition 3. *Let $\overline{X} = (X_1, \ldots, X_n)$ and $\overline{\omega} = (\omega_1, \ldots, \omega_n)$ be n-tuples on $GF(2)$ and $\overline{X}.\overline{\omega} = X_1\omega_1 \oplus \ldots \oplus X_n\omega_n$. Let $f(\overline{X})$ be a Boolean function whose domain is the vector space over $GF(2)^n$. Then the Walsh transform of $f(\overline{X})$ is a real valued function over $GF(2)^n$ that can be defined as $F(\overline{\omega}) = \sum_{\overline{X}} (-1)^{f(\overline{X}) \oplus \overline{X}.\overline{\omega}}$,*

where the sum is over all $\overline{X}$ in $GF(2)^n$.

The following result provides the relationship between Walsh distance and Walsh transform.

Proposition 1. $F(\overline{\omega}) = wd(f, \bigoplus_{i=1}^{i=n} \omega_i X_i)$.

The following characterization of correlation immunity, based on Walsh transform, was given in [6].

Theorem 1. *([6]) A function $f(X_n, X_{n-1}, \ldots, X_1)$ is mth order correlation immune iff its Walsh transform F satisfies $F(\overline{\omega}) = 0$, for $1 \leq wt(\overline{\omega}) \leq m$.*

Proposition 2. *Let $h, f \in \Omega_n$. Then (a) $wd(h, f) = -wd(h^c, f)$ and (b) $wd(h, f^r) = wd(h^r, f)$. Consequently, $wd(h, f) = 0$ iff $wd(h, f^c) = 0$.*

We use the following definition of correlation immunity which follows from Proposition 1, Theorem 1 and Proposition 2.

Definition 4. *A function $f(X_n, X_{n-1}, \ldots, X_1)$ is said to be mth $(1 \leq m \leq n-1)$ order correlation immune if $wd(f, h) = 0$ where $h \in L(n)$ and h is a non degenerate function of i variables with $1 \leq i \leq m$. Moreover, if f is balanced then f is called m-resilient.*

From this definition it is clear that *if a function is mth order correlation immune, then it is kth order correlation immune for* $1 \leq k \leq m$. We define,

1. $C_n(m) = \{f \in \Omega_n \mid f$ is correlation immune of order m but not correlation immune of order $m+1\}$.
2. $A_n(m) = \bigcup_{m \leq k \leq n-1} C_n(k)$, is the set of all correlation immune functions of order m or more.
3. A function is called *correlation immune* if it is at least correlation immune of order one. Also $A_n = A_n(1)$, is the set of all correlation immune functions of n variables.

Let f be a balanced function of degree d, and $f \in C_n(m)$. Then f is optimized with respect to balancedness, degree and order of correlation immunity if $m + d = n - 1$. The maximum value of $nl(f)$ for such functions is not known. In Theorem 11 we describe methods to construct such optimized functions with sufficiently large values of $nl(f)$. We next define three constructions P, Q, R as follows. These constructions have also been used in [8] to obtain the currently best known lower bounds on (balanced) correlation immune Boolean functions.

Definition 5. *1.* $P : \Omega_{n-1} \times \Omega_{n-1} \to \Omega_n,\ P(f,g) = f^u g^u g^l f^l$.
2. $Q : \Omega_{n-1} \times \Omega_{n-1} \to \Omega_n,\ Q(f,g) = fg = f^u f^l g^u g^l$.
3. $R : \Omega_{n-1} \times \Omega_{n-1} \to \Omega_n,\ R(f,g) = f^u g^u f^l g^l$.

Later we will use these constructions to recursively build correlation immune functions. The construction Q appears in [17], although in a different form. Note that, the generic construction functions P, Q, R should not be viewed as linear combination of two Boolean functions. As example, if we consider the Boolean function $Q(f, f^r)$, then the nonlinearity of $Q(f, f^r)$ will be twice that of f and the number of terms with highest algebraic degree will increase. We discuss it elaborately in the next section.

2 Nonlinearity, Algebraic Degree, and Balancedness

We provide a few technical results in this section related to nonlinearity, algebraic degree and balancedness.

Theorem 2. *Let* $f, g \in \Omega_{n-1}$ *and* $F = \Psi(f,g)$ *where* $\Psi \in \{P, Q, R\}$. *Then* $nl(F) \geq nl(f) + nl(g)$. *Moreover, if* $g = f$, $g = f^c$ *or* $g = f^r$, *then* $nl(F) = nl(f) + nl(g) = 2nl(f)$.

Next we state without proof the following result on the degree of the constructed function. The proof consists in checking the different cases.

Theorem 3. *Let* $f \in \Omega_n$ *and* $F = \Psi(f, f^\tau)$, *where* $\Psi \in \{P, Q, R\}$ *and* $\tau \in \{c, r\}$. *Then,* $deg(F) = deg(f)$.

The special case of Theorem 3 with $\Psi = Q$ and $\tau = c$ was mentioned in [5]. The importance of this result lies in the fact that the degree of the constructed function is equal to the degree of the original function. It is known [13] that the degree of bent functions of n variables for $n \geq 4$ is at most $\frac{n}{2}$. We propose the following simple but *powerful* method to improve the degree. Note that, $X_1 \dots X_n$ means logical AND of X_1 to X_n.

