On an Improved Correlation Analysis of Stream Ciphers Using Muti-Output Boolean Functions and the Related Generalized Notion of Nonlinearity [*]

Claude Carlet
Université Paris 8, Département de mathématiques
2, rue de la Liberté; 93526 - SAINT-DENIS cedex 02, France
email: claude.carlet@inria.fr

Khoongming Khoo, Chu-Wee Lim and Chuan-Wen Loe
DSO National Laboratories
20 Science Park Dr
S118230, Singapore
email: kkhoongm@dso.org.sg, lchuwee@dso.org.sg, lchuanwe@dso.org.sg

December 20, 2007

Abstract

We investigate the security of n-bit to m-bit vectorial Boolean functions in stream ciphers. Such stream ciphers have higher throughput than those using single-bit output Boolean functions. However, as shown by Zhang and Chan at Crypto 2000, linear approximations based on composing the vector output with any Boolean functions have higher bias than those based on the usual correlation attack. In this paper, we introduce a new approach for analyzing vector Boolean functions called generalized correlation analysis. It is based on approximate equations which are linear in the input x but of free degree in the output $z = F(x)$. The complexity for computing the generalized nonlinearity for this new attack is reduced from $2^{2^m \times n + n}$ to 2^{2n}. Based on experimental results, we show that the new generalized correlation attack gives linear approximation with much higher bias than the Zhang-Chan and usual correlation attack. We confirm this with a theoretical upper bound for generalized nonlinearity, which is much lower than for the unrestricted nonlinearity (for Zhang-Chan's attack) and *a fortiori* for usual nonlinearity. We also prove a lower bound for generalized nonlinearity which allows us to construct vector Boolean functions with high generalized nonlinearity from bent and almost bent functions. We derive the generalized nonlinearity of some known secondary constructions for secure vector Boolean functions. Finally, we prove that if a vector Boolean function has high nonlinearity or even a high unrestricted nonlinearity, it cannot ensure that it will have high generalized nonlinearity.

Keywords. Vectorial Boolean Functions, Unrestricted Nonlinearity, Generalized Nonlinearity.

[*] This is an extended version of a paper [3] presented at the FSE 2007 conference. New results are presented in Sections 6, 7, 8 and 9 of this paper.

1 Introduction

In this paper, we consider n-bit to m-bit vectorial Boolean functions when they are used in stream ciphers. There are two basic designs for such stream ciphers based on linear feedback shift registers (LFSR). One is the combiner generator [11] which consists of n LFSR's and a vector function $F(x)$. At each clock, one bit is tapped from the secret state of each LFSR as an input bit of $F(x)$ to produce m bits of output keystream. This keystream is then XORed with the plaintext to form the ciphertext. The other model is the filter function generator [11] where n bits are tapped from one LFSR as input to $F(x)$ to produce the keystream output. The advantage of using vector Boolean functions is that the stream ciphers have then higher throughput, since the encryption and decryption speed is m times faster than with single output Boolean functions. However, we need to study its security when compared to the single-bit output case.

A basic attack on these stream ciphers is the correlation attack of Siegenthaler [13]. In [13], a linear approximation is formed between the LFSR state bits and output keystream. If the approximation has probability $p \neq 1/2$, then we can recover the secret LFSR bits when enough keystream bits are known. Siegenthaler's attack was described for single-output Boolean functions but it can be generalized naturally to the vector output case where we take any linear combination of output bits.

This attack can be improved as shown by Zhang and Chan at Crypto 2000 [14] where they consider linear approximation of any combination (instead of just linear combination) of the output vector bits. Since there are 2^{2^m+n} linear approximations to choose from in the Zhang-Chan approach compared to just 2^{n+m} linear approximations in the usual approach, it seems easier to choose one with higher bias, i.e. where probability p is further away from $1/2$. This has been confirmed by an upper bound on the parameter quantifying the resistance of the function to the Zhang-Chan attack, called its unrestricted nonlinearity [4].

In Section 2, we introduce the generalized correlation attack by considering linear approximations which are linear in the input x as for the Zhang-Chan attack, but of free degree in the output $z = F(x)$. Now there are $2^{2^m \times (n+1)}$ linear approximations from which we can choose one with even higher bias than the Zhang-Chan and usual correlation attack. However, choosing the best linear approximation out of that many choices is infeasible. Therefore in Section 3, we reduce the complexity of choosing the best linear approximation for generalized correlation attack from $2^{2^m \times (n+1)+n}$ to 2^{2n}, which is much more manageable.

The generalized nonlinearity is an analogue of the usual nonlinearity, which measures the effectiveness of a function against generalized correlation attack. Based on efficient computation for finding the best generalized linear approximation, we computed the generalized nonlinearity of highly nonlinear vector functions and randomly generated vector functions in Section 3.2. We observe that the generalized nonlinearity is much lower than the usual nonlinearity and unrestricted nonlinearity (corresponding to Zhang-Chan's attack) for these functions. For example, when the inverse function on $GF(2^8)$ is restricted to $5, 6, 7$ output bits, the usual and unrestricted nonlinearities are non-zero while the generalized nonlinearity is already zero. That means the stream cipher can be attacked as a deterministic linear system while the Zhang-Chan and usual correlation attack are still probabilistic.

Theoretical bounds on the generalized nonlinearity are also derived. In Section 4, we derive an upper bound for generalized nonlinearity which is much lower than the upper bound for usual correlation attack (covering radius bound [2]) and that for Zhang-Chan's attack (unrestricted nonlinearity bound [4]). Thus it gives further evidence that generalized correlation attack is more effective than the other correlation attacks on vector Boolean functions. In Section 6, we prove a lower bound for generalized nonlinearity in terms of the nonlinearity of a function. Based on this lower bound, we can derive vector Boolean functions with high generalized nonlinearity from bent and almost bent functions. Furthermore, we prove in Section 7 that when the chosen almost bent functions are the Gold or the Kasami power functions, the resulting generalized nonlinearity can be improved.

The generalized correlation attack on single-bit output Boolean functions corresponds to bilinear cryptanalysis. However, we can deduce from the lower bound on generalized nonlinearity that biliniear cryptanalysis does not improve on the usual correlation attack.

In Section 8, we investigate the generalized nonlinearity of some secondary constructions for vector Boolean functions that are resilient and/or possess high nonlinearity. Some consequences of our study include the following. Input composition of a vector Boolean function with an invertible linear function preserves generalized nonlinearity. Output composition (e.g. dropping output bits) of balanced vector functions may increase generalized nonlinearity (the generalized nonlinearity is preserved when the output is composed with a bijection). The construction of Zhang-Zheng [15] for obtaining nonlinear resilient functions from linear resilient functions is insecure (this fact has also been noted in [2]). For a concatenated function to possess high generalized nonlinearity, we require all component functions to possess high generalized nonlinearity.

In [4], a function with high unrestricted nonlinearity $2^{n-1} - 2^{n/2}$ (n even) is constructed. This unrestricted nonlinearity is the best known as it is the same as the highest nonlinearity of balanced vectorial functions known in the literature [2]. However, we shall show in Section 9 that its generalized nonlinearity is zero. Thus, a high unrestricted nonlinearity is not sufficient to ensure high generalized nonlinearity.

In Section 10, we summarize our findings and pose some open problems for further research.

2 Generalized Correlation Analysis of Vector Output Stream Ciphers

In this section, we consider a stream cipher where the state bits of one or more linear feedback shift registers are filtered by a vector Boolean function $F : GF(2)^n \to GF(2)^m$ to form keystream bits. The keystream bits will be XORed with the plaintext to form the ciphertext.

Traditionally, an adversary who wants to perform correlation attack on this stream cipher tries to find an approximation of a linear combination of output bits by a linear combination of input bits $u \cdot F(x) \approx w \cdot x$. For correlation attack to be successful, we require that the bias defined by:

$$Bias = |Pr(u \cdot F(x) = w \cdot x) - 1/2|,\ u \in GF(2)^m,\ w \in GF(2)^n,$$

is large. Conversely, if all linear approximations of $u \cdot F(x)$ have small bias, then F is secure against correlation attack.

A concept related to the correlation attack is the Hadamard (or Walsh) transform $\hat{f} : GF(2)^n \to \mathbb{R}$ of a Boolean function $f : GF(2)^n \to GF(2)$ which is defined as:

$$\hat{f}(w) = \sum_{x \in GF(2)^n} (-1)^{f(x)+w\cdot x}.$$

Based on the Hadamard transform, we can define the nonlinearity [2] of $F(x)$ as:

$$N_F = 2^{n-1} - 1/2 \max_{0 \neq u \in GF(2)^m, w \in GF(2)^n} |\widehat{u \cdot F}(w)|. \tag{1}$$

From the above equation, we deduce that a high nonlinearity ensures protection against correlation attack. It is well known that $0 \leq N_F \leq 2^{n-1} - 2^{n/2-1}$ (see e.g. [2]).

