Nonlinearity, Local and Global Avalanche Characteristics of Balanced Boolean Functions

Pantelimon Stănică *

Auburn University Montgomery, Department of Mathematics
Montgomery, AL 36117, e-mail: stanpan@strudel.aum.edu.

Abstract

For a Boolean function f, define $\Delta_f(\alpha) = \sum_x \hat{f}(x)\hat{f}(x \oplus \alpha)$, $\hat{f}(x) = (-1)^{f(x)}$, the *absolute* indicator $\Delta_f = \max_{\alpha \neq 0} |\Delta_f(\alpha)|$, and the *sum-of-squares* indicator $\sigma_f = \sum_\alpha \Delta_f^2(\alpha)$. We construct a class of functions with good local avalanche characteristics, but bad global avalanche characteristics, namely we show that $2^{2n}(1+p) \leq \sigma_f \leq 2^{3n-2}, \Delta_f = 2^n$, where p is the number of linear structures (with even Hamming weight) of the first half of an SAC balanced Boolean function f. We also derive some bounds for the nonlinearity of such functions. It improves upon the results of Son *et al.* [5] and Sung *et al.* [7]. In our second result we construct a class of highly nonlinear balanced functions with good local and global avalanche characteristics. We show that for these functions, $2^{2n+2} \leq \sigma_f \leq 2^{2n+2+\epsilon}$ ($\epsilon = 0$ for n even and $\epsilon = 1$ for n odd).

Keywords: Cryptography; Boolean functions; Nonlinearity; Avalanche Characteristics

1 Definitions and Preliminaries

The design and evaluation of cryptographic functions requires the definition of design criteria. The Strict Avalanche Criterion (SAC) was introduced by Webster and Tavares [8] in a study of these criteria. A Boolean function is said to satisfy the SAC if comple-

*On leave from the Institute of Mathematics of Romanian Academy, Bucharest, Romania

menting a single bit results in changing the output bit with probability exactly one half. In [3], Preneel *et al.* introduced the *propagation criterion of degree k* (PC of degree k or $PC(k)$), which generalizes the SAC: a function satisfies the $PC(k)$ if by complementing at most k bits the output changes with probability exactly one half. Obviously $PC(1)$ is equivalent to the SAC property. The $PC(k)$ can be stated in terms of autocorrelation function. Let $V_n = \{\alpha_i | 1 \leq i \leq 2^n\}$ be the set of vectors of $\mathbf{Z}_2^n$ in lexicographical order. For a function on V_n, we say that f satisfies the $PC(k)$ if and only if

$$\sum_{x \in V_n} f(x) \oplus f(x \oplus c) = 2^{n-1}, \tag{1}$$

for all elements c with *Hamming weight* (the number of nonzero bits) $1 \leq wt(c) \leq k$, or equivalently, $\Delta_f(c) = 0$, where

$$\Delta_f(c) = \sum_{x \in V_n} \hat{f}(x)\hat{f}(x \oplus c)$$

is the autocorrelation function and $\hat{f}(x) = (-1)^{f(x)}$. There is also another variation of the PC, when one requires to have the above relation for an arbitrary subset of V_n, not necessarily for *all* x with $1 \leq wt(x) \leq k$ (see also [2]).

As many authors observed, the PC is a very important concept in designing cryptographic primitives used in data encryption algorithms and hash functions. However, the PC captures only local properties of the function. In order to improve the global analysis of cryptographically strong functions, Zhang and Zheng [11] introduced another criterion, which measures the *Global Avalanche Characteristics* (GAC) of a Boolean function. They proposed two indicators related to the GAC: the *absolute* indicator

$$\Delta_f = \max_{\alpha \neq 0} |\Delta_f(\alpha)|,$$

and the *sum-of-squares* indicator

$$\sigma_f = \sum_{\alpha} \Delta_f^2(\alpha).$$

The smaller σ_f, Δ_f the better the GAC of a function. Zhang and Zheng obtained some bounds on the two indicators:

$$2^{2n} \leq \sigma_f \leq 2^{3n}, 0 \leq \Delta_f \leq 2^n.$$

The upper bound for σ_f holds if and only if f is affine and the lower bound holds if and only if f is *bent* (satisfies the PC with respect to all $c \neq 0$).

There is an interest in computing bounds of the two indicators for various classes of Boolean functions. Recently, Son, Lim, Chee and Sung [5] proved

$$\sigma_f \geq 2^{2n} + 2^{n+3}, \tag{2}$$

when f is a balanced Boolean function, and Sung, Chee and Park [7] proved that if f also satisfies the PC with respect to $A \subset V_n$, $t = |A|$, then

$$\sigma_f \geq \begin{cases} 2^{2n} + 2^6(2^n - t - 1), \text{ if } 0 \leq t \leq 2^n - 2^{n-3} - 1, t \text{ odd} \\ 2^{2n} + 2^6(2^n - t + 2), \text{ if } 0 \leq t \leq 2^n - 2^{n-3} - 1, t \text{ even} \\ \left(1 + \frac{1}{2^n - 1 - t}\right) 2^{2n}, \text{ if } 2^n - 2^{n-3} - 1 < t \leq 2^n - 2. \end{cases} \tag{3}$$

The result (3) improves upon (2). Using the above result the authors of [7] have derived some new bounds for the nonlinearity of a balanced Boolean function satisfying the PC with respect to t vectors. We will improve their results significantly.