Theorem 4. *Let $h \in \Omega_n$ be of degree less than n and $f = h \oplus X_1 \dots X_n$. Then $deg(f) = n$ and $nl(f) \geq nl(h) - 1$. Moreover, if h is a bent function then $nl(f) = nl(h) - 1$.*

Proof. Let, $g \in L(n)$. Then $D(f, g)$ is either $D(h, g) - 1$ or $D(h, g) + 1$. If h is bent, $nl(f) \leq nl(h)$, and so $nl(f) = nl(h) - 1$. □

If we start with a bent function $h \in \Omega_8$ and use the above theorem then we can get a function $f \in \Omega_8$ of degree 8 and nonlinearity 119. Using this f, one can get $F = \Psi(f, f^c) \in \Omega_9$, *which is balanced, has degree* 8 *and nonlinearity* 238. Generalizing, we can get balanced functions $F \in \Omega_{2p+1}$ having degree $2p$ and nonlinearity $2^{2p} - 2^p - 2$. We now discuss the following negative results which can be obtained from Siegenthaler's inequality. Let $f \in \Omega_n$.
1. If $deg(f) = n - 1$, then f is not both correlation immune and balanced.
2. If $deg(f) = n$, then f is neither correlation immune nor balanced.
Both 1 and 2 follow from Siegenthaler's inequality $m + d \leq n - 1$ for balanced functions $f \in C_n(m)$ having degree d. To see that if $deg(f) = n$, then f is not balanced, suppose the converse, i.e., $deg(f) = n$ and f is balanced. Since f is balanced, using Theorem 8 in the next section, $Q(f, f^c) = ff^c \in C_{n+1}(1)$. Also, ff^c is balanced and has degree n. Thus, Siegenthaler's inequality is violated for $ff^c \in \Omega_{n+1}$.
Note that item 2 shows that a function of n variables cannot both have degree n and be balanced. Thus it relates two simple properties of Boolean functions. However, it requires the use of correlation immunity, which is a much more specialized property. This shows that there cannot exist balanced functions $F \in \Omega_n$ of degree n. Filiol and Fontaine [5] provided examples of balanced $F \in \Omega_9$ having nonlinearity 240 but degree upto 7. *It is interesting to find out whether there exists balanced $F \in \Omega_9$, having degree* 8 *and nonlinearity* 240.

Theorem 4 shows that the degree can be increased significantly with insignificant change in nonlinearity. Moreover, it can be checked that though f in the above theorem has only one term of degree n, the number of terms of degree n in $\Psi(f, f^r) \in \Omega_{n+1}$ is *more than one*. We state one specific result regarding this.

Proposition 3. *Let $f \in \Omega_n$ with degree n. Then $Q(f, f^r) \in \Omega_{n+1}$ contains n terms of degree n.*

The linear complexity of the output sequence produced by the Boolean function depends on the algebraic normal form of the function and the lengths of the input LFSRs [14,4]. Having more terms of degree n ensures that the linear complexity of the output sequence is higher. See Example 1 in the last section

for further illustration regarding the number of high degree terms. Proper use of this technique will ensure that the functions designed using Construction 1 (see later), will have this property. This has direct implication towards the stability of the generated sequence [4]. We would like to point out that this phenomenon does not hold for the construction $Q(f, f^c)$ given in [17,5]. Next, we list a few simple results on balancedness.

Proposition 4. *(a) A function of the form ff^c is balanced. (b) If f is a balanced function then both f^r and f^c are balanced. (c) Let $f, g \in \Omega_n$ be two balanced functions, and $F = \Psi(f, g)$, where $\Psi \in \{P, Q, R\}$. Then F is also balanced.*

3 Correlation Immunity

Here we provide generalized construction methods for correlation immune functions. First we state the following two results which have been proved in different forms in [12,1] and [17] respectively.

Proposition 5. *Let $h \in \Omega_n$. Then $Q(h, h^r) = hh^r \in A_{n+1}$.*

Proposition 6. *Let $f \in \Omega_n$. Then $Q(f, f^c) = ff^c \in A_{n+1}$ iff f is balanced.*

Next we state without proof the following basic result.

Lemma 1. *Let $f \in A_n(m)$ (respectively $C_n(m)$). Then $f^r, f^c \in A_n(m)$ (respectively $C_n(m)$).*

In [17, Section IV] Siegenthaler proposed a construction of $F \in A_{n+1}(m)$ from $f, g \in A_n(m)$ as follows.

Theorem 5. *([17]) If $Z_1 = f_1(X_1, X_2, \ldots, X_n)$ and $Z_2 = f_2(X_1, X_2, \ldots, X_n)$ are mth-order correlation immune functions of n binary variables such that $Prob(Z_1 = 1) = Prob(Z_2 = 1)$, then the binary-valued function f of $n + 1$ random variables defined by the GF(2) expression*
$f(X_1, X_2, \ldots, X_{n+1}) = X_{n+1} f_1(X_1, X_2, \ldots, X_n) + (X_{n+1} + 1) f_2(X_1, X_2, \ldots, X_n)$
is also mth order correlation immune.

The condition $Prob(Z_1 = 1) = Prob(Z_2 = 1)$ is equivalent to the condition $wt(f_1) = wt(f_2)$. Note that the construction in the above theorem corresponds to our construction Q. We further generalize the construction to include P, R also.

Lemma 2. *Let $f, g \in A_n(m)$ and F be of the form $F = P(f, g) = f^u g^u g^l f^l$. If (a) $m = 1$ or (b) $m > 1$ and $wt(f) = wt(g)$, then $F \in A_{n+1}(m)$.*

Proof. Let f, g be functions of $\{X_1, X_2, \ldots, X_n\}$ and F be a function of $\{X_1, X_2, \ldots, X_{n+1}\}$. We use the characterization of correlation immunity given in Definition 4. Let us consider any linear/affine function $H \in L(n + 1)$, where H is a non degenerate function of k variables $(1 \leq k \leq m)$.

Now we will have four cases.