At Crypto 2000, Zhang and Chan [14] observed that instead of taking linear combination of the output bit functions $u \cdot F(x)$, we can compose $F(x)$ with any Boolean function $g : GF(2)^m \to GF(2)$ and consider the probability:

$$Pr(g(z) = w \cdot x) \text{ where } z = F(x). \tag{2}$$

Because $z = F(x)$ corresponds to the output keystream which is known, then $g(z)$ is also known. Therefore $g(z) \approx w \cdot x$ is a linear approximation which can be used in correlation attacks. Since we are choosing from a larger set of equations now, we can find linear approximations with larger bias $|Pr(g(z) = w\cdot x) - 1/2|$. Let us define the unrestricted nonlinearity [4] which measures the effectiveness of the Zhang-Chan attack. Denote by $wt(f)$ the number of ones among the output of $f : GF(2)^n \to GF(2)$.

Definition 1. *Let $F : GF(2)^n \to GF(2)^m$ and let $\mathcal{G}$ be the set of m-bit Boolean functions $g : GF(2)^m \to GF(2)$.*

We define the unrestricted nonlinearity *as:*

$$UN_F = \min\{\min_{0 \neq u \in GF(2)^m}(wt(u \cdot F), 2^n - wt(u \cdot F)), nonlin_{UN}F\}$$

where

$$nonlin_{UN}F = 2^{n-1} - \frac{1}{2}\max_{w \neq 0, g \in \mathcal{G}} \widehat{g \circ F}(w). \tag{3}$$

Remark 1. *If $w = 0$ in equation (2), then it does not involve the input x and it is not useful for correlation attack. Thus we let $w \neq 0$ when computing $nonlin_{UN}F$ which gauges the effectiveness of equation (2) for correlation attack. The other part $\min_{u \neq 0}(wt(u \cdot F), 2^n - wt(u \cdot F))$ ensures that $F(x)$ is close to balanced when UN_F is high. This is essential because an unbalanced keystream will reveal statistical information on the plaintext.*

From equation (3), we deduce that a high unrestricted nonlinearity is required for protection against correlation attack on $g \circ F(x)$.

In this paper, we introduce a linear approximation for performing correlation attack, which is more effective than the Zhang-Chan attack [14]. The idea is to consider implicit equations which are linear in the input variable x and of any degree in the output variable $z = F(x)$, i.e. we consider the probability of the expression:

$$Pr(g(z) + w_1(z)x_1 + w_2(z)x_2 + \cdots + w_n(z)x_n = 0), \tag{4}$$

where $z = F(x)$ and $w_i : GF(2)^m \to GF(2)$. Because $z = F(x)$ corresponds to the output keystream which is known, $g(z)$ and $w_i(z)$ are known for all $i = 1, \ldots, n$. Thus equation (4) is a useful linear approximation.

We call the attack based on this linear approximation the *generalized correlation attack*. This attack can be considered as a generalization of Zhang-Chan's correlation attack because if we let $w_i(z) = 0$ or 1 for $i = 1 \ldots n$, equation (4) becomes equation (2). Since we are choosing from a larger set than that of Zhang and Chan, it is easier to find a linear approximation with larger bias $|Pr(g(z) + w_1(z)x_1 + w_2(z)x_2 + \cdots w_n(z)x_n = 0) - 1/2|$.

In relation to the approximation of equation (4), we make the following definition:

Definition 2. *Let $F : GF(2)^n \to GF(2)^m$. The* generalized Hadamard transform $\hat{F} : (GF(2)^{2^m})^{n+1} \to \mathbb{R}$ *is defined as:*

$$\hat{F}(g(\cdot), w_1(\cdot), \ldots, w_n(\cdot)) = \sum_{x \in GF(2)^n} (-1)^{g(F(x)) + w_1(F(x))x_1 + \cdots w_n(F(x))x_n}.$$

where the input is an $(n+1)$-tuple of Boolean functions $g, w_i : GF(2)^m \to GF(2)$, $i = 1, \ldots, n$. Let $\mathcal{G}$ be defined as in Definition 1 and let $\mathcal{W}$ be the set of all n-tuple functions $w(\cdot) = (w_1(\cdot), \ldots, w_n(\cdot))$

where $w_i \in \mathcal{G}$ and such that $w(z) = (w_1(z), \ldots, w_n(z)) \neq (0, \ldots, 0)$ for all $z \in GF(2)^m$. The generalized nonlinearity *is defined as:*

$$GN_F = \min\{\min_{0 \neq u \in GF(2)^m} (wt(u \cdot F), 2^n - wt(u \cdot F)), nonlin_{gen} F\},$$

where

$$nonlin_{gen} F = 2^{n-1} - \frac{1}{2} \max_{g \in \mathcal{G}, w \in \mathcal{W}} \hat{F}(g(\cdot), w_1(\cdot), \ldots, w_n(\cdot)). \tag{5}$$

Remark 2. *We introduce the set $\mathcal{W}$ to give a meaningful definition to the generalized nonlinearity. This is because if there exists $z \in GF(2)^m$ such that $(w_1(z), \ldots, w_n(z)) = (0, \ldots, 0)$, then equation (4) does not involve the input x and it is not useful for correlation attack. Thus we let $w \in \mathcal{W}$ when computing $nonlin_{gen} F$. The other part $\min_{u \neq 0}(wt(u \cdot F), 2^n - wt(u \cdot F))$ ensures that $F(x)$ is close to balanced when GN_F is high. This is essential because an unbalanced keystream will reveal statistical information on the plaintext.*

From equation (5), we deduce that a high generalized nonlinearity is required for protection against generalized correlation attack.

In Proposition 1, we show that the generalized nonlinearity is lower than the other nonlinearity measures and thus provides linear approximations with better bias for correlation attack. The proof follows naturally from the definitions of the various nonlinearities.

Proposition 1. *Let $F : GF(2)^n \to GF(2)^m$. Then the nonlinearity, unrestricted nonlinearity and generalized nonlinearity are related by the following inequality:*

$$GN_F \leq UN_F \leq N_F. \tag{6}$$

I.e., the generalized correlation attack is more effective than the Zhang-Chan's correlation attack, which itself is more effective than the usual correlation attack.

Proof. Let us first consider the case when $w = 0$ in the definition of N_F. We see that

$$2^{n-1} - 1/2 \max_{0 \neq u \in GF(2)^m} |\widehat{u \cdot F}(0)| = \min_{0 \neq u \in GF(2)^m} (wt(u \cdot F), 2^n - wt(u \cdot F)).$$

This measure indicates how balanced the output of $F(x)$ is and has been included in the definition of UN_F and GN_F.

Let us now consider the case when $w \neq 0$, that is, when the input x is involved in the linear approximation (which is more useful for correlation attack). Because the set of functions $\{u \cdot F(x) + w \cdot x + c\,;\, u \in GF(2)^m \backslash \{0\}; w \in GF(2)^m \backslash \{0\}, c \in GF(2)\}$ is a subset of the set of functions $\{g \circ F(x) + w \cdot x \,|\, g \in \mathcal{G}, w \in$

$GF(2)^m\backslash\{0\}\}$, which itself is a subset of the set of functions $\{g(F(x))+w_1(F(x))x_1+\cdots w_n(F(x))x_n \mid g \in \mathcal{G}, w \in \mathcal{W}\}$, we have

$$\max_{0\neq u\in GF(2)^m, 0\neq w\in GF(2)^n} |\widehat{u\cdot F}(w)| \leq \max_{w\neq 0, g\in\mathcal{G}} \widehat{g\circ F}(w) \leq \max_{g\in\mathcal{G}, w\in\mathcal{W}} \hat{F}(g(\cdot), w_1(\cdot), \ldots, w_n(\cdot)).$$

By substituting this inequality in the definition of $nonlin_{UN}F$ and $nonlin_{gen}F$, we have:

$$nonlin_{gen}F \leq nonlin_{UN}F \leq 2^{n-1} - 1/2 \max_{0\neq u\in GF(2)^m, 0\neq w\in GF(2)^n} |\widehat{u\cdot F}(w)|$$

By combining the two cases $w = 0$ and $w \neq 0$, we conclude that $GN_F \leq UN_F \leq N_F$. □

Remark 3. *A vector function $F : GF(2)^n \to GF(2)^m$ is said to be balanced if $|F^{-1}(z)| = 2^{n-m}$ for all $z \in GF(2)^m$. It is well-known that $wt(u\cdot F) = 2^{n-1}$ for all $u \in GF(2)^m - \{0\}$ if and only if F is balanced (see e.g. [2]). Thus*

$$GN_F = \min\{\min_{0\neq u\in GF(2)^m}(wt(u\cdot F), 2^n - wt(u\cdot F)), nonlin_{gen}F\} = \min(2^{n-1}, nonlin_{gen}F) = nonlin_{gen}F,$$

because $GN_F \leq N_F \leq 2^{n-1} - 2^{(n-1)/2}$ (see e.g. [2]). Therefore $GN_F = nonlin_{gen}F$ if F is balanced. In a similar way, $UN_F = nonlin_{UN}F$ if F is balanced.