We need the following

Definition 1.

1. *We call* e_i *the* i*-th* basis vector *of* V_n.

2. *An* affine *function is a Boolean function of the form* $f(x) = \oplus_{i=1}^{n} c_i x_i \oplus c$. f *is called* linear *if* $c = 0$.

3. *The* truth table *of* f *is the binary sequence* $f = (v_1, v_2, \ldots, v_{2^n})$, *where* $v_i = f(\alpha_i)$.

4. *The* Hamming weight *of a binary vector* v, *denoted by* $wt(v)$ *is defined as the number of ones it contains. The* Hamming distance *between two functions* $f, g : V_n \to V_1$, *denoted by* $d(f, g)$ *is defined as* $wt(f \oplus g)$. f *is* balanced *is* $wt(f) = 2^{n-1}$.

5. *The* nonlinearity *of a function* f, *denoted by* N_f *is defined as* $\min_{l \in A_n} d(f, l)$, *where* A_n *is the class of all affine function on* V_n.

6. *A vector* $0 \neq \alpha \in V_n$ *is a* linear structure *of* f *if* $f(x) \oplus f(x \oplus \alpha)$ *is constant for all* x.

7. *If* X, Y *are two strings of the same length,* $(X|Y)$ *means that* X *and* Y *occupy the same positions in the first and the second half of some function.*

8. *Define the set of* 4*-bit blocks* $T = \{A = 0,0,1,1;\ \bar{A} = 1,1,0,0;\ B = 0,1,0,1;\ \bar{B} = 1,0,1,0;\ C = 0,1,1,0;\ \bar{C} = 1,0,0,1;\ D = 0,0,0,0;\ \bar{D} = 1,1,1,1\}$.

9. *If some bits of an affine function* l *agree with the the corresponding bits in a function* f, *we say that* l cancels *those bits in* f.

10. *If* u *is a given string and* g *is a Boolean function, we use* $u^g =$ the string of bits in g which occupy the same positions as the bits in the string u.

11. If a Boolean string is a concatenation of either $A/\bar{A}$ or $B/\bar{B}$ or $C/\bar{C}$ or $D/\bar{D}$ we say that it is based on *A or B or C or D.*

12. By $MSB(\cdot)$ we denote the most significant bit *of the enclosed argument.*

2 The First Result

In this section the function f will denote a balanced Boolean function which satisfies the *SAC*. We will consider *SAC* functions constructed using some ideas of [9, 10] (see also [1] for another version of the construction). Define $\mathbf{1} \cdot x = \oplus_{i=1}^{n-1} x_i$, for $x = (x_1, \ldots, x_{n-1})$. Let $g : V_{n-1} \to V_1$ denote the Boolean function $\mathbf{1} \cdot x \oplus b$, $b \in V_1$, which satisfies $g(x) = \bar{g}(x \oplus a)$, for any element a of odd Hamming weight. For a vector $v \in V_n$, we denote by $v' \in V_{n-1}$ the $n-1$ least significant bits in v. In [9, 10, 1] or [6] it is proved that functions of the form

$$f = (h \mid h \oplus g), \text{ or } f = (h \mid l \oplus g), \tag{4}$$

are *SAC* functions, where h is an arbitrary function on V_{n-1} and $l(x) = h(x \oplus a)$, $wt(a) = odd$. Let $\bar{x}$ be the complement of x.

Proposition 2. *The functions (4) can be written as* $f(x_1, \ldots, x_{n-1}, x_n) =$

$$\bar{x}_n h(x_1, \ldots, x_{n-1}) \oplus x_n \left(h(x_1, \ldots, x_{n-1}) \oplus_{i=1}^{n-1} x_i \oplus b \right) \text{ or}$$

$$\bar{x}_n h(x_1, \ldots, x_{n-1}) \oplus x_n \left(h(x_1, \ldots, \bar{x}_k, \ldots, x_{n-1}) \oplus_{i=1}^{n-1} x_i \oplus b \right),$$

(an odd number of input bits x_k are complemented), for an arbitrary Boolean function h defined on V_{n-1} and $b \in V_1$.

Proof. Straightforward using the definition of g and concatenation. □

First, we consider the case of balanced Boolean functions f defined on $V_n, n \geq 3$ of the form (4) such that h has linear structures. We denote by $\mathcal{L}_h^{even}$ the number of nonzero linear structures of h with even Hamming weight. We take a to be an element of odd Hamming weight. In our next theorem we compute the indicators for a class of functions satisfying the SAC. We remark that the global characteristics are not good for these functions although the local ones are (the functions are SAC).

Theorem 3. *If f is a balanced Boolean function of the form $f = (h|l \oplus g)$, $l(x) = h(x)$ or $l(x) = h(x \oplus a)$, h an arbitrary Boolean function with $\mathcal{L}_h^{even} \geq 1$ and g as before, we have*

$$2^{2n}(1 + \mathcal{L}_h^{even}) \leq \sigma_f \leq 2^{3n-2}. \tag{5}$$

Proof. Zhang and Zheng [12] proved that for functions satisfying the *SAC*, the nonlinearity satisfies

$$N_f \geq 2^{n-2}. \tag{6}$$

In [5] the following inequality is obtained:

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{\sigma_f/2^n}. \tag{7}$$

Using (6) and (7) we obtain easily the right inequality of (5), that is

$$\sigma_f \leq 2^{3n-2}.$$

From the proof of Lemma 1 of [7] we get that σ_f satisfies

$$\sigma_f = \sum_x \Delta_f^2(x) = 2^6 \sum_x (b_x - 2^{n-3})^2 + 2^{n+4} \sum_x (b_x - 2^{n-3}),$$

where $b_x = \frac{1}{2}\sum_y f(y)f(y \oplus x)$. Using the trivial identity $ab = \frac{1}{2}(a + b - a \oplus b)$ and the fact that f is balanced, we get $b_x = \frac{1}{4}\sum_y (f(y) + f(y \oplus x) - f(y) \oplus f(y \oplus x)) = 2^{n-2} - \frac{1}{4}\sum_y f(y) \oplus f(y \oplus x)$. We note that f satisfies the PC with respect to x if and only if $b_x = 2^{n-3}$. Since f is balanced, $\sum_x (b_x - 2^{n-3}) = 0$. It follows that

$$\sigma_f = 2^{2n} + 2^6 \sum_{wt(x) \geq 2} (b_x - 2^{n-3})^2.$$

We want to evaluate $\sum_{wt(x)\geq 2}(b_x - 2^{n-3})^2$. In order to do that we have to compute

$$S_x = \sum_{y \in V_n} f(y) \oplus f(y \oplus x).$$

Case 1: $MSB(x) = 0$.

In this case

$$\begin{aligned} S_x = \sum_{y \in V_n} f(y) \oplus f(y \oplus x) = \sum_{i=1}^{2^{n-1}} h(v_i') \oplus h(v_i' \oplus x') + \\ \sum_{i=1}^{2^{n-1}} h(v_i') \oplus h(v_i' \oplus x') \oplus g(v_i') \oplus g(v_i' \oplus x'). \end{aligned} \tag{8}$$

Case 1.1: $wt(x') = even$.

In this case, since g satisfies $g(x) = \bar{g}(x \oplus a)$ for any element with odd Hamming weight, it follows that $g(v_i' \oplus x') = g(v_i')$. Therefore, the equation (8) becomes

$$S_x = 2 \sum_{i=1}^{2^{n-1}} h(v_i') \oplus h(v_i' \oplus x').$$

When x' is a linear structure of h, $S_x = 2^n c$, where $c = h(0) \oplus h(0 \oplus x')$.

Case 1.2: $wt(x') = odd$.

Then $g(v_i' \oplus x') = \bar{g}(v_i')$ and (8) becomes

$$S_x = \sum_{i=1}^{2^{n-1}} h(v_i') \oplus h(v_i' \oplus x') + \sum_{i=1}^{2^{n-1}} h(v_i') \oplus h(v_i' \oplus x') \oplus 1 = 2^{n-1}.$$

Case 2: $MSB(x) = 1$.

In this case, S_x can be evaluated as follows:

$$S_x = \sum_{i=1}^{2^{n-1}} h(v_i') \oplus h(v_i' \oplus x') \oplus g(v_i' \oplus x') + \sum_{i=1}^{2^{n-1}} h(v_i') \oplus h(v_i' \oplus x') \oplus g(v_i').$$

Case 2.1: $wt(x') = even$.

Since $g(v_i') = g(v_i' \oplus x')$, we get

$$S_x = 2 \sum_{i=1}^{2^{n-1}} h(v_i') \oplus h(v_i' \oplus x') \oplus g(v_i').$$

Case 2.2: $wt(x') = odd$.

Since $g(v_i' \oplus x') = \bar{g}(v_i')$, we get

$$S_x = \sum_{i=1}^{2^{n-1}} h(v_i') \oplus h(v_i' \oplus x') + \sum_{i=1}^{2^{n-1}} h(v_i') \oplus h(v_i' \oplus x') \oplus 1 = 2^{n-1}.$$

From the above analysis we deduce that:

Case 1.1: $b_x = 2^{n-2} - 2^{-2} S_x$, and if x' is a linear structure for h, $b_x = 2^{n-2}$ or $b_x = 0$.

Case 1.2: $b_x = 2^{n-3}$.

Case 2.1: $b_x = 2^{n-2} - 2^{-2} S_x$, and if x' is a linear structure for h, $b_x = 2^{n-3}$.

Case 2.2: $b_x = 2^{n-3}$.

We observe that the only cases where we do not know precisely b_x are when x is an element of odd Hamming weight with x' not a linear structure for h.

We deduce that in the case 1.1 with x' a linear structure for h,

$$(b_x - 2^{n-3})^2 = 2^{2(n-3)}.$$

Now, returning to the computation of σ_f, with the new results we get

$$\sigma_f = 2^{2n} + 2^6 \sum_{wt(x)\geq 2} (b_x - 2^{n-3})^2 \geq$$
$$2^{2n} + 2^6 2^{2(n-3)} \mathcal{L}_h^{even} = 2^{2n} \left(1 + \mathcal{L}_h^{even}\right).$$

□

With the same data as in the previous theorem we obtain

Corollary 4. *For $n \geq 3, \Delta_f = 2^n$.*

Proof. The corollary follows from the proof of the theorem. For a Boolean balanced function, $\Delta_f(x) = 2^3 b_x - 2^n$. Therefore for any x, such that x' is a linear structure of h of even Hamming weight, we have $b_x = 0$ or 2^n. Thus $\Delta_f = \max_{x \in V_n} |\Delta_f(x)| = 2^n$. □

The previous corollary can also be deduced from Lemma 7 of [11], observing that if x' is a linear structure of h with even Hamming weight, then $(0, x')$ is a linear structure for f.