1. If H contains k variables from $\{X_1, X_2, \ldots, X_{n-1}\}$ then H is of the form $hhhh$. Now, $wd(F,H) = wd(f^u g^u g^l f^l, hhhh) = wd(f, hh) + wd(g, hh) = 0$, as f, g are mth order correlation immune.
2. If H contains X_n and the remaining $k-1$ variables from $\{X_1, X_2, \ldots, X_{n-1}\}$ then H is of the form $hh^c hh^c$. Then, $wd(F,H) = wd(f^u g^u g^l f^l, hh^c hh^c) = wd(f, hh^c) + wd(g, h^c h) = 0$.
3. If H contains X_{n+1} and the remaining $k-1$ variables from $\{X_1, X_2, \ldots, X_{n-1}\}$ then H is of the form $hhh^c h^c$.
   Now, $wd(F,H) = wd(f^u g^u g^l f^l, hhh^c h^c) = wd(f, hh^c) + wd(g, hh^c) = 0$.
4. If H contains X_n, X_{n+1} and the remaining $k-2$ variables from $\{X_1, X_2, \ldots, X_{n-1}\}$ then H is of the form $hh^c h^c h$. Now two cases arise.
   (a) If $k-2 > 0$, then $wd(F,H) = wd(f^u g^u g^l f^l, hh^c h^c h) = wd(f, hh) + wd(g, h^c h^c) = 0$.
   (b) If $k-2 = 0$, then H is of the form $0^{n-1}1^{n-1}1^{n-1}0^{n-1}$ and hence, $wd(F,H) = wd(f^u g^u g^l f^l, 0^{n-1}1^{n-1}1^{n-1}0^{n-1}) = wd(f, 0^n) + wd(g, 1^n) = 0$, if $wt(f) = wt(g)$. Note that the weight condition is not required if $m = 1$.

Hence by Definition 4, F is mth order correlation immune. □

The case for the construction R is similar. Hence we get,

Theorem 6. *Let $f, g \in A_n(m)$, with $wt(f) = wt(g)$ and $F = \Psi(f, g)$, where $\Psi \in \{P, Q, R\}$. Then $F \in A_{n+1}(m)$.*

In [17] only a construction with two correlation immune functions f, g of same order was considered. However, if the correlation immunity of f, g are of different orders then we get the following result.

Theorem 7. *Let $f \in C_n(m_1)$ and $g \in A_n(m_2)$ with $m_1 < m_2$. Then $F \in C_{n+1}(m_1)$ if (a) $\Psi = P$ and $m_1 = 1$ or (b) $\Psi = P, Q$ or R and $wt(f) = wt(g)$.*

Proof. The proof that F belongs to $A_{n+1}(m_1)$ is similar to the above theorem. It can be checked that if $m_1 = 1$ then the weight condition $wt(f) = wt(g)$ is not required for P. To see that $F \in C_{n+1}(m_1)$, note that there exists a function $h \in L(n)$, which is non degenerate of $(m_1 + 1)$ variables such that $wd(f, h) \neq 0$ but $wd(g, h) = 0$. Depending on Ψ we can use this h to build a linear function $H \in L(n+1)$ which is non degenerate of (m_1+1) variables such that $wd(F,H) \neq 0$. Hence F is not correlation immune of order $(m_1 + 1)$. □

Next we consider construction of $(m+1)$th order correlation immune function from mth order correlation immune functions.

Proposition 7. *Let f be an n variable balanced function with mth order correlation immunity. Then $F = Q(f, f^c) = ff^c$ is an $(n+1)$ variable function with $(m+1)$th order correlation immunity.*

In a different form, this was first observed in [17] and later in [1]. This is the basic technique of construction used in [5]. We show that the same result can be achieved using R also.

Theorem 8. *Let $f \in C_n(m)$ and $F = \Psi(f, f^c)$ where $\Psi \in \{Q, R\}$. Then $F \in C_{n+1}(m+1)$ iff f is balanced. Moreover, F is balanced.*

Proof. We prove this theorem for $\Psi = R$, the other case being similar. Let us consider any linear/affine function $H \in L(n+1)$ which is a non degenerate function of k variables ($1 \leq k \leq m+1$). For ($1 \leq k \leq m$) the proof that $wd(F, H) = 0$ is similar to that of Lemma 2. If H is a non degenerate function of $(m+1)$ variables then H can be of the forms $hhhh$, hhh^ch^c, hh^chh^c and hh^ch^ch. Let H be of the form hh^chh^c, where $h \in L(n-1)$ is non degenerate of m variables. So, $wd(R(f, f^c), H) = wd(R(f, f^c), hh^chh^c) = wd(f, hh) + wd(f^c, h^ch^c) = 2wd(f, hh) = 0$ as $f \in C_n(m)$ and hh is a non degenerate function of m variables. It can be checked that for the other cases also $wd(R(f, f^c), H) = 0$. This shows that $F \in A_{n+1}(m+1)$.

The resulting $R(f, f^c)$ will not be in $A_{n+1}(m+2)$. We show a function $H \in L(n+1)$ which is a non degenerate function of $(m+2)$ variables, such that $wd(R(f, f^c), H) \neq 0$. Since f is not correlation immune of order $(m+1)$, there exists a non degenerate function $h_1 \in L(n)$ of $(m+1)$ variables such that $wd(f, h_1) \neq 0$. Now two cases arise.

Case 1: h_1 is of the form hh, where $h \in L(n-1)$. Then h is nondegenerate of $(m+1)$ variables and let $H \in L(n+1)$ be of the form hh^chh^c. Then, $wd(R(f, f^c), H) = wd(f^u(f^u)^cf^l(f^l)^c, hh^chh^c) = wd(f, hh) + wd(f^c, h^ch^c) = 2wd(f, hh) \neq 0$.