2.1 The Single-Bit Output Case and Bilinear Cryptanalysis

It is easy to see that in the single output case ($m = 1$), the Zhang-Chan correlation attack is equivalent to the usual correlation attack, i.e. $UN_F = N_F$. However, it is not so obvious whether the generalized correlation attack is better than the usual correlation attack. The expression used for the generalized correlation attack is a bilinear approximation:

$$Pr(a_0z + b_0 + (a_1z + b_1)x_1 + (a_2z + b_2)x_2 + \cdots + (a_nz + b_n)x_n = 0),\ a_i, b_i \in GF(2),$$

where for any $z \in GF(2)$, some $a_1z + b_1, \ldots, a_nz + b_n$ is a non-zero function. We state the following fact without proof since it can be seen as a corollary of Theorem 4 in Section 6.

Fact 1. *Let $f : GF(2)^n \to GF(2)$. Then $GN_f = N_f$.*

Thus we see that generalized correlation attack does not improve on the usual correlation attack when the number of output bits is $m = 1$. But in Section 3.2, we will give many examples where generalized correlation attack yields better results than the usual and Zhang-Chan correlation attack when the number of output bits is $m \geq 2$.

3 Efficient Computation of the Generalized Nonlinearity

To compute the generalized nonlinearity GN_F, we first compute $\min_{0\neq u\in GF(2)^m}(wt(u\cdot F), 2^n - wt(u\cdot F))$ with complexity approximatively 2^{m+n}. Then we need to compute $nonlin_{gen}F$ which requires computation of the generalized Hadamard transform over all input. But the complexity of computing $\hat{F}(g(\cdot), w_1(\cdot), \ldots, w_n(\cdot))$ in a naive way, for all possible $(n+1)$-tuples of m-bit functions is $\approx 2^n \times 2^{2^m\times(n+1)}$: for each fixed $(g(\cdot), w_1(\cdot), \ldots, w_n(\cdot))$, we sum over 2^n elements x to compute $\hat{F}$ and there are approximately[1] $2^{2^m\times(n+1)}$ tuple of functions $g, w_i : GF(2)^m \to GF(2)$, $i = 1, \ldots, n$. This computation quickly becomes unmanageable even for small values of n, m. Since the bulk of the computational time comes from $nonlin_{gen}F$, we need to make it more efficient to compute.

First, we state Lemma 1 which rewrites the generalized Hadamard transform as a double sum.

Lemma 1. *Let $F : GF(2)^n \to GF(2)^m$ and $w_i : GF(2)^m \to GF(2)$. Let $w(\cdot)$ denote the n-tuple of m-bit Boolean functions $(w_1(\cdot), \ldots, w_n(\cdot))$. Then the generalized Hadamard transform can be expressed as:*

$$\begin{aligned}\hat{F}(g(\cdot), w_1(\cdot), \ldots, w_n(\cdot)) &= \sum_{x\in GF(2)^n} (-1)^{g(F(x))+w(F(x))\cdot x} \\ &= \sum_{z\in GF(2)^m} (-1)^{g(z)} \sum_{x\in F^{-1}(z)} (-1)^{w(z)\cdot x}.\end{aligned}$$

Based on Lemma 1, we get the following theorem, which is an analogue of Theorem 1 of [14].

Theorem 1. *Let $F : GF(2)^n \to GF(2)^m$ and $w(\cdot)$ denote the n-tuple of m-bit Boolean functions $(w_1(\cdot), \ldots, w_n(\cdot))$. The generalized nonlinearity measure $nonlin_{gen}F$ can be computed as:*

$$nonlin_{gen}F = 2^{n-1} - 1/2 \sum_{z\in GF(2)^m} \max_{w(z)\in GF(2)^n - \{0\}} \left| \sum_{x\in F^{-1}(z)} (-1)^{w(z)\cdot x} \right|$$

Proof. Based on Lemma 1, we have:

$$\begin{aligned}& \max_{g\in\mathcal{G}, w\in\mathcal{W}} \hat{F}(g(\cdot), w_1(\cdot), \ldots, w_n(\cdot)) \\ = & \max_{g\in\mathcal{G}, w\in\mathcal{W}} \sum_{z\in GF(2)^m} (-1)^{g(z)} \sum_{x\in F^{-1}(z)} (-1)^{w(z)\cdot x} \\ = & \sum_{z\in GF(2)^m} \max_{g(z)\in GF(2), w(z)\in GF(2)^n - \{0\}} (-1)^{g(z)} \sum_{x\in F^{-1}(z)} (-1)^{w(z)\cdot x}.\end{aligned}$$

[1] We say approximately $2^{2^m\times(n+1)}$ functions because we do not range over all tuples of functions $(w_1(\cdot), \ldots, w_n(\cdot))$ but only over those in the set $\mathcal{W}$ of Defintion 2

To maximize this expression, we choose $g(z) = 0$ if $\sum_{x\in F^{-1}(z)}(-1)^{w(z)\cdot x} > 0$, else we choose $g(z) = 1$. Thus we can equivalently write the expression as:

$$\max_{g\in\mathcal{G},w\in\mathcal{W}} \hat{F}(g(\cdot), w_1(\cdot), \ldots, w_n(\cdot)) = \sum_{z\in GF(2)^m} \max_{w(z)\in GF(2)^n-\{0\}} \left| \sum_{x\in F^{-1}(z)} (-1)^{w(z)\cdot x} \right|.$$

By substituting this expression in equation (5), we get $nonlin_{gen}F$. □

Remark 4. *The proof of Theorem 1 also provides the functions $g(\cdot), w_i(\cdot)$, $i = 1, \ldots, n$, for the best generalized linear approximation. At each z, the optimal $g(z)$ is the one that makes the inner sum positive while and the optimal tuple $(w_1(z), \ldots, w_n(z))$ is the n-bit vector that maximizes the inner sum.*

3.1 Reduction in Complexity

To compute $nonlin_{gen}F$ based on Theorem 1, we first perform a pre-computation to identify the sets $\{x : x \in F^{-1}(z)\}$ with complexity 2^n and store them with memory of size $n \times 2^n$. This is needed in computing the sum $\sum_{x\in F^{-1}(z)}(-1)^{w(z)\cdot x}$. We consider then the 2^m elements $z \in GF(2^m)$. For each z, we find $w(z) \in GF(2)^n$ which maximizes the sum $\left|\sum_{x\in F^{-1}(z)}(-1)^{w(z)\cdot x}\right|$. The additional computational complexity is:

$$\begin{aligned} \text{Complexity} &= \sum_{z\in GF(2)^m} 2^n \times |\{x : x \in F^{-1}(z)\}| = 2^n \sum_{z\in GF(2)^m} |\{x : x \in F^{-1}(z)\}| \\ &= 2^n \times |Domain(F)| = 2^n \times 2^n = 2^{2n}. \end{aligned}$$

Together with a complexity of 2^{m+n} to compute $\min_{0\neq u\in GF(2)^m}(wt(u \cdot F), 2^n - wt(u \cdot F))$, the total complexity for computing GN_F is:

$$\text{Precomputation } = 2^n, \text{ Memory } = n \times 2^n, \text{ Time Complexity } = 2^{m+n} + 2^{2n}.$$

This is much less than a time complexity of $2^{m+n} + 2^{n+2^m\times(n+1)}$ by the direct approach.

3.2 Experimental Results

Based on Theorem 1, we can compute the generalized nonlinearity of some highly nonlinear functions. We also computed the unrestricted nonlinearity of these functions for comparison. The bulk of the complexity for computing UN_F comes from $nonlin_{UN}F$. To compute $nonlin_{UN}F$ efficiently, we recall Theorem 1 of [14].

We give now some examples of computations of N_F, UN_F and GN_F and of the corresponding optimum bias. First, let us look at bent functions, which have the highest nonlinearity $N_F = 2^{n-1} - 2^{n/2-1}$.