The following is an easy consequence of the previous theorem. It shows that the theorem gives tight bounds.

Corollary 5. *For a balanced Boolean* SAC *function f given by* (4)*, where h is affine we have the following equation*

$$\sigma_f = 2^{3n-2}.$$

Proof. This follows from the fact that any nonzero element of V_n is a linear structure for an affine function. □

Now we turn our attention to the nonlinearity of such functions. Using

$$N_f \leq 2^{n-1} - 2^{-n/2-1}\sqrt{\sigma_f},$$

and $\sigma_f \geq 2^{2n}(1 + \mathcal{L}_h^{even})$, we get the corollary

Corollary 6. *Let f be as in the Theorem 3. Then, the nonlinearity satisfies*

$$2^{n-2} \leq N_f \leq 2^{n-1} - 2^{n/2-1}\sqrt{1 + \mathcal{L}_h^{even}}. \tag{9}$$

If f satisfies the conditions of Corollary 5, then we have

$$N_f = 2^{n-2}. \tag{10}$$

Since $2^n + 2^{n/2+3} + 2^4 < 2^n\,(1 + \mathcal{L}_h^{even})\,,$ if $\mathcal{L}_h^{even} \geq 1$, it follows that the bounds (9) or (10) are better than the result of Zhang and Zheng, who proved in [12] that

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + 2^{n/2+3} + 2^4},\ \text{if } n \text{ is even.}$$

Sung *et al.* [7] obtained the following upper bound for the nonlinearity

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + 2^6 - \frac{(n+1)2^6}{2^n}},\ \text{if } n > 2 \text{ is odd and}$$
$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + 2^6 - \frac{(n-1)2^6}{2^n}},\ \text{if } n \text{ is even,}$$

which is certainly weaker than the bound we have obtained.

3 Highly nonlinear balanced SAC functions with good GAC

In the previous section we constructed a class of balanced functions with good local avalanche characteristics, but bad global avalanche characteristics. In this section we

will use some results from [6] to construct balanced Boolean SAC functions of nonlinearity at least $2^n - 2^{[(n+1)/2]}$, with good GAC.

From a result we like to call *Folklore Lemma* (see [6]), we know that for any affine function l, if L is the first string of length 2^s in l, then the next string of the same length will be L or $\bar{L}$. A consequence of this fact is that any affine function is made up as a concatenation of blocks $A/\bar{A}$ or $B/\bar{B}$ or $C/\bar{C}$ or $D/\bar{D}$.

Our next theorem was proved initially in a more general form. However, its proof relied heavily on results available only in [6], so we decided to provide here a complete proof for a slightly restricted subclass. Moreover, for this subclass we can provide better results, especially for even dimensions, which makes it all worthwhile. For the purpose of easy computation, we define a transformation $\mathcal{O}(g)$ ("opposite") which maps an affine function based on $M \in T$, into an affine function based on the same block M, having the self-invertible property $\mathcal{O}(\mathcal{O}(g)) = g$. If $g = X_1 X_2 \dots X_{2^{n-2}}$, then $\mathcal{O}(g) = Y_1 Y_2 \dots Y_{2^{n-2}}$ is constructed by the following *Algorithm*, supported by the Folklore Lemma:

Step 1. $Y_1 = X_1$.

Step $i+2$. *For any* $0 \leq i \leq n-3$, *if* $X_{2^i+1} \dots X_{2^{i+1}} = X_1 \dots X_{2^i}$, *then* $Y_{2^i+1} \dots Y_{2^{i+1}} = \bar{Y}_1 \dots \bar{Y}_{2^i}$. *If* $X_{2^i+1} \dots X_{2^{i+1}} = \bar{X}_1 \dots \bar{X}_{2^i}$, *then* $Y_{2^i+1} \dots Y_{2^{i+1}} = Y_1 \dots Y_{2^i}$.

Remark 7. *The results will not change if we take the first block* $Y_1 = \bar{X}_1$.

By induction we can easily prove

Lemma 8. $\mathcal{O}(\bar{g}) = \overline{\mathcal{O}(g)}$.

The following theorem is a construction for balanced functions of high nonlinearity with very good local and global avalanche characteristics. Define $\lfloor x \rfloor$ (*the floor function*) to be the largest integer less than or equal to x. For easy writing we let $h_i = \mathcal{O}(g_i)$.

Theorem 9. *For $n = 2k \geq 8$ (or $n = 2k+1 \geq 9$) let f to be the function obtained by concatenating 2^{k-1} segments T_i. For each $1 \leq i \leq 2^{k-2}$, T_i is of the form*

$$(g_i h_i g_i \bar{h}_i | \bar{h}_i g_i h_i g_i) \tag{11}$$

and the segment $T_{i+2^{k-2}}$ is of the form

$$(h_i \bar{g}_i \bar{h}_i \bar{g}_i | \bar{g}_i \bar{h}_i \bar{g}_i h_i), \tag{12}$$

respectively, where the functions g_i are affine functions on V_1^{k-2} (or V_1^{k-1}). Furthermore, we impose the following conditions:

(i) Exactly a quarter of the functions g_i are based on each of the 4-bit blocks A, B, C, D.