Case 2: h_1 is of the form hh^c, where $h \in L(n-1)$. In this case h is non degenerate of m variables and take $H \in L(n+1)$ to be of the form hh^ch^ch. Now, $wd(R(f, f^c), H) = wd(f^u(f^u)^cf^l(f^l)^c, hh^ch^ch) = wd(f, hh^c) + wd(f^c, h^ch) = 2wd(f, hh^c) \neq 0$. □

The above result does not in general hold for the construction P. If h_1 in the above proof is of the form hh^c, and we choose H to be of the form hh^chh^c, which is non degenerate of $(m+1)$ variables, then
$wd(P(f, f^c), H) = wd(f^u(f^u)^c(f^l)^cf^l, hh^chh^c) = wd(f, hh^c) + wd(f^c, h^ch) = 2wd(f, hh^c) \neq 0$. However, the following result holds.

Lemma 3. *Let $f \in \Omega_n - A_n$ be such that $wt(f^u) = wt(f^l)$. Then $P(f, f^c) \in A_{n+1}$ and is balanced.*

If f is a correlation immune function of even order then we can use f^r instead of f^c in Theorem 8.

Theorem 9. *Let $f \in C_n(m)$ and $\Psi \in \{Q, R\}$.*
1. Let $F = \Psi(f, f^r)$. Then, $F \in C_{n+1}(m+1)$ iff m is even. Moreover, F is balanced iff f is balanced.
2. Let $F = \Psi(f, (f^r)^c)$. Then, $F \in C_{n+1}(m+1)$ iff m is odd. Moreover, F is balanced.

Proof. We only prove (1) for $\Psi = R$. We show that if $H \in L(n+1)$, and H is a non degenerate function of k ($1 \leq k \leq m+1$) variables, then $wd(R(f, f^r), H) = 0$. The case where $1 \leq k \leq m$ is similar to Lemma 2. Now for $k = m+1$ four cases arise.

1. H is of the form $hhhh$. Then $h \in L(n-1)$ and h is a non degenerate function of $(m+1)$ variables. Since m is even, $(m+1)$ is odd and so $h^r = h^c$. Therefore, $wd(R(f, f^r), hhhh)) = wd(f, hh) + wd(f^r, hh) = wd(f, hh) + wd(f, h^r h^r) = wd(f, hh) + wd(f, h^c h^c) = wd(f, hh) - wd(f, hh) = 0$.
2. H is of the form $hh^c hh^c$. Then hh^c is a non degenerate function of $(m+1)$ variables and hence h is a non degenerate function of m variables. Therefore, $wd(R(f, f^r), hh^c hh^c) = wd(f, hh) + wd(f^r, h^c h^c) = 0 + 0 = 0$ as $f, f^r \in C_n(m)$.
3. H is of the form $hhh^c h^c$. Then hh is a non degenerate function of m variables and hence hh^c is a non degenerate function of $(m+1)$ variables. Therefore, $wd(R(f, f^r), hhh^c h^c) = wd(f, hh^c) + wd(f^r, hh^c) = wd(f, hh^c) + wd(f, (hh^c)^r) = wd(f, hh^c) - wd(f, hh^c) = 0$.
4. H is of the $hh^c h^c h$. Then h is a non degenerate function of $(m-1)$ variables and so hh^c is a non degenerate function of m variables.
   Hence, $wd(R(f, f^r), hh^c h^c h) = wd(f, hh^c) + wd(f^r, h^c h) = 0 + 0 = 0$.

Hence $wd(R(f, f^r), H) = 0$ and so $R(f, f^r) \in A_{n+1}(m+1)$. The proof that $R(f, f^r) \notin A_{n+1}(m+2)$ is similar to Theorem 8.
If m is odd, then it can be checked that $F \notin C_{n+1}(m+1)$. □

Camion et al. [1] had earlier proved one side of both (1) and (2) of the above theorem for $\Psi = Q$ only.

Remark 1. In Theorem 8 and Theorem 9 we can obtain a weaker result by replacing $C_n(m)$ and $C_{n+1}(m+1)$ by $A_n(m)$ and $A_{n+1}(m+1)$ respectively.

We also have the following result which is similar to Lemma 3.

Lemma 4. *Let $f \in \Omega_n - A_n$ be such that $wt(f^u) = wt(f^l)$. Then $P(f, f^r) \in A_{n+1}$.*

We will be using the results of Section 2 and Section 3 to design cryptographically strong Boolean functions in the next section.

4 Generalized Construction

Here we describe a recursive procedure to design *highly nonlinear* Boolean functions which *optimizes balancedness, degree* and *order of correlation immunity.* Such functions are ideally suited for stream cipher applications since they can resist all known types of attacks. First we require the following definition. We use the convention that $f^{rc} = (f^c)^r = (f^r)^c$.

Definition 6. *Let $(S_i)_{1 \le i \le q}$ be a finite sequence, where,*
$S_i \in \{Q, R\} \times \{c, r, rc\}$. Given a function $h \in \Omega_k$ and a sequence S_i of length q we define a function $F \in \Omega_{q+k}$ as follows.
$F_0 = h$ and $F_i = \Psi_i(F_{i-1}, F_{i-1}^{\tau_i})$
where $S_i = (\Psi_i, \tau_i)$, for $i \ge 1$, and $F = F_q$. We say that F is represented by $(h, S_1, \ldots, S_q)$ and the length of the representation is q.

First we observe that given a function $h \in \Omega_k$ it is easy to design a linear time (on the number of inputs to the function) algorithm that generates a function $F \in \Omega_{q+k}$ represented by $(h, S_1, \ldots, S_q)$. Though the size of the function F may be large, we need not store the whole truth table for F. Using the representation of F, the storage space required is not much larger than h. The penalty is that we require an algorithm to calculate the output of F. This can be done in $O(q)$ time (specifically, q clocks in hardware circuit) if h is implemented as a truth table. However, very low cost pipelined circuit (using flip flops) can be developed which produces a output at each clock pulse after an initial latency period of q clocks. Both the hardware and the algorithm are interesting which we omit here due to space constraint. Now we state some important properties of functions constructed by the above procedure.