Table 1: Truth Table of $F(x)$ from Example 1.

x	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
$F(x)$	00	00	00	00	00	01	10	11	11	00	10	01	11	01	00	10

Example 1. *Consider the bent function $F : GF(2)^4 \to GF(2)^2$ defined by $F(x_1, x_2, x_3, x_4) = (z_1, z_2) = (x_1 + x_1x_4 + x_2x_3, x_1 + x_1x_3 + x_1x_4 + x_2x_4)$. The truth table of F is listed in Table 1. The various nonlinearity and bias take the following values:*

$$\begin{aligned} \textit{Usual nonlinearity } N_F &= 6 \Rightarrow Bias = 0.125 \\ \textit{Unrestricted nonlinearity } UN_F &= 5 \Rightarrow Bias = 0.1875 \\ \textit{Generalized nonlinearity } GN_F &= 2 \Rightarrow Bias = 0.375. \end{aligned}$$

From Remark 4, we deduce that the following approximation holds with bias 0.375*.*

$$Pr(z_1 + z_2 = (z_1 + 1)(z_2 + 1)x_2 + z_1x_3 + z_2x_4) = \frac{14}{16},$$

where $x = 0100, 1110$ are the only two points not satisfying the relation.

We experimented with other bent functions, and observed that the generalized nonlinearity is strictly lower than the other nonlinearities for these functions. In the next example, we look at functions formed from dropping $(n - m)$ output bits of the inverse function x^{-1} over the finite field $GF(2^n)$. They are balanced functions with high nonlinearity.

Example 2. *Let $GF(2^8)$ be the finite field defined by the relation $\alpha^8 + \alpha^4 + \alpha^3 + \alpha^2 + 1 = 0$. Consider the S-box $Inv : GF(2)^8 \to GF(2)^8$ of the x^{-1} function with the correspondence*

$$(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) \leftrightarrow x_1\alpha^7 + x_2\alpha^6 + \cdots + x_7\alpha + x_8$$

Consider $Inv(x)$ restricted to the least significant m bits. Then the nonlinearity, unrestricted nonlinearity and generalized nonlinearity are given by Table 2. We see that the generalized nonlinearity for the inverse function restricted to m output bits is lower than the usual and unrestricted nonlinearities. Therefore generalized correlation attack works better in this case.

Moreover, for $m \geq 5$ output bits, the generalized nonlinearity is already 0 *which means the system can be broken by linear algebra with very few keystream bits. In comparison, the linear relations in the usual and Zhang-Chan correlation attack are still probabilistic and require more keystream bits to determine the right key.*

Table 2: Nonlinearities for x^{-1} on $GF(2^8)$ restricted to m least significant output bits.

n	8	8	8	8	8	8	8
m	1	2	3	4	5	6	7
N_F	112	112	112	112	112	112	112
UN_F	112	108	100	94	84	70	56
GN_F	112	80	66	40	0	0	0

Table 3: Average nonlinearity for randomly generated balanced functions, $n = 2m$

n	6	8	10	12	14
m	3	4	5	6	7
N_F	18	100	443	1897	7856
UN_F	16	88	407	1768	7454
GN_F	6	36	213	1101	5224

Example 3. *Lastly in Table 3, we tabulate the average nonlinearity measures for 100 randomly generated balanced functions $F : GF(2)^n \to GF(2)^m$, $n = 2m$, for various n. Again, we see that the average generalized nonlinearity is much lower than the unrestricted and usual nonlinearities, and that generalized correlation attack is more effective.*

4 Upper Bound on Generalized Nonlinearity

In this Section, we prove an upper bound for the generalized nonlinearity. This allows us to gauge theoretically the effectiveness of the generalized correlation attack.

Theorem 2. *Let $F : GF(2)^n \to GF(2)^m$. Then the following inequality holds.*

$$nonlin_{gen}F \leq 2^{n-1} - \frac{1}{4} \sum_{z \in GF(2)^m} \sqrt{\frac{2^{n+2}|F^{-1}(z)| - 4|F^{-1}(z)|^2}{2^n - 1}}.$$

Furthermore if $F(x)$ is balanced, then we have:

$$GN_F \leq 2^{n-1} - 2^{n-1}\sqrt{\frac{2^m - 1}{2^n - 1}}$$

Proof. According to Theorem 1, we have:

$$nonlin_{gen}F = 2^{n-1} - 1/2 \sum_{z \in GF(2)^m} \max_{a \in GF(2)^n - \{0\}} \left| \sum_{x \in F^{-1}(z)} (-1)^{a \cdot x} \right|.$$

Let $\phi_z(x)$ be the indicator function of $F^{-1}(z)$. I.e., $\phi_z(x) = 1$ if $F(x) = z$ else $\phi_z(x) = 0$. Then:

$$\begin{aligned}\sum_{x\in F^{-1}(z)} (-1)^{a\cdot x} &= \sum_{x\in GF(2)^n} \phi_z(x)(-1)^{a\cdot x} = \sum_{x\in GF(2)^n} \frac{1-(-1)^{\phi_z(x)}}{2}(-1)^{a\cdot x} \\ &= -\frac{1}{2}\sum_{x\in GF(2)^n} (-1)^{\phi_z(x)+a\cdot x} = -\frac{1}{2}\widehat{\phi_z}(a), \text{ when } a \neq 0.\end{aligned}$$

Thus

$$nonlin_{gen}F = 2^{n-1} - 1/4 \sum_{z\in GF(2)^m} \max_{a\in GF(2)^n-\{0\}} \left|\widehat{\phi_z}(a)\right|.$$

In a similar way to the computation of $\sum_{x\in F^{-1}(z)}(-1)^{a\cdot x}$, we can prove that $|F^{-1}(z)| = \sum_{x\in F^{-1}(z)}(-1)^{0\cdot x} = 2^{n-1} - \frac{1}{2}\widehat{\phi_z}(0)$. This implies $\widehat{\phi_z}(0) = 2^n - 2|F^{-1}(z)|$.

By Parseval's relation,

$$\begin{aligned}&\sum_{a\in GF(2)^n-\{0\}} \widehat{\phi_z}(a)^2 = 2^{2n} - \widehat{\phi_z}(0)^2 \\ = \;& 2^{2n} - (2^n - 2|F^{-1}(z)|)^2 = 2^{n+2}|F^{-1}(z)| - 4|F^{-1}(z)|^2.\end{aligned}$$

By the pigeon hole principle, we deduce that

$$\max_{a\in GF(2)^n-\{0\}} \widehat{\phi_z}(a)^2 \geq \frac{2^{n+2}|F^{-1}(z)| - 4|F^{-1}(z)|^2}{2^n-1}.$$

and therefore

$$nonlin_{gen}F \leq 2^{n-1} - \frac{1}{4}\sum_{z\in GF(2)^m} \sqrt{\frac{2^{n+2}|F^{-1}(z)| - 4|F^{-1}(z)|^2}{2^n-1}}.$$

When $F(x)$ is balanced, $nonlin_{gen}F = GN_F$, $|F^{-1}(z)| = 2^{n-m}$ for all $z \in GF(2)^m$ and we deduce:

$$GN_F \leq 2^{n-1} - 2^{m-2}\sqrt{\frac{2^{2n-m+2} - 2^{2n-2m+2}}{2^n-1}} = 2^{n-1} - 2^{n-1}\sqrt{\frac{2^m-1}{2^n-1}}.$$

□

This upper bound is much lower than the covering radius bound and the upper bound for UN_F deduced in [4]:

$$UN_F \leq 2^{n-1} - \frac{1}{2}\left(\frac{2^{2m}-2^m}{2^n-1} + \sqrt{\frac{2^{2n}-2^{2n-m}}{2^n-1} + \left(\frac{2^{2m}-2^m}{2^n-1} - 1\right)^2 - 1}\right). \tag{7}$$

when $F : GF(2)^n \to GF(2)^m$ is balanced.

From [4], it is stated that the upper bound for UN_F in equation (7) is higher than the covering radius bound $2^{n-1} - 2^{n/2-1}$ when $m \leq n/2$. Therefore it is not a useful bound when $m \leq n/2$. In comparison,

Table 4: Comparison of Upper Bound for N_F, UN_F and GN_F when $m = n/2$

n	6	8	10	12	14	16
m	3	4	5	6	7	8
N_F	28	120	496	2016	8128	32640
UN_F	29	121	497	2017	8129	32641
GN_F	22	97	423	1794	7471	30724

Table 5: Comparison of Upper Bound for N_F, UN_F and GN_F when $m = \lfloor 3n/4 \rfloor$

n	6	8	10	12	14	16
m	4	6	7	9	10	12
N_F	28	120	496	2016	8128	32640
UN_F	27	110	487	1972	8090	32460
GN_F	17	65	332	1325	6145	24577

the upper bound for GN_F in Theorem 2 is clearly lower than the covering radius bound, for every m. In Table 4, we illustrate this fact for n even and $m = n/2$.

When $m > n/2$, the upper bound for UN_F is lower than the covering radius bound $2^{n-1} - 2^{(n-1)/2}$. However it is not as low as the upper bound of GN_F. We demonstrate this fact in Table 5 for n even and $m = \lfloor 3n/4 \rfloor$.