(ii) For any $1 \leq i \neq j \leq 2^{k-2}$, the functions $g_i \oplus g_j$ are balanced.

Then the function f is balanced, satisfies the SAC*, has the nonlinearity $N_f \geq 2^{n-1} - 2^{\lfloor \frac{n+1}{2} \rfloor}$ and the sum-of-squares indicator satisfies*

$$2^{2n+2} \leq \sigma_f \leq 2^{2n+2+\epsilon},$$

where $\epsilon = 0, 1$ if n is even, respectively, odd.

Proof. We will prove the theorem for the case of n even, that is $n = 2k$, pointing out, whenever necessary, the differences for the case of odd n. The function f can be written as

$$\begin{array}{ll} \big(g_1 h_1 g_1 \bar{h}_1 \cdots g_{2^{k-2}} h_{2^{k-2}} g_{2^{k-2}} \bar{h}_{2^{k-2}} & h_1 \bar{g}_1 \bar{h}_1 \bar{g}_1 \cdots h_{2^{k-2}} \bar{g}_{2^{k-2}} \bar{h}_{2^{k-2}} \bar{g}_{2^{k-2}} \\ \bar{h}_1 g_1 h_1 g_1 \cdots \bar{h}_{2^{k-2}} g_{2^{k-2}} h_{2^{k-2}} g_{2^{k-2}} & \bar{g}_1 \bar{h}_1 \bar{g}_1 h_1 \cdots \bar{g}_{2^{k-2}} \bar{h}_{2^{k-2}} \bar{g}_{2^{k-2}} h_{2^{k-2}} \big). \end{array} \tag{13}$$

The fact that f is balanced can be seen by pairing the functions g with $\bar{g}$ and h with $\bar{h}$ in the two segments T_i and $T_{i+2^{k-2}}$. To show that f satisfies the SAC we use some results of Cusick and Stănică, that is Lemma 1 or relation (8) of [1], which says that a function $f = (v_1, \ldots, v_{2^n}) = X_1 \cdots X_{2^{n-2}}$ satisfies the SAC if and only if

$$\begin{aligned} &(w_1 w_{2^{i-1}+1} + w_2 w_{2^{i-1}+2} + \cdots + w_{2^{i-1}} w_{2^i}) + \\ &(w_{2^i+1} w_{2^i+2^{i-1}+1} + \cdots + w_{2^i+2^{i-1}} w_{2^{i+1}}) + \cdots + \\ &(w_{2^n-2^i+1} w_{2^n-2^{i-1}+1} + \cdots + w_{2^n-2^{i-1}} w_{2^n}) = 0, \end{aligned} \tag{14}$$

for each $i = 1, 2, \ldots, n$, where $w_i = (-1)^{v_i}$, or equivalently (if $i \geq 3$),

$$(X_1 \odot X_{2^{i-3}+1} + \cdots + X_{2^{i-3}} \odot X_{2^{i-2}}) + \cdots = 0, \tag{15}$$

for each $i = 3, 4, \ldots, n$, where $M \odot N$ is equal to the number of 0's minus the number of 1's in $M \oplus N$ If we associate the 4-bit blocks $\{A, \bar{A}\} \Longleftrightarrow \{B, \bar{B}\}$ and $\{C, \bar{C}\} \Longleftrightarrow \{D, \bar{D}\}$, we see that, for $i \leq 2$, the relation (14) holds. Obviously, if $M \oplus N$ is balanced, then $M \odot N = 0$. Thus, in the sum (15) the sum in each parenthesis is zero, except perhaps the ones based entirely on $D, \bar{D}$ (which are the only unbalanced 4-bit blocks in T). However, those terms will have an antidote in another parenthesis. For instance, since $D \odot D = -D \odot \bar{D} = 4$, $D \odot D$ will have the antidote $D \odot \bar{D}$, according to the form of our functions.

In order to compute the nonlinearity of f we have counted the bits at which our function differ from any linear or affine function. Intuitively, we need to prove that on average an affine function cannot cancel to many blocks in a segment. Precisely, we show that given any two segments U_1, U_2 in the same half of f, based on the same block $M \in T$, then $wt(U_1 U_2 \oplus U_1^l U_2^l) \geq 2^{k-1} + 2^k$, for any affine function l based on the

same block M. This is shown easily using the folklore lemma, and observing that on the positions of U_1U_2, l can have only the following forms: $(LLLLLLLL|LLLLLLLL)$, $(LLLLLLLL|\bar{L}\bar{L}\bar{L}\bar{L}\bar{L}\bar{L}\bar{L}\bar{L})$, $(LL\bar{L}\bar{L}LLL\bar{L}\bar{L}|LL\bar{L}\bar{L}LLL\bar{L}\bar{L})$, etc. Since all cases are treated similarly, we may assume that $(U_1^lU_2^l) = (LLLLLLLL|LLLLLLLL)$ (recall the definition of U^l). Without loss of generality we may assume that U_1, U_2 are in the first half of f and $U_1 = (g_1h_1g_1\bar{h}_1)|\bar{h}_1g_1h_1g_1)$, $U_2 = (g_2h_2g_2\bar{h}_2)|\bar{h}_2g_2h_2g_2)$. Thus

$$\begin{aligned} wt(U_1U_2 \oplus U_1^lU_2^l) &= 2wt(g_1 \oplus L) + wt(h_1 \oplus L) + wt(\bar{h}_1 \oplus L) \\ &\quad +2wt(g_2 \oplus L) + wt(h_2 \oplus L) + wt(\bar{h}_2 \oplus L) \\ &\quad +wt(\bar{h}_1 \oplus L) + wt(h_1 \oplus L) + 2wt(g_1 \oplus L) \\ &\quad +wt(\bar{h}_2 \oplus L) + wt(h_2 \oplus L) + 2wt(g_2 \oplus L) \\ &= 4wt(g_1 \oplus L) + 4wt(g_2 \oplus L) + 2^k \\ &\geq 4wt(g_1 \oplus g_2) + 2^k = 2^{k-1} + 2^k. \end{aligned}$$