Theorem 10. *Let $h \in \Omega_k$ and $F \in \Omega_{m+k+1}$ be represented by $(h, S_1, \ldots, S_{m+1})$ where $S_i = (\Psi_i, \tau_i)$, $\tau_{2i+1} \in \{c, rc\}$ and $\tau_{2i+2} \in \{c, r\}$ for $i \geq 0$. Then F is balanced and (1) $nl(F) = 2^{m+1}nl(h)$ (2) $deg(F) = deg(h)$, (3) If $m \geq 1$, then $F \in A_{m+k+1}(m)$. Moreover, if degree of h is k, then $F \in C_{m+k+1}(m)$.*

Proof. (1) Follows from Theorem 2. (2) Follows from Theorem 3. (3) Follows from Theorem 8, Theorem 9 and Remark 1. Moreover, if degree of h is k, then F can not be correlation immune of order $m+1$ due to the Siegenthaler's inequality. □

Note that there are four possible options of S_i for $i > 0$. Moreover, the construction P can also be used in the first step S_1, since the purpose of the first step is to attain balancedness. This generalizes the construction method of [5, Section 5], which uses the sequence $S_i = (Q, c)$ for all $i \geq 1$.

Corollary 1. *Let $h \in \Omega_k$ be balanced and $F \in \Omega_{m+k}$ be represented by $(h, S_1, \ldots, S_m)$, where $S_i = (\Psi_i, \tau_i)$, $\tau_{2i+1} \in \{c, r\}$ and $\tau_{2i+2} \in \{c, rc\}$ for $i \geq 0$. Then F is in $A_{m+k}(m)$. Moreover, if degree of h is $(k-1)$, the maximum degree attained for a balanced function, then $F \in C_{m+k}(m)$.*

It is important to realize that there are different trade-offs involved among the parameters, algebraic degree $deg(.)$, order of correlation immunity m, nonlinearity $nl(.)$, balancedness and the number of input variables n. The first result from [17], is that for any Boolean function f, $deg(f) + m \leq n$ and for balanced Boolean functions, $deg(f)+m \leq n-1$. The next result is that the maximum value of nonlinearity for even n is achieved for bent functions and it is known [13] that for $n \geq 4$, the degree of such functions cannot exceed $\frac{n}{2}$. Let us now consider the following construction which provides a good trade-off among the parameters.

Construction 1. *On input n, m we provide a method to construct a balanced n variable mth order correlation immune function with algebraic degree $k = n - m - 1$. Let $h \in \Omega_k$ of degree k be as follows.*
If k is even, then h is formed by adding the term $X_1 \ldots X_k$ (logical AND of X_1 to X_k) to a bent function g of k variables. If k is odd then h is formed by adding the term $X_1 \ldots X_k$ to a function g of k variables, where g is formed by

concatenating two bent functions of $(k-1)$ *variables.*
Let $F \in \Omega_n$ *where* $n = m+k+1$ *and* $m \geq 1$. F *is represented by* $(h, S_1, \ldots, S_{m+1})$ *where* $S_i = (\Psi_i, \tau_i)$, $\Psi_i \in \{Q, R\}$, $\tau_{2i+1} \in \{c, rc\}$ *and* $\tau_{2i+2} \in \{c, r\}$ *for* $i \geq 0$.

It is clear from the above discussion that Construction 1 provides functions which optimize Siegenthaler's inequality. We now find out the exact expression of nonlinearity obtained by the above construction. The result follows from Theorem 10, Corollary 1 and the nonlinearity of bent functions.

Theorem 11. *Consider* $F \in C_n(m)$ *as in Construction 1.*
(1) If $n \not\equiv m \bmod 2$, *then* $nl(F) = 2^{n-1} - 2^{\frac{n+m-1}{2}} - 2^{m+1}$.
(2) If $n \equiv m \bmod 2$ *then* $nl(F) \geq 2^{n-1} - 2^{\frac{n+m}{2}} - 2^{m+1}$.

Proof. (1) We take a bent function g of $k = n - m - 1$ variable and $h = g \oplus X_1X_2 \ldots X_{n-m-1}$. Thus by Theorem 4, $nl(g) = 2^{n-m-2} - 2^{\frac{n-m-1}{2}-1} - 1$. Then we apply our method of Definition 6 to get $nl(f) = 2^{m+1}\ (2^{n-m-2} - 2^{\frac{n-m-1}{2}-1} - 1)$.
(2) We take a bent function g_1 of $n-m-2$ variables. Then we use bent concatenation to get g of $n-m-1$ variables with nonlinearity $nl(g) = 2^{n-m-2} - 2^{\frac{n-m-2}{2}}$. Now, $h = g \oplus X_1X_2 \ldots X_{n-m-1}$. Thus, $nl(h) \geq 2^{n-m-2} - 2^{\frac{n-m-2}{2}} - 1$. Hence, $nl(f) \geq 2^{m+1}\ (2^{n-m-2} - 2^{\frac{n-m-2}{2}} - 1)$. □

This also shows that by varying the order of correlation immunity, one can adjust the nonlinearity of the optimized functions.

5 Direct Construction

Here we provide a simpler interpretation of the construction method provided in [16] and show that this also gives simpler proofs for the order of correlation immunity and nonlinearity of the constructed functions. Let $L(n, k)$ be the set of all $f \in L(n)$, which are the sum modulo 2 (XOR) of exactly k variables and $MU(n, k) = L(n, k) \cup L(n, k+1) \cup \ldots \cup L(n, n)$. Also let $ML(n, k) = L(n, 1) \cup L(n, 2) \cup \ldots \cup L(n, k)$.

Definition 7. *Let* $n = n_1 + n_2$ *and choose* 2^{n_1} *functions* $f_0, \ldots, f_{2^{n_1}-1}$ *from the set* $MU(n_2, m+1)$. *Let* $f = f_0 \ldots f_{2^{n_1}-1}$, *and denote by* $\Gamma(n, n_2, m)$ *the set of all such functions. Clearly* $\Gamma(n, n_2, m) \subseteq \Omega_n$.