Thus Theorem 2 provides further evidence that generalized correlation attack is more effective than the usual and Zhang-Chan correlation attacks on vector Boolean functions.

5 Spectral Characterization of Generalized Correlation

In Theorem 3, we express the generalized correlation in terms of the values of the Hadamard transform of F (its Hadamard, or Walsh, spectrum). This allows us to deduce general correlation properties based on the spectral distribution.

Theorem 3. *Let $F : GF(2)^n \rightarrow GF(2)^m$ and $w_i : GF(2)^m \rightarrow GF(2)$. Let $w(\cdot)$ denote the n-tuple of m-bit Boolean functions $(w_1(\cdot), \ldots, w_n(\cdot))$. Then the generalized Hadamard transform can be expressed as:*

$$\hat{F}(g(\cdot), w_1(\cdot), \ldots, w_n(\cdot)) = \frac{1}{2^m} \sum_{z \in GF(2)^m} (-1)^{g(z)} \sum_{v \in GF(2)^m} (-1)^{v \cdot z} \widehat{v \cdot F}(w(z)).$$

Proof. Let $\phi_z(x)$ be defined as in the proof of Theorem 2. For a fixed $z \in GF(2)^m$:

$$\begin{aligned}
\sum_{x\in F^{-1}(z)} (-1)^{w(z)\cdot x} &= \frac{1}{2^m} \sum_{x\in GF(2)^n} (-1)^{w(z)\cdot x} \times 2^m \phi_z(x) \\
&= \frac{1}{2^m} \sum_{x\in GF(2)^n} (-1)^{w(z)\cdot x} \times \sum_{v\in GF(2)^m} (-1)^{v\cdot(F(x)+z)} \\
&\quad \text{(because } \sum_{v\in GF(2)^m} (-1)^{v\cdot a} = 2^m \text{ if and only if } a = 0) \\
&= \frac{1}{2^m} \sum_{v\in GF(2)^m} (-1)^{v\cdot z} \times \sum_{x\in GF(2)^n} (-1)^{w(z)\cdot x + v\cdot F(x)} \\
&= \frac{1}{2^m} \sum_{v\in GF(2)^m} (-1)^{v\cdot z} \widehat{v\cdot F}(w(z)).
\end{aligned}$$

By substituting this expression in Lemma 1, the proof is complete. □

Remark 5. *Based on Theorem 3 and equation (5), we get the following expression for $nonlin_{gen}F$.*

$$nonlin_{gen}F = 2^{n-1} - \frac{1}{2^{m+1}} \sum_{z\in GF(2)^m} \max_{w(z)\in GF(2)^n - \{0\}} \left| \sum_{v\in GF(2)^m} (-1)^{v\cdot z} \widehat{v\cdot F}(w(z)) \right|. \quad (8)$$

Assume the Hadamard transform distribution of $F(x)$ is known, then we can have a more efficient computation of GN_F. By equation (8), we compute $nonlin_{gen}F$ by an outer sum over 2^m elements z, each of which finds the maximum inner sum (over 2^m elements v) for 2^n choices of $w(z)$. Thus the complexity of computing $nonlin_{gen}F$ is 2^{n+2m}. Together with a complexity of 2^{m+n} for determining the balanceness of $F(x)$, the complexity for computing GN_F is $2^{m+n} + 2^{n+2m}$. This is more efficient than the computation of Theorem 1 because usually, m is much smaller than n in applications. Furthermore, we do not need pre-computation and memory to store the sets $\{x : x \in F^{-1}(z)\}$ as in Theorem 1. Some examples of vectorial Boolean functions on which this optimization can be applied is the Maiorana-McFarland class of functions, e.g. see [2, 5].

Besides enabling more efficient computation of the generalized nonlinearity when the spectral distribution is known, Theorem 3 also allows us to compute a lower bound for generalized nonlinearity (through equation (8)) in Section 6.

6 Lower Bound on Generalized Nonlinearity

In Section 4, we derived an upper bound for the generalized nonlinearity of a vectorial Boolean function. In this section, we shall derive a lower bound for the generalized nonlinearity.

Theorem 4. *Let $F : GF(2)^n \rightarrow GF(2)^m$, then*

$$GN_F \geq 2^{n-1} - (2^m - 1)(2^{n-1} - N_F).$$

Proof. From equation (8) in Section 5 of our paper, we see that $nonlin_{gen}F$ depends on the following sum, that we shall bound in terms of the nonlinearity N_F.

$$\begin{array}{rl} & \displaystyle\sum_{z \in GF(2)^m} \max_{a \in GF(2)^n - \{0\}} \left| \sum_{v \in GF(2)^m} (-1)^{v \cdot z} \widehat{v \cdot F}(a) \right| \\ \leq & \displaystyle\sum_{z \in GF(2)^m} \sum_{v \in GF(2)^m} \max_{a \neq 0} |\widehat{v \cdot F}(a)| \\ = & \displaystyle\sum_{z \in GF(2)^m} \sum_{v \in GF(2)^m - \{0\}} \max_{a \neq 0} |\widehat{v \cdot F}(a)| \text{ (because } \widehat{v \cdot F}(a) = 0 \text{ when } v = 0, a \neq 0) \\ \leq & 2^m(2^m - 1) \displaystyle\max_{v \neq 0, a \neq 0} |\widehat{v \cdot F}(a)| \leq 2^{m+1}(2^m - 1)(2^{n-1} - N_F). \\ & \text{(because } \displaystyle\max_{v \neq 0, a \neq 0} |\widehat{v \cdot F}(a)| \leq \max_{v \neq 0, a} |\widehat{v \cdot F}(a)| = 2^n - 2N_F) \end{array}$$

By substituting this inequality in equation (8) we have the following inequality.

$$nonlin_{gen}F \geq 2^{n-1} - (2^m - 1)(2^{n-1} - N_F).$$

Also,

$$\begin{array}{rl} & \displaystyle\min_{0 \neq u \in GF(2)^m} (wt(u \cdot F), 2^n - wt(u \cdot F)) \\ \geq & N_F \geq 2^{n-1} - (2^m - 1)(2^{n-1} - N_F). \end{array}$$

Thus

$$\begin{array}{rcl} GN_F & = & \min\{\displaystyle\min_{0 \neq u \in GF(2)^m} (wt(u \cdot F), 2^n - wt(u \cdot F)), nonlin_{gen}F\} \\ & \geq & 2^{n-1} - (2^m - 1)(2^{n-1} - N_F). \end{array}$$

□

The following corollary is an immediate consequence of Theorem 4.

Corollary 1. *Let $f : GF(2)^n \rightarrow GF(2)$. Then $GN_f = N_f$.*

Proof. From Proposition 1, we see that $GN_f \leq N_f$. From Theorem 4, we see that $GN_f \geq N_f$ when the number of output bits is $m = 1$. Therefore $GN_f = N_f$ when $m = 1$. □

Based on Theorem 4, we can construct vector Boolean functions with relatively high generalized nonlinearity from those with high nonlinearity. Two well known classes of vector Boolean functions with high nonlinearity are the bent functions and almost bent functions. The bent functions $F : GF(2)^n \to GF(2)^m$ have optimal nonlinearity $2^{n-1} - 2^{n/2-1}$ and they exist only when n is even and $m \leq n/2$. The almost bent functions $S : GF(2)^n \to GF(2)^n$ have optimal nonlinearity $2^{n-1} - 2^{(n-1)/2}$ when n is odd. By composing $S(x)$ with a surjective linear function $L : GF(2)^n \to GF(2)^m$, we can construct vector Boolean functions $F : GF(2)^n \to GF(2)^m$, $F(x) = L \circ S(x)$ with high nonlinearity $2^{n-1} - 2^{(n-1)/2}$ (obviously, this is not the only way of reaching such nonlinearity for n odd, but it is a simple one).

The following proposition on construction of vector Boolean functions with relatively high generalized nonlinearity is a direct application of Theorem 4.

Proposition 2. *1. Let n be even and $S : GF(2)^n \to GF(2)^k$, $k \leq n/2$, be a bent function. Let $L : GF(2)^k \to GF(2)^m$ be a surjective linear function and $F : GF(2)^n \to GF(2)^m$ be defined by $F(x) = L \circ S(x)$. Then*

$$GN_F \geq 2^{n-1} - 2^{n/2-1}(2^m - 1).$$

2. Let n be odd and $S : GF(2)^n \to GF(2)^n$ be an almost bent function. Let $L : GF(2)^n \to GF(2)^m$ be a surjective linear function and $F : GF(2)^n \to GF(2)^m$ be defined by $F(x) = L \circ S(x)$. Then

$$GN_F \geq 2^{n-1} - 2^{(n-1)/2}(2^m - 1).$$

7 Functions with Improved Generalized Nonlinearity

In this Section, we shall show that when $S(x)$ is some specially chosen almost bent function in Proposition 2 part (2), we can improve (increase) the lower bound on generalized nonlinearity from $2^{n-1} - 2^{(n-1)/2}(2^m - 1)$ to $2^{n-1} - 2^{(n-1)/2+m-1}$.