Here we used $wt(a \oplus c) + wt(b \oplus c) \geq wt(a \oplus b)$, the fact that $g_i \oplus g_j$ is balanced and $wt(a \oplus b) + wt(a \oplus \bar{b}) = 2^{k-2}$, if $a, b, c \in V_{k-2}$. Next, we compute $wt(f \oplus l)$. One may assume that l is based on A. From the part of f that does not contain $A, \bar{A}$ we get $3 \cdot 2^{2k-3} = 2^{2k-1} - 2^{2k-3}$ units for the weight (we recall that only a quarter of all blocks contain $A, \bar{A}$). We consider now the part of f based on A. Using the previous result, we deduce that in the worst case (minimum weight), l cancels completely at most four functions from each half, and from the rest of the part of f based on A, half of the blocks are cancelled. Since there are 2^k functions based on A and we cancel 8 functions, we gather that there remain $2^k - 8$ functions uncancelled. Since

each uncancelled function contributes 2^{k-3} units to the weight (recall that if two affine functions g, l are not equal or complementary, their sum is balanced), we get $2^{2k-3} - 2^k$ units contributed to the weight by the part based on A, so the nonlinearity is at least $2^{2k-1} - 2^{2k-3} + 2^{2k-3} - 2^k = 2^{2k-1} - 2^k$. In the odd case we get $N_f \geq 2^{2k-1} - 2^{k+1}$ (the lengths of the affine functions g_i, h_i double, while the number of segments remains the same), by a similar argument.

Now, since $N_f \leq 2^{n-1} - 2^{-\frac{n}{2}-1}\sqrt{\sigma_f}$ and from the above analysis $N_f \geq 2^{n-1} - 2^{\lfloor \frac{n+1}{2} \rfloor}$ we get

$$2^{n-1} - 2^{\lfloor \frac{n+1}{2} \rfloor} \leq 2^{n-1} - 2^{-\frac{n}{2}-1}\sqrt{\sigma_f},$$

which will produce our right hand side inequality

$$\sigma_f \leq 2^{2n+2}, \text{ if } n \text{ is even, and } \sigma_f \leq 2^{2n+3}, \text{ if } n \text{ is odd.}$$

In order to evaluate S_x for suitably chosen x we apply the same technique as in the proof of Theorem 3. For $x = e_i \oplus e_j, i < j$, let

$$\begin{aligned} S_x = \sum_{y \in V_n} f(y) \oplus f(y \oplus x) = \sum_{s=1}^{2^n} f(v_s) \oplus f(v_s \oplus e_i \oplus e_j) = \\ 2[f(v_1) \oplus f(v_{2^{j-1}+2^{i-1}+1}) + \cdots + f(v_{2^{i-1}}) \oplus f(v_{2^{j-1}+2^i}) + \\ f(v_{2^{i-1}+1}) \oplus f(v_{2^{j-1}+1}) + \cdots + f(v_{2^{i-1}+2^{i-1}}) \oplus f(v_{2^{j-1}+2^{i-1}})] + \cdots . \end{aligned} \tag{16}$$

Using the form of our functions and taking $x = e_{n-1} \oplus e_n$, we get

$$S_{e_{n-1} \oplus e_n} = 2 \sum_{g_i, h_i} (g_i \oplus \bar{g}_i + h_i \oplus \bar{h}_i + g_i \oplus \bar{g}_i + \bar{h}_i \oplus h_i) = 2^n.$$

Thus, $(b_{e_{n-1} \oplus e_n} - 2^{n-3})^2 = 2^{2n-6}$.

Now, we take $x = e_i \oplus e_j \oplus e_r, i < j < r$. Thus, we get

$$\begin{aligned} S_x = \sum_{y \in V_n} f(y) \oplus f(y \oplus x) = \\ \sum_{s=1}^{2^n} f(v_s) \oplus f(v_s \oplus e_i \oplus e_j \oplus e_r) = \\ 2[f(v_1) \oplus f(v_{2^{r-1}+2^{j-1}+2^{i-1}+1}) + \cdots + \\ f(v_{2^{i-1}}) \oplus f(v_{2^{r-1}+2^{j-1}+2^i}) + \\ f(v_{2^{i-1}+1}) \oplus f(v_{2^{r-1}+2^{j-1}+1}) + \cdots + \\ f(v_{2^{i-1}+2^{i-1}}) \oplus f(v_{2^{r-1}+2^{j-1}+2^{i-1}})] + \cdots . \end{aligned} \tag{17}$$