We first state a simple result which is crucial to understand the cryptographic properties of the construction provided by Definition 7. The proof is a simple consequence of the fact that the XOR of two linear functions is also a linear function.

Proposition 8. *Let* $l_1, l_2 \in L(n)$. *(a) If* $l_1 = l_2$ *then* $D(l_1, l_2) = 0$. *(b) If* $l_1 = l_2^c$ *then* $D(l_1, l_2) = 2^n$ *and (c) If* $l_1 \neq l_2$ *or* l_2^c *then* $D(l_1, l_2) = 2^{n-1}$. *Consequently, the Walsh distances are respectively,* 2^n, -2^n *and* 0.

The following result is easy to see.

Proposition 9. *The construction provided in Definition 7 is same as that given by Equation 5 of [16].*

This proves (using [16, Corollary 8]) that any function in $\Gamma(n, n_2, m)$ is an mth order CI function. Here we provide a much simpler direct proof as follows.

Theorem 12. $\Gamma(n, n_2, m) \subseteq A_n(m)$.

Proof. Let $f \in \Gamma(n, n_2, m)$. We show that for any $l \in L(n, k)$, $wd(f, l) = 0$ for all $1 \leq k \leq m$. We write $f = f_0 \dots f_{2^{n_1}-1}$, where each $f_i \in MU(n_2, m+1)$. It is not difficult to see that l can be written as $l_0 \dots l_{2^{n_1}-1}$, where each $l_i \in ML(n_2, m)$. Now $wd(f, l) = wd(f_0 \dots f_{2^{n_1}-1}, l_0 \dots l_{2^{n_1}-1}) = \sum_{i=0}^{2^{n_1}-1} wd(f_i, l_i) = 0$, using Proposition 8, since $f_i, l_i \in L(n_2)$ and $f_i \neq l_i$ or l_i^c. □

Visualizing the construction as above, it is easy to obtain the nonlinearity as follows.

Theorem 13. *Let $f \in \Gamma(n, n_2, m)$ be of the form $f_0 \dots f_{2^{n_1}-1}$, where each $f_i \in MU(n_2, m+1)$. Then $nl(f) \geq 2^{n-1} - t2^{n_2-1}$, where t is the maximum number of times a function h or its complement h^c are together repeated in the construction $f_0 \dots f_{2^{n_1}-1}$ for some $h \in MU(n_2, m+1)$.*

Proof. Let $l \in L(n)$. We have to show that $D(f, l)$ is at least as large as the given bound. Note that l can be written as $l_0 \dots l_{2^{n_1}-1}$, where each l_i is either g or g^c for some $g \in L(n_2)$. Then at most t of the l_i's and f_i's can be equal. Using Proposition 8, it follows $D(l, f) \geq (2^{n_1} - t)2^{n_2-1} = 2^{n-1} - t2^{n_2-1}$. □

Theorem 13 is first proved in [16, Theorem 14]. However, our proof is much shorter and clearer. One can show as in [16, Theorem 12], that the degree of such functions is $n - n_2 + 1$, provided there are at least two functions g_1, g_2 among the f_i's of Theorem 13, such that $g_1 \neq g_2$ or g_2^c and there is a variable which occurs in an odd number of these f_i's. Thus maximum degree is attained if $n_2 = m + 2$, in which case the constructed function optimizes Siegenthaler's inequality.

Let us now estimate the nonlinearity of functions constructed using the method of [16], for functions which optimize Siegenthaler's inequality.
Let $\Omega_{k,n} = MU(n, k+1)$. *By $nld(n)$, we denote the lower bound on nonlinearity of n-variable optimized functions achieved by the direct construction of Definition 7 (see also [16]).* Note that Siegenthaler's inequality is optimized if $n_2 = m + 2$ and in this case,
$| \Omega_{m,m+2} | = \binom{m+2}{m+1} + \binom{m+2}{m+2} = m + 3$. Since one has to choose 2^{n_1} functions from $\Omega_{m,m+2}$, the repetition factor t is at least $\lceil \frac{2^{n-m-2}}{m+3} \rceil$ and hence the nonlinearity obtained is
$nld(n) = 2^{n-1} - \lceil \frac{2^{n-m-2}}{m+3} \rceil \ 2^{m+1}$.

Remark 2. The construction method provided in Section 4 is a recursive concatenation of highly nonlinear Boolean functions. On the other hand, the construction provided in Definition 7 is a direct concatenation of linear functions.

6 Results and Comparison to Previous Research

First we compare $nld(n)$, the nonlinearity of [16], with our method. *For n variable functions, let the lower bound of nonlinearity obtained by our recursive construction be* $nlr(n)$.

1. **When** $n \not\equiv m \bmod 2$.

$nlr(n) = 2^{n-1} - (2^{\frac{n-m-3}{2}} + 1)\ 2^{m+1}$. $nld(n) = 2^{n-1} - \lceil \frac{2^{n-m-2}}{m+3} \rceil\ 2^{m+1}$. So, our method works favorably when

$$\lceil \frac{2^{n-m-2}}{m+3} \rceil > 2^{\frac{n-m-3}{2}} + 1. \quad (I)$$

We consider it more conservatively, i.e., we replace $(2^{\frac{n-m-3}{2}} + 1)$ by $2^{\frac{n-m-2}{2}}$. Hence, our method performs better when, $\lceil \frac{2^{n-m-2}}{m+3} \rceil > 2^{\frac{n-m-2}{2}}$, i.e., $\lceil \frac{2^{n-m-2}}{2^{\log_2(m+3)}} \rceil > 2^{\frac{n-m-2}{2}}$, i.e., $n - m - 2 - \log_2(m+3) > \frac{n-m-2}{2}$, i.e., when,

$$n > m + 2\log_2(m+3) + 2. \quad (IA)$$

2. **When** $n \equiv m \bmod 2$**.**

$nlr(n) = 2^{n-1} - (2^{\frac{n-m-2}{2}} + 1)\ 2^{m+1}$. $nld(n) = 2^{n-1} - \lceil \frac{2^{n-m-2}}{m+3} \rceil\ 2^{m+1}$. So, our method works favorably when $\lceil \frac{2^{n-m-2}}{m+3} \rceil > 2^{\frac{n-m-2}{2}} + 1$. (II)