Theorem 5. *Let n be an odd integer and m be an integer dividing n. Let $F : GF(2^n) \to GF(2^m)$ be defined by*

1. *$F(x) = Tr_m^n(x^k)$ where $k = 2^r + 1$, $\gcd(r, n) = 1$ or*
2. *$F(x) = Tr_m^n(x^k)$ where $k = 2^{2r} - 2^r + 1$, $3r \equiv 1 \mod n$.*

Then the generalized nonlinearity satisfies $GN_F \geq 2^{n-1} - 2^{(n-1)/2+m-1}$.

Before we prove Theorem 5, we present Lemma 2 and 3 which we need for the proof.

Lemma 2. *Let n be odd and $f(x) = Tr_1^n(x^k)$ on $GF(2^n)$.*

1. *(Gold [8]) Let $k = 2^r + 1$ where $\gcd(r, n) = 1$. Then $\hat{f}(\lambda) = 0 \Longleftrightarrow Tr_1^n(\lambda) = 0$.*

2. (Dillon [7, Theorem 7]) Let $k = 2^{2r} - 2^r + 1$ *where* $3r \equiv 1 \mod n$. *Then* $\hat{f}(\lambda) = 0 \Longleftrightarrow Tr_1^n(\lambda^{2^r+1}) = 0$.

We also need the following lemma on the trace function of subfield elements.

Lemma 3. *Let* n *be odd and* $GF(2^m)$ *be a subfield of* $GF(2^n)$, *i.e.* $m|n$. *Let* u *be an integer such that* $\gcd(u, 2^m - 1) = 1$ *and* $\lambda \in GF(2^n)$. *Then*

$$|\{b \in GF(2^m)|Tr_1^n(\lambda b^u) = 0\}| \geq 2^{m-1}.$$

Proof. Note that

$$Tr_1^n(\lambda b^u) = Tr_1^m(Tr_m^n(b^u\lambda)) = Tr_1^m(b^u Tr_m^n(\lambda))$$

because $b^u \in GF(2^m)$. Since $Tr_m^n(\lambda)$ is fixed, Tr_1^m is balanced and $b \mapsto b^u$ is a permutation on $GF(2^m)$, half of $Tr_1^n(\lambda b^u)$ are 0's when we vary b if $Tr_m^n(\lambda) \neq 0$. $Tr_1^n(\lambda b^u)$ is 0 for all $b \in GF(2^m)$ if $Tr_m^n(\lambda) = 0$. □

Proof of Theorem 5: For any linear combination of output bits $v \cdot F(x)$, $v \in GF(2)^m$, there exists a unique element $b \in GF(2^m)$ such that $v \cdot F(x) = Tr_1^m(bF(x))$. And $Tr_1^m(bF(x)) = Tr_1^n(bx^k)$. We see that

$$\begin{aligned}
\widehat{v \cdot F}(\lambda) &= \sum_{x \in GF(2^n)} (-1)^{Tr_1^n(bx^k)+Tr_1^n(\lambda x)} \\
&= \sum_{y \in GF(2^n)} (-1)^{Tr_1^n(y^k)+Tr_1^n(\lambda b^{-k^{-1}} y)},\ y = b^{k^{-1}} x \\
&= \hat{f}(\lambda b^{-k^{-1}}) \text{ where } f(x) := Tr_1^n(x^k), \lambda \neq 0.
\end{aligned}$$

By Lemma 2,

$$\begin{aligned}
\hat{f}(\lambda b^{-k^{-1}}) = 0 &\Longleftrightarrow Tr_1^n(\lambda b^{-k^{-1}}) = 0 \text{ for } k = 2^r + 1 \\
\hat{f}(\lambda b^{-k^{-1}}) = 0 &\Longleftrightarrow Tr_1^n(\lambda^{2^r+1} b^{-k^{-1}(2^r+1)}) = 0 \text{ for } k = 2^{2r} - 2^r + 1.
\end{aligned}$$

By Lemma 3, for every λ, at least 2^{m-1} elements $b \in GF(2^m)$ satisfy this condition. Therefore, $\widehat{v \cdot F}(\lambda) = 0$ for at least 2^{m-1} elements $v \in GF(2)^m$. For the other $\leq 2^{m-1}$ elements $v \in GF(2)^m$, $\widehat{v \cdot F}(\lambda) = \pm 2^{(n+1)/2}$ because the permutation x^k is almost bent. Thus for each $z \in GF(2)^m$,

$$\begin{aligned}
&\max_{a \in GF(2)^n - \{0\}} \left| \sum_{v \in GF(2)^m} (-1)^{v \cdot z} \widehat{v \cdot F}(a) \right| \\
\leq\ & 2^{m-1} \times \max_{v,a} |\widehat{v \cdot F}(a)| = 2^{m-1} \times 2^{(n+1)/2} = 2^{m+(n-1)/2},
\end{aligned}$$

Table 6: Comparison of Generalized Nonlinearity Bounds for Constructions from Proposition 2 and Theorem 5.

n	9	15	15	21	21
m	3	3	5	3	7
Lower Bound of GN_F from Proposition 2 part (2)	144	15488	12416	1041408	918528
Lower Bound of GN_F from Theorem 5	192	15872	14336	1044480	983040
N_F of Almost Bent functions	240	16256	16256	1047552	1047552

and

$$\begin{aligned}&\sum_{z\in GF(2)^m} \max_{a\in GF(2)^n-\{0\}} \left| \sum_{v\in GF(2)^m} (-1)^{v\cdot z} \widehat{v\cdot F}(a) \right| \\ \leq\ & 2^m \times 2^{m+(n-1)/2} = 2^{2m+(n-1)/2}.\end{aligned}$$

By substituting this inequality in equation (8), we have

$$nonlin_{gen}F \geq 2^{n-1} - \frac{1}{2^{m+1}} \times 2^{2m+(n-1)/2} = 2^{n-1} - 2^{(n-1)/2+m-1}.$$

Since $F(x)$ is a balanced function, we have $GN_F = nonlin_{gen}F$ and we are done. □

In Table 6, we illustrate that the generalized nonlinearity of the functions constructed from Theorem 5 is higher than that from Proposition 2 part (2). We also list the nonlinearity of almost bent functions for comparison.

8 Generalized Nonlinearity of Secondary Constructions

Secondary constructions produce Boolean functions with high nonlinearity, resiliency and other good cryptographic properties from other Boolean functions as building blocks. With respect to the generalized correlation attack, it would be useful to check if these constructions yield functions with high generalized nonlinearity. The first secondary construction we will look at is input composition with an invertible linear function. As in the case of nonlinearity, generalized nonlinearity is preserved in this case.

Proposition 3. *Let $F : GF(2)^n \rightarrow GF(2)^m$ be a vectorial Boolean function and let $L : GF(2)^n \rightarrow GF(2)^n$ be an invertible linear function. Then $GN_{F\circ L} = GN_F$.*

Proof. When computing $GN_{F\circ L}$, we compute the expression:

$$\begin{aligned} nonlin_{gen}F\circ L &= 2^{n-1} - 1/2\max_{g,w}\sum_x (-1)^{g(F(L(x))+w(F(L(x))\cdot x} \\ &= 2^{n-1} - 1/2\max_{g,w}\sum_y (-1)^{g(F(y))+w(F(y))\cdot L^{-1}(y)} \\ &= 2^{n-1} - 1/2\max_{g,w}\sum_y (-1)^{g(F(y))+(L^{-1})^*(w(F(y)))\cdot y} \\ &= 2^{n-1} - 1/2\max_{g,w'}\sum_y (-1)^{g(F(y))+w'(F(y))\cdot y} \\ &= nonlin_{gen}F. \end{aligned}$$

where L^* is the adjoint (transpose) of the linear function L and $w' : GF(2)^n \to GF(2)^n$ is defined by $w'(x) = [(L^{-1})^* \circ w](x)$. Furthermore $wt(F) = wt(F \circ L)$, therefore $GN_F = GN_{F\circ L}$. □

The next secondary construction we look at is output composition. One common candidate for output composition is the projection function, i.e. dropping output bits. For example, there are many known permutations with high nonlinearity [1] and by dropping output bits, we form vectorial Boolean functions with the same or higher nonlinearity.