Now, taking $x = e_{k-1} \oplus e_k \oplus e_n$ and $n = 2k$, we obtain

$$\begin{aligned} S_{e_{k-1} \oplus e_k \oplus e_n} = 2 \Big[\big(f(v_1) \oplus f(v_{2^{n-1}+2^{k-1}+2^{k-2}+1}) + \cdots + \\ f(v_{2^{k-2}}) \oplus f(v_{2^{n-1}+2^k}) \big) + \\ \big(f(v_{2^{k-2}+1}) \oplus f(v_{2^{n-1}+2^{k-1}+1}) + \cdots + \\ f(v_{2^{k-2}+2^{k-2}}) \oplus f(v_{2^{n-1}+2^{k-1}+2^{k-2}}) \big) + \\ \big(f(v_{2^{k-1}+1}) \oplus f(v_{2^{n-1}+2^{k-2}+1}) + \cdots + \\ f(v_{2^{k-1}+2^{k-2}}) \oplus f(v_{2^{n-1}+2^{k-1}}) \big) + \\ \big(f(v_{2^{k-1}+2^{k-2}+1}) \oplus f(v_{2^{n-1}+1}) + \cdots + \\ f(v_{2^k}) \oplus f(v_{2^{n-1}+2^{k-2}}) \big) \Big] + \cdots \end{aligned}$$

for any function f. In particular, for the functions in our class, we get

$$\begin{aligned} S_{e_{k-1} \oplus e_k \oplus e_n} &= 2 \sum_{s=1}^{2^{k-2}} (g_s \oplus g_s + h_s \oplus h_s + g_s \oplus g_s + \bar{h}_s \oplus \bar{h}_s) \\ &+ 2 \sum_{s=1}^{2^{k-2}} (h_s \oplus h_s + \bar{g}_s \oplus \bar{g}_s + \bar{h}_s \oplus \bar{h}_s + \bar{g}_s \oplus \bar{g}_s) = 0. \end{aligned}$$

Similarly, $S_{e_{k-1}\oplus e_k\oplus e_{n-1}} = 2^n$. Thus, $b_{e_{k-1}\oplus e_k\oplus e_n} = 2^{n-2}$ and $b_{e_{k-1}\oplus e_k\oplus e_{n-1}} = 0$.

In any of the three cases $x = e_{n-1} \oplus e_n, e_{k-1} \oplus e_k \oplus e_{n-1}, e_{k-1} \oplus e_k \oplus e_n$, we have $(b_x - 2^{n-3})^2 = 2^{2n-6}$. Thus,

$$\sigma_f \geq 2^{2n} + 2^6 2^{2n-6} + 2^6 2^{2n-6} + 2^6 2^{2n-6} = 2^{2n+2}.$$

□

Corollary 10. *For f given by Theorem 9, we have $\Delta_f = 2^n$.*

Proof. We know that $\Delta_f(x) = 2^3 b_x - 2^n$. Therefore,

$$\Delta_f(e_{k-1} \oplus e_k \oplus e_n) = 2^3 \cdot 2^{n-2} - 2^n = 2^n,$$

and the result follows. □

Corollary 11. *If n is even and f is given as in Theorem 9, then $\sigma_f = 2^{2n+2}$, $N_f = 2^{n-1} - 2^{\frac{n}{2}}$, and f is PC with respect to all but four vectors. Moreover, the three nonzero vectors, which do not satisfy the propagation criterion, are linear structures for f.*

Proof. We proved that, if n is even, then $\sigma_f = 2^{2n+2}$. If there is an x not equal to the four displayed vectors in the proof of Theorem 9, for which f is not PC, then $b_x \neq 2^{n-3}$. If so, then by the same argument we would get $\sigma_f > 2^{2n+2}$, which is not true. So f is PC with respect to all but four vectors. In [13], Zhang and Zheng proved that, if a function satisfies the PC with respect to all but four vectors, then n must be even, the nonzero vectors, where the propagation criterion is not satisfied, must be linear structures and $N_f = 2^{n-1} - 2^{n/2}$. We have the result. □

As we can see the bounds are extremely good, not too far from that of bent functions, improving upon any known ones. We suspect we can modify the construction to improve the nonlinearity for the odd dimension as well, and we will pursue this idea elsewhere.

Remark 12. *If the conditions imposed in Theorem 9 hold for g_i, they certainly hold for $h_i = \mathcal{O}(g_i)$ as well.*

4 Examples and Further Research

An example of a function satisfying the conditions of Theorem 9 with $h_i = \mathcal{O}(g_i)$, for $n = 8$ is

$$AAA\bar{A}BBB\bar{B}CCC\bar{C}DDDD\bar{D}A\bar{A}\bar{A}\bar{A}B\bar{B}\bar{B}\bar{B}C\bar{C}\bar{C}\bar{C}D\bar{D}\bar{D}\bar{D}$$

$$\bar{A}AAA\bar{B}BBB\bar{C}CCC\bar{D}DDD\bar{A}\bar{A}\bar{A}A\bar{B}\bar{B}\bar{B}B\bar{C}\bar{C}\bar{C}C\bar{D}\bar{D}\bar{D}D,$$

which is balanced, SAC (actually, it is PC with respect to all but $\mathbf{0}, e_7 \oplus e_8, e_3 \oplus e_4 \oplus e_8, e_3 \oplus e_4 \oplus e_7$), has nonlinearity 112 and the sum-of-squares indicator attains the upper bound, $\sigma_f = 262,144 = 2^{2 \cdot 8+2}$. The algebraic normal form is $x_1 + x_7 + x_1x_5 + x_1x_6 + x_2x_5 + x_2x_6 + x_3x_8 + x_4x_7 + x_4x_8 + x_5x_6$.