We consider it more conservatively, i.e., we replace $(2^{\frac{n-m-2}{2}} + 1)$ by $2^{\frac{n-m-1}{2}}$. Hence, our method performs better when, $\lceil \frac{2^{n-m-2}}{2^{\log_2(m+3)}} \rceil \geq 2^{\frac{n-m-1}{2}}$, i.e., $\lceil \frac{2^{n-m-2}}{2^{\log_2(m+3)}} \rceil > 2^{\frac{n-m-1}{2}}$, i.e., $n - m - 2 - \log_2(m+3) > \frac{n-m-1}{2}$, i.e., when

$$n > m + 2\log_2(m+3) + 3. \quad (IIA)$$

One can look at (I) and (II) in two ways.

1. If we fix a particular value of m, then there is a certain N, such that $nlr(n) > nld(n)$ for all $n \geq N$. For example for $m = 1$, our method performs better for all $n \geq 8$.

2. Similarly, if we fix a value of n, we get an upper bound $M(n)$ on m, such that for all $m \leq M(n)$, we have $nlr(n) > nld(n)$. Moreover, from (IA) and (IIA), it is clear that this upper bound $M(n)$ becomes close to n, as n increases.

This clearly shows that in a majority of cases the functions obtained by our method are better than those obtained using [16]. It should also be noted that if we take $m = 1$, then $nld(n) = 2^{n-1} - 2^{n-3}$. Whereas,

(1) If n even, then $nlr(n) = 2^{n-1} - 2^{\frac{n}{2}} - 4$. (2) If n odd, then $nlr(n) \geq 2^{n-1} - 2^{\frac{n+1}{2}} - 4$.

It should be noted that high nonlinearity can be obtained by the direct construction provided in [16] without optimizing the Siegenthaler's inequality. Currently no known general method can provide balanced CI functions with such nonlinearity. However, the nonlinearity of this method decreases when the optimization criterion is considered. From [16, Theorem 12, 14], if one does not want to optimize the Siegenthaler's inequality, then the nonlinearity for first order correlation immune functions is (we denote it as $nlx(n)$ for n variable function) $nlx(n) = 2^{n-1} - min_{3 \leq r < n}(\lceil \frac{2^{n-r}}{2^r - r - 1} \rceil 2^{r-1})$ with algebraic degree $n - r + 1$. In the following table we compare nonlinearities of first order CI functions. In second, third and fourth columns we respectively provide nonlinearities of nonoptimized

$(m + d < n - 1)$ functions constructed using the method of [16], optimized functions constructed using the method of [16] and optimized functions constructed using our recursive method. Each of the entries are <**nonlinearity, algebraic degree**>. Note that the Equations $(IA), (IIA)$ provides a clear analysis of when our nonlinearity is better than that of [16]. The table only illustrates this point for small values of n.

n	$nlx(n)$	$nld(n)$	$nlr(n)$	n	$nlx(n)$	$nld(n)$	$nlr(n)$
8	112, 5	96, 6	108, 6	11	992, 7	768, 9	956, 9
9	240, 5	192, 7	220, 7	12	1984, 7	1536, 10	1980, 10
10	480, 6	384, 8	476, 8	13	4032, 7	3072, 11	3964, 11

Let us consider the class of all optimized functions constructed using the method of [16] for each n. Then also the maximum (lower bound on) nonlinearity achieved is 96, 208, 448, 896, 1792, 3584 for n from 8 to 13 respectively. Column 4 of the table shows that the nonlinearities obtained by our method are better.

Example 1. The class of functions constructed by our recursive method contains significantly different functions. As a simple example for $n = 10$, and $m = 1$, let $f_1 = (h, (Q, c), (Q, r))$ and $f_2 = (h, (R, c), (R, r))$, where $h \in \Omega_8$ and is modified from bent functions as in Construction 1. As a concrete example, $h = \bigoplus_{i=1}^{4} X_i Y_i \oplus X_1 X_2 X_3 X_4 \oplus X_1 \dots X_4 Y_1 \dots Y_4$. Then both f_1, f_2 contains 8 terms of degree 8. Moreover, the function $f_1 \oplus f_2$ is nondegenerate and contains 14 terms of degree 8. The algebraic normal forms of f_1, f_2 and $f_1 \oplus f_2$ are complicated and too long to be written down here.

Next we compare the performance of our construction with [5]. In [5, Section 5], balanced $g \in \Omega_9$ with $nl(g) = 224$, correlation immunity of order 2 and degree 6 has been reported. The function is optimized with respect to Siegenthaler's inequality. The function g has the representation $(f, (Q, c), (Q, c))$, where $f \in \Omega_7$ and has degree 6 (only one term) and nonlinearity 56. It was remarked in [5, Example 5] that g is the representative of all such functions obtained which are well-suited for stream cipher application. Using this function f as our initial function, one can construct *more functions* of the form (f, S_1, S_2) as in Corollary 1, with the same parameters as g above. As an example one can construct a function of the form $(f, (Q, r), (Q, c))$, which contains 6 terms of degree 6. However, it seems difficult to get such good functions f for a larger number of input variables. The particular function $f \in \Omega_7$ reported in [5, Section 5] was obtained by exhaustive search over a particular subset (the idempotents) of Boolean functions. It seems infeasible to carry out such an exhaustive search for functions of larger number of input variables. Using our method from scratch, if one starts with a bent function $f_1 \in \Omega_6$ and apply Construction 1, we get a balanced, second order correlation immune function g_1 with degree 6 and $nl(g_1) = 216$. The direct construction method in [16] provides a nonlinearity $2^{9-1} - \lceil \frac{2^{9-2-2}}{2+3} \rceil \, 2^{2+1} = 200$.