Proposition 4. *Let $F : GF(2)^n \to GF(2)^m$ and $G : GF(2)^m \to GF(2)^k$ be balanced functions. Then $GN_{G\circ F} \geq GN_F$. If $G(z)$ is a permutation, then $GN_{G\circ F} = GN_F$.*

Proof. Let $\mathcal{G}, \mathcal{W}$ and $\mathcal{G}', \mathcal{W}'$ be the set of m-bit and k-bit Boolean functions in Definitions 1 and 2 respectively.

$$\max_{g'\in\mathcal{G}',w'\in\mathcal{W}'} \widehat{G\circ F}(g', w'_1, \ldots, w'_n) = \max_{g'\in\mathcal{G}',w'\in\mathcal{W}'} \hat{F}(g'\circ G, w'_1\circ G, \ldots, w'_n \circ G) \leq \max_{g\in\mathcal{G},w\in\mathcal{W}} \hat{F}(g, w_1, \ldots, w_n).$$

Therefore by equation (5), $nonlin_{gen}G\circ F \geq nonlin_{gen}F$. Note that $w' \in W'$ implies $w'\circ G \in W$ in the above inequality.

Since $F(x)$ is balanced, $nonlin_{gen}F = GN_F$ by remark 3. It is easy to deduce that $G\circ F$ is balanced if both F and G are balanced. Thus $nonlin_{gen}G\circ F = GN_{G\circ F}$ by remark 3 and we have $GN_{G\circ F} \geq GN_F$.

If $G(z)$ is a permutation, then $\{g\circ G | g \in \mathcal{G}\} = \mathcal{G}$ and $\{(w_1 \circ G, \ldots, w_n \circ G)|w \in \mathcal{W}\} = \mathcal{W}$. Thus we have $nonlin_{gen}G\circ F = nonlin_{gen}F$ which implies $GN_{G\circ F} = GN_F$. □

By Proposition 4, we see that output composition, e.g. dropping output bits, is good for enhancing security as it may increase the generalized nonlinearity.

However, output composition is not always good. It is well-known that linear error correction codes correspond to resilient vectorial Boolean functions [6], so they are easy to construct. However, one drawback is that they are linear and thus not useful for stream cipher applications. In [15], Zhang and

Zheng proved that if we compose the output of a resilient function with a permutation, then the order of resiliency is preserved in the resulting function. Thus they composed linear resilient functions with highly nonlinear permutation to form large sets of resilient vector functions with relatively high algebraic degree and nonlinearity. By applying Proposition 4 to Zhang and Zheng's construction, we immediately deduce that their construction is not secure.

Corollary 2. *Let $F : GF(2)^n \to GF(2)^m$ be a linear function and $G : GF(2)^m \to GF(2)^m$ be any bijective map. Then $G \circ F$ has null generalized nonlinearity. Thus the construction of Zhang and Zheng in [15, Section 4] is insecure for stream cipher applications.*

Proof. This is because $G(z)$ is a permutation implies $GN_{G\circ F} = GN_F$ by Proposition 4 and $GN_F = 0$ because $F(x)$ is linear. □

The consequence of Corollary 2 is that we can form exact linear equations between the output keystream and secret LFSR state bits. Thus the secret key can be recovered easily by simple linear algebra with less keystream bits than correlation attacks. This fact has also been pointed out in [2] where they remarked that the unrestricted nonlinearity of the Zhang-Zheng construction is 0.

Another common construction for vectorial resilient functions is concatenation. Let us look at the known results on this construction.

Proposition 5. *([15, Corollary 4]) Let $F_1 : GF(2)^{n_1} \to GF(2)^{m_1}$ be a t_1-resilient function and $F_2 : GF(2)^{n_2} \to GF(2)^{m_2}$ be a t_2-resilient function. Then $F_1||F_2 : GF(2)^{n_1+n_2} \to GF(2)^{m_1+m_2}$ defined by*

$$F_1||F_2(x, y) = (F_1(x), F_2(y))$$

is a t-resilient function where $t = \min(t_1, t_2)$.

By Proposition 5, given two smaller vectorial Boolean functions which are t-resilient, we can form a bigger Boolean function which is t-resilient. With respect to generalized correlation attack, we would like to know its generalized nonlinearity.

Proposition 6. *Let $F_1 : GF(2)^{n_1} \to GF(2)^{m_1}$ and $F_2 : GF(2)^{n_2} \to GF(2)^{m_2}$ be balanced functions. Then the generalized nonlinearity of their concatenation $F(x, y) = F_1(x)||F_2(y)$ satisfies:*

$$GN_F \leq 2^{n_1+n_2-1} - \frac{1}{2}(2^{n_1} - 2GN_{F_1})(2^{n_2} - 2GN_{F_2}).$$

Proof. Consider any $g_i : GF(2)^{m_i} \to GF(2)$, $i = 1, 2$ and any $w_{i,1}, \ldots, w_{i,n_i} : GF(2)^{m_i} \to GF(2)$,

$i = 1, 2$ where for all $z \in GF(2)^{m_i}$, $(w_{i,1}(z), \ldots, w_{i,n_i}(z)) \neq (0, \ldots, 0)$. We see that:

$$\begin{aligned}
& \widehat{F_1}(g_1(\cdot), w_{1,1}(\cdot), \ldots, w_{1,n_1}(\cdot)) \times \widehat{F_2}(g_2(\cdot), w_{2,1}(\cdot), \ldots, w_{2,n_2}(\cdot)) \\
= \;& \sum_x (-1)^{g_1(F_1(x)) + w_{1,1}(F_1(x))x_1 + \ldots + w_{1,n_1}(F_1(x))x_{n_1}} \sum_y (-1)^{g_2(F_2(y)) + w_{2,1}(F_2(y))y_1 + \ldots + w_{2,n_2}(F_2(y))y_{n_2}} \\
= \;& \sum_{x,y} (-1)^{g(F_1(x), F_2(y)) + w_1(F_1(x), F_2(y))x_1 + \ldots + w_{n_1+n_2}(F_1(x), F_2(y))y_{n_2}} \\
= \;& \widehat{(F_1, F_2)}(g(\cdot), w_1(\cdot), \ldots, w_{n_1+n_2}(\cdot)).
\end{aligned}$$

where we let $g : GF(2)^{m_1+m_2} \to GF(2)$ be defined by $g(z_1, z_2) = g_1(z_1) + g_2(z_2)$. Let

$$w_1(z_1, z_2) = w_{1,1}(z_1), \ldots, w_{n_1}(z_1, z_2) = w_{1,n_1}(z_1), w_{n_1+1}(z_1, z_2) = w_{2,1}(z_2), \ldots, w_{n_1+n_2}(z_1, z_2) = w_{2,n_2}(z_2).$$

Then for all $(z_1, z_2) \in GF(2)^{m_1+m_2}$, it is obvious that $(w_1(z_1, z_2), \ldots, w_{n_1+n_2}(z_1, z_2)) \neq (0, \ldots, 0)$.

Since on the left hand side of the above equations $g(\cdot)$ and $w_{i,j}(\cdot)$ can be any functions while the g, w_i defined on the right hand side are only functions on $(z_1, z_2) \in GF(2)^{m_1+m_2}$ of a special form, we have:

$$\max_{g_1, w_{1,i}} \widehat{F_1}(g_1(\cdot), w_{1,i}(\cdot)) \times \max_{g_2, w_{2,i}} \widehat{F_2}(g_2(\cdot), w_{2,i}(\cdot)) \leq \max_{g, w_i} \widehat{(F_1||F_2)}(g(\cdot), w_1(\cdot), \ldots, w_{n_1+n_2}(\cdot).$$

By substituting this inequality in equation (5), we get

$$nonlin_{gen}(F_1||F_2) \leq 2^{n_1+n_2-1} - \frac{1}{2}(2^{n_1} - 2nonlin_{gen}F_1)(2^{n_2} - 2nonlin_{gen}F_2). \tag{9}$$

Since $F_1(x)$ and $F_2(y)$ are balanced functions, we have $nonlin_{gen}F_i = GN_{F_i}$ by remark 3. Furthermore, it is easy to see that $(F_1(x), F_2(y))$ is a balanced function. Thus $nonlin_{gen}(F_1||F_2) = GN_{(F_1||F_2)}$ by remark 3. Thus we can substitute all the $nonlin_{gen}F$ in equation (9) by GN_F and we are done. □

By Proposition 6, we see that for a concatenated function to possess high generalized nonlinearity, both the component functions have to possess high generalized nonlinearity.

9 Generalized Nonlinearity of a Function with Very High Unrestricted Nonlinearity

In Corollary 2, we saw that although the Zhang-Zheng function has very high nonlinearity, it has zero unrestricted and generalized nonlinearity. A related question which we may ask is this: if a vectorial Boolean function has very high unrestricted nonlinearity, does it ensure that it has high generalized nonlinearity? We shall answer this question in this Section.