We can define the transformation $\mathcal{O}$ using the same algorithm starting with the first bit, rather than the first block, so $\mathcal{O}(A) = B, \mathcal{O}(C) = D$, etc., obtaining a result similar to our Theorem 9. It seems that the algebraic degree increases for that class, but we were not able to prove that in its full generality. An example of a function constructed using this idea, for $n = 8$, is

$$AB A\bar{B}BAB\bar{A}CDC\bar{D}DCD\bar{C}B\bar{A}\bar{B}\bar{A}A\bar{B}\bar{A}\bar{B}C\bar{D}\bar{C}\bar{D}D\bar{C}\bar{D}\bar{C}$$

$$\bar{B}ABA\bar{A}BAB\bar{D}CDC\bar{C}DCD\bar{A}\bar{B}\bar{A}B\bar{B}\bar{A}\bar{B}A\bar{D}\bar{C}\bar{D}C\bar{C}\bar{D}\bar{C}D.$$

It turns out that the above function is balanced, has nonlinearity precisely 112, it is SAC (in fact, it is PC with respect to 252 vectors), the sum-of-squares indicator attains the upper bound, $\sigma_f = 262,144 = 2^{2\cdot 8+2}$. The algebraic normal form is $x_1+x_7+x_1x_5+x_1x_6+x_1x_7+x_1x_8+x_2x_5+x_2x_6+x_2x_7+x_2x_8+x_3x_8+x_4x_7+x_4x_8+x_5x_6+x_6x_7+x_6x_8+x_2x_3x_7+x_2x_3x_8$.

Another venue of further research would be the construction of a class of functions with these good local and global avalanche characteristics and high nonlinearity, using blocks in the complementary set of T, namely $T' = \{U = 1,0,0,0;\ \bar{U} = 0,1,1,1;\ V = 0,0,0,1;\ \bar{V} = 1,1,1,0; X = 0,1,0,0;\ \bar{X} = 1,0,1,1;\ Y = 0,0,1,0;\ \bar{Y} = 1,1,0,1\}$. Our experiments showed that this approach seems to increase the algebraic degree of the functions involved, but we were not able to find and control all the mentioned cryptographic parameters, yet.

Acknowledgements. The author would like to thank the anonymous referees for their helpful comments, which improved significantly the presentation of the paper.

References

[1] T.W. Cusick, P. Stănică, Bounds on the number of functions satisfying the Strict Avalanche Criterion, *Information Processing Letters* **60**, No. 4 (1996), pp. 215-219.

[2] R. Forré, The Strict Avalanche Criterion: Spectral Properties of Boolean Functions and an Extended Definition, *Advances in Cryptology – Crypto' 88*, LNCS Springer-Verlag, Vol. 403 (1989), pp. 450-468.

[3] B. Preneel, W. Van Leekwijck, L. Van Linden, R. Govaerts, J. Vandewalle, Propagation characteristics of Boolean functions, *Advances in Cryptology – Eurocrypt' 90* LNCS Springer-Verlag, Vol. 473 (1991), pp. 161-173.

[4] J. Seberry, X-M. Zhang, Y. Zheng, Nonlinearly balanced functions and their propagation characteristics, *Advances in Cryptology – Crypto' 93*, LNCS Springer-Verlag, Vol. 773 (1994), pp. 49-60.

[5] J.J. Son, J.I. Lim, S. Chee, S.H. Sung, Global avalanche characteristics and nonlinearity of balanced boolean functions, *Information Processing Letters* **65**, No. 3 (1998), pp. 139-144.

[6] P. Stănică, Chromos, Boolean functions and Avalanche Characteristics, *Ph.D. Thesis*, State University of New York at Buffalo, Buffalo, 1998.

[7] S.H. Sung, S. Chee, C. Park, Global avalanche characteristics and propagation criterion of balanced boolean functions, *Information Processing Letters* **69**, No. 1 (1999), pp. 21-24.

[8] A.F. Webster, S.E. Tavares, On the design of S-boxes, *Advances in Cryptology – Crypto' 85* (1986), LNCS Springer-Verlag, Vol. 218 (1987), pp. 523-534.

[9] A.M. Youssef, T.W. Cusick, P. Stănică, S.E. Tavares, New bounds on the number of functions satisfying the Strict Avalanche Criterion, *Selected Areas in Cryptography' 96*, Kingston-Ontario, Canada, pp. 49-56.

[10] A.M. Youssef, S.E Tavares, Comment on "Bounds on the number of functions satisfying the Strict Avalanche Criterion", *Information Processing Letters* **60**, No. 5 (1997), pp. 271-275.

[11] X-M. Zhang, Y. Zheng, GAC - The criterion for global avalanche characteristics of cryptographic functions, *J. Universal Computer Science* **1**, No. 5 (1995), pp. 320-337.

[12] X-M. Zhang, Y. Zheng, Autocorrelation and new bounds on the nonlinearity of Boolean functions, *Advances in Cryptology - Eurocrypt' 96*, LNCS Springer-Verlag, Vol. 1070 (1996), pp. 294-306.

[13] X-M. Zhang, Y. Zheng, Characterizing the Structures of Cryptographic Functions Satisfying the Propagation Criterion for Almost all Vectors, *Designs, Codes and Cryptography* **7**, No. 1/2 (1996), pp. 111-134.