We next compare the nonlinearities obtained in [11] with the following simple construction. Algebraic degree and optimization criteria is not considered in [11]

and hence we also do not consider it for this comparison. For n even, we start with a bent function h of $(n-2)$ variables and construct F represented by (h, S_1, S_2) as in Theorem 10. Then F is a balanced correlation immune function of order 1 and $nl(F) = 4nl(h)$. For n odd, we start with a best known example of balanced nonlinear function h of $(n-1)$ variables as in [11, Table 1] and construct F represented by any sequence of length one. Then by Corollary 1, F is a balanced first order correlation immune function with $nl(F) = 2nl(h)$. We compare the result using 3 tuples (n, Nonlinearity [11], Our Nonlinearity) : **(8, 112, 112), (9, 232, 232), (10, 476, 480), (11, 976, 984), (12, 1972, 1984).** Note that, better nonlinearity for functions of 10 variables and onwards can be found using a deterministic technique compared to an evolutionary one.

The recursive method proposed here can be used effectively to construct functions with large number of variables. As an example, if we take a bent function $h \in \Omega_{20}$ and consider a F represented by a sequence $(f, S_1, \ldots, S_{30})$ satisfying Theorem 11, then $F \in A_{50}(29)$ with nonlinearity $2^{49} - 2^{39} - 2^{30}$ and $deg(F) = 20$. *Currently, there are no known methods which can construct such an optimized function with better or even equal nonlinearity.* Direct implementation of F using truth table will take 2^{50} bits, which is not feasible to store. However, using the representation of F as $(f, S_1, \ldots, S_{30})$, it is possible to implement F using 1 Megabit, i.e., 128 Kilobytes by implementing f by truth table. Moreover, if f is of the form $f(X_1, \ldots, X_{10}, Y_1, \ldots, Y_{10}) = \bigoplus_{i=1}^{10} X_i Y_i \oplus g(Y_1, \ldots, Y_{10}) \oplus X_1 \ldots X_{10} Y_1 \ldots Y_{10}$, implementation of f requires 1 Kilobit (128 bytes) of memory as we need to represent g by truth table only.

References

1. P. Camion, C. Carlet, P. Charpin, and N. Sendrier. On correlation immune functions. In *Advances in Cryptology - CRYPTO'91*, pages 86–100. Springer-Verlag, 1992.
2. C. Carlet. More correlation immune and resilient functions over Galois fields and Galois rings. In *Advances in Cryptology - EUROCRYPT'97*, pages 422–433. Springer-Verlag, May 1997.
3. C. Carlet and P. Guillot. A characterization of bent functions. *Journal of Combinatorial Theory, Series A*, 76(2):328–335, September 1996.
4. C. Ding, G. Xiao, and W. Shan. *The Stability Theory of Stream Ciphers.* Lecture Notes in Computer Science. Springer-Verlag, 1991.
5. E. Filiol and C. Fontaine. Highly nonlinear balanced Boolean functions with a good correlation-immunity. In *Advances in Cryptology - EUROCRYPT'98.* Springer-Verlag, 1998.
6. X. G. Zhen and J. Massey. A spectral characterization of correlation immune combining functions. *IEEE Transactions on Information Theory*, 34(3):569–571, May 1988.
7. R. W. Hamming. *Coding And Information Theory.* Prentice Hall Inc., 1980.
8. S. Maitra and P. Sarkar. Enumeration of correlation immune Boolean functions. In *4th Australasian Conference on Information, Security and Privacy.* Springer Verlag, Lecture Notes in Computer Science, No 1587, 7-9 April 1999.

9. J. Massey. Shift-Register Synthesis and BCH Decoding. *IEEE Transactions on Information Theory*, IT-15:122–127, January 1969.
10. W. Meier and O. Staffelbach. Nonlinearity criteria for cryptographic functions. In *Advances in Cryptology - EUROCRYPT'89*, pages 549–562. Springer-Verlag, 1990.
11. W. Millan, A. Clark, and E. Dawson. Heuristic design of cryptographically strong balanced Boolean functions. In *Advances in Cryptology - EUROCRYPT'98*. Springer-Verlag, 1998.
12. C. J. Mitchell. Enumerating Boolean functions of cryptographic significance. *Journal of Cryptology*, 2(3):155–170, 1990.
13. O. S. Rothaus. On bent functions. *Journal of Combinatorial Theory, Series A20*, pages 300–305, 1976.
14. R. A. Rueppel and O. J. Staffelbach. Products of Linear Recurring Sequences with Maximum Complexity. *IEEE Transactions on Information Theory*, IT-33:124–131, January 1987.
15. J. Seberry, X. M. Zhang, and Y. Zheng. Nonlinearly balanced Boolean functions and their propagation characteristics. In *Advances in Cryptology - CRYPTO'93*, pages 49–60. Springer-Verlag, 1994.
16. J. Seberry, X. M. Zhang, and Y. Zheng. On constructions and nonlinearity of correlation immune Boolean functions. In *Advances in Cryptology - EUROCRYPT'93*, pages 181–199. Springer-Verlag, 1994.
17. T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information Theory*, IT-30(5):776–780, September 1984.
18. T. Siegenthaler. Decrypting a class of stream ciphers using ciphertext only. *IEEE Transactions on Computers*, C-34(1):81–85, January 1985.