The following function from [4] has high unrestricted nonlinearity:

Proposition 7. *(Carlet-Prouff [4, Proposition 4]) Let $F : GF(2^{n/2}) \times GF(2^{n/2}) \to GF(2^{n/2})$ be defined by $F(x, y) = x/y$ if $y \neq 0$ and $F(x, y) = x$ when $y = 0$. Then $UN_F = 2^{n-1} - 2^{n/2}$.*

There are many contructions for balanced vectorial Boolean functions with high nonlinearity but to the best of our knowledge, there are none with nonlinearity higher than $2^{n-1} - 2^{n/2}$ when n is even [2]. Therefore this function has the best possible unrestricted nonlinearity. Next we shall compute its generalized nonlinearity.

Theorem 6. *Let $F : GF(2^{n/2}) \times GF(2^{n/2}) \to GF(2^{n/2})$ be defined by $F(x, y) = x/y$ if $y \neq 0$ and $F(x, y) = x$ when $y = 0$. Then $GN_F = 0$.*

Proof. From the proof of Theorem 2, we see that:

$$GN_F = 2^{n-1} - 1/4 \sum_{z \in GF(2^{n/2})} \max_{(a,b)\neq(0,0)} \left| \widehat{\phi_z}(a, b) \right|, \tag{10}$$

where $\phi_z(x, y) = 1$ if $z = F(x, y)$ and $\phi_z(x, y) = 0$ otherwise. Here we write GN_F for $nonlin_{gen}F$ because $F(x)$ is balanced.

For a subset $A \subset GF(2^{n/2}) \times GF(2^{n/2})$, let $Ind_A(x, y)$ be the indicator function of A, i.e. $Ind_A(x, y) = 1$ if $(x, y) \in A$, $Ind_A(x, y) = 0$ otherwise.

Let $z = F(x, y)$. We know that, for every z,

$$\phi_z(x, y) = Ind_{GF(2^{n/2})\times(z,1)}(x, y) - Ind_{(0,0)}(x, y) + Ind_{(z,0)}(x, y). \tag{11}$$

We need to compute the Walsh transform $\widehat{\phi_z}(a, b)$ when $(a, b) \neq (0, 0)$:

$$\begin{aligned} \widehat{\phi_z}(a, b) &= \sum_{x,y} (-1)^{\phi_z(x,y)+Tr(ax+by)} \\ &= \sum_{x,y} (-1)^{Tr(ax+by)} - 2 \sum_{x,y} \phi_z(x, y)(-1)^{Tr(ax+by)} \\ &= -2 \times \sum_{x,y} \phi_z(x, y)(-1)^{Tr(ax+by)} \text{ (because } (a, b) \neq (0, 0)). \end{aligned}$$

From equation (11), we see that:

$$\begin{aligned} \sum_{x,y} \phi_z(x, y)(-1)^{Tr(ax+by)} &= \sum_{y} (-1)^{Tr((az+b)y)} - 1 + (-1)^{Tr(az)} \\ &= 2^{n/2} Ind_{GF(2^{n/2})\times(1,z)}(a, b) - 1 + (-1)^{Tr(az)}. \end{aligned}$$

Thus the maximum of $|\widehat{\phi_z}(a, b)|$ is $2^{n/2+1}$ and we see that $GN_F = 0$ by substituting $\max_{(a,b)\neq(0,0)} |\widehat{\phi_z}(a, b)|$ in equation (10):

$$GN_F = 2^{n-1} - 1/4 \sum_{z \in GF(2^{n/2})} 2^{n/2+1} = 0.$$

□

Theorem 6 shows the surprising result that although a vectorial function may have very high unrestricted nonlinearity, it may still be possible for it to have zero generalized nonlinearity.

10 Conclusion

In this paper, we have introduced the generalized correlation attack on stream ciphers with vector output pseudo-random generators. When the generalized correlation attack is applied to the single output case, it is equivalent to a bilinear cryptanalysis. Unfortunately, we showed that it does not improve on the usual correlation attack in this case. However, for the case of multiple output bits, we found promising results which show that generalized correlation attack improves on current methods. An upper bound for generalized correlation is proved which shows that generalized correlation attack is more effective than the Zhang-Chan and usual correlation attack for vector Boolean functions in general. Efficient ways to find the best generalized linear approximations were also investigated. From experimental results, we saw that the bias of generalized linear approximations is much larger than that of the Zhang-Chan and usual correlation attack. Furthermore, there were several cases where the generalized linear equation is exact while the usual correlation and/or Zhang-Chan approach can only yield linear approximations (e.g. see example 2, Zhang-Zheng resilient functions). We also proved a lower bound for generalized nonlinearity and used it to construct functions with high generalized nonlinearity. Next, we found the generalized nonlinearity of some secondary constructions like pre-composition, post-composition and concatenation. Finally, we showed through the Zhang-Zheng function from [15] and Carlet-Prouff function from [4] that high nonlinearity and/or high unrestricted nonlinearity does not ensure high generalized nonlinearity.

There are still more open questions to investigate for further research. One is to find better lower and upper bounds for generalized correlation. We can also find the generalized nonlinearity of other useful secondary constructions like the direct sum (and its generalization by Carlet) and the vector Maiorana-McFarland construction. Another direction is to investigate practical implementations of the attack on actual stream ciphers.

References

[1] A. Canteaut, P. Charpin and H. Dobbertin, "Binary m-sequences with three-valued cross correlation: a proof of Welch's conjecture", *IEEE Trans. Inform. Theory*, vol. 46 no. 1, pp. 4-8, 2000.

[2] C. Carlet, "Vectorial Boolean Functions for Cryptography", to appear in *Boolean Methods and Models* published by Cambridge University Press, Eds Yves Crama and Peter Hammer. Can be found at http://www-rocq.inria.fr/codes/Claude.Carlet/chap-vectorial-fcts.pdf.

[3] C. Carlet, K. Khoo, C.W. Lim, C.W. Loe, "Generalized Correlation Analysis of Vectorial Boolean Functions", in *Proceedings of Fast Software Encryption 2007*, to appear in Lecture Notes in Computer Science, Springer-Verlag, 2007.

[4] C. Carlet and E. Prouff, "On a New Notion of Nonlinearity Relevant to Multi-Output Pseudo-Random Generators", LNCS 3006, *Selected Areas in Cryptography 2003*, pp. 291-305, Springer-Verlag, 2003.

[5] C. Carlet and E. Prouff, "Vectorial Functions and Covering Sequences", LNCS 2948, *International Conference on Finite Fields and Applications*, pp. 215-248, Springer-Verlag, 2003.

[6] B. Chor, O. Goldreich, J. Hastad, J. Friedman, S. Rudich and R. Smolensky, "The Bit Extraction Problem or t-resilient Functions", *IEEE Symposium on Foundations of Computer Science 26*, pp. 396-407, 1985.

[7] J.F. Dillon, "Multiplicative Difference Sets via Additive Characters", *Designs, Codes and Cryptography*, vol. 17, pp. 225-235, 1999.

[8] R. Gold, "Maximal Recursive Sequences with 3-valued Cross Correlation Functions", *IEEE Trans. Inform. Theory*, vol. 14, pp. 154-156, 1968.

[9] K.C. Gupta and P. Sarkar, "Improved Construction of Nonlinear Resilient S-boxes", LNCS 2501, *Asiacrypt 2002*, pp. 466-483, Springer-Verlag, 2002.

[10] E. Pasalic and S. Maitra, "Linear Codes in Constructing Resilient Functions with High Nonlinearity", LNCS 2259, *Selected Areas in Cryptography 2001*, pp. 60-74, Springer-Verlag, 2001.

[11] R. Rueppel, *Analysis and Design of Stream Ciphers*, Springer-Verlag, 1986.

[12] P. Sarkar, "The Filter-Combiner Model for Memoryless Synchronous Stream Ciphers", LNCS 2442, *Crypto 2002*, pp. 533-548, Springer-Verlag, 2002.

[13] T. Siegenthaler, "Decrypting a Class of Stream Ciphers using Ciphertexts only", *IEEE Transactions on Computers*, vol. C34, no. 1, pp. 81-85, 1985.

[14] M. Zhang and A. Chan, "Maximum Correlation Analysis of Nonlinear S-boxes in Stream Ciphers", LNCS 1880, *Crypto'2000*, pp. 501-514, Springer-Verlag, 2000.

[15] X.M. Zhang and Y. Zheng, "On Cryptographically Resilient Functions", *IEEE Transaction on Information Theory*, Vol. 43, no.5, pp. 1740-1747, 1997. (Also presented at *Eurocrypt'95*, LNCS 921, pp. 274-288, Springer-Verlag, 1995).