# On the Guaranteed Number of Activations in XS-circuits *

Sergey Agievich

Research Institute for Applied Problems of Mathematics and Informatics
Belarusian State University
agievich@{bsu.by,gmail.com}

**Abstract**

XS-circuits describe cryptographic primitives that utilize 2 operations on binary words of fixed length: X) bitwise modulo 2 addition and S) substitution. The words are interpreted as elements of a field of characteristic 2. In this paper, we develop a model of XS-circuits according to which several instances of a simple round circuit containing only one S operation are linked together and form a compound circuit called a cascade. S operations of a cascade are interpreted as independent round oracles. When a cascade processes a pair of different inputs, some round oracles get different queries, these oracles are activated. The more activations, the higher security guarantees against differential cryptanalysis the cascade provides. We introduce the notion of the guaranteed number of activations, that is, the minimum number of activations over all choices of the base field, round oracles and pairs of inputs. We show that the guaranteed number of activations is related to the minimum distance of the linear code associated with the cascade. It is also related to the minimum number of occurrences of units in segments of binary linear recurrence sequences whose characteristic polynomial is determined by the round circuit. We provide an algorithm for calculating the guaranteed number of activations. We show how to use the algorithm to deal with linear activations related to linear cryptanalysis.

**Keywords**: circuit, differential cryptanalysis, linear cryptanalysis, linear code, linear recurrence sequence.

## 1 Introduction

XS-circuits describe cryptographic primitives that utilize 2 operations on binary words of fixed length: X) bitwise modulo 2 addition and S) substitution. A circuit may describe a block cipher when instantiating S with key-dependent round functions which usually have a complicated internal structure being circuits of the same (of smaller word length) or other types. Or a circuit may describe an encryption or authentication mode when S is a keyed permutation of a block cipher. One of the directions here is constructing wide-block and variable-input-length ciphers, that is, extending the block length of the underlying cipher to some fixed or even arbitrary length.

We interpret binary words that are processed in an XS-circuit as elements of a field of characteristic 2. A circuit becomes arithmetic if all its operations S are instantiated using only

*Related programs and materials can be found at https://github.com/agievich/xs.

the addition (actually, X) and multiplication in the base field. Arithmetic circuits designed for symmetric cryptography are demanded in universal ZK-proof systems, especially if the circuits have low multiplicative complexity (an example of the approach can be found in [2]). We see another area of application of XS-circuits.

In this paper, we follow the model of XS-circuits proposed in [1]. According to this model, several instances of a simple round circuit containing only one S operation are linked together and form a compound circuit called a cascade. S operations of a cascade are interpreted as independent round oracles. We extensively use notions and notation from [1]. In particular, the notion of regular circuits that are in a sense the best elementary circuits and the only ones worth considering when constructing cascades.

When a cascade processes a pair of different inputs, some round oracles get different queries, these oracles are called *activated*. The more activations, the higher security guarantees against differential cryptanalysis [4] the cascade provides.

In Section 2 we introduce the notion of the guaranteed number of activations, that is, the minimum number of activations over all choices of the base field, round oracles and pairs of inputs. In Section 3 we show that the guaranteed number of activations is related to the minimum distance of the linear code associated with the target cascade. This number is also related to the minimum number of occurrences of units in segments of binary linear recurrence sequences whose characteristic polynomial is determined by the round circuit. This is shown in Section 4. Finally, in Section 5 we provide an algorithm for calculating the guaranteed number of activations. This algorithm can also be used to deal with linear activations related to linear cryptanalysis [12].

Bringing the problem of lower bounding the number of activations to the context of coding theory and showing how to solve it algorithmically, we introduce a systematic approach for constructing sound cryptographic mappings. Interestingly, another systematic approach of this kind, the so-called Wide trail strategy, also relates to coding theory. This approach was proposed in [6, 7] and was implemented in numerous block ciphers including AES and Kuznyechik (see [3] for a fairly complete list).

The Wide trail strategy allows to achieve a high activation rate, close to $1/2$, when MDS (Maximum Distance Separable) codes are used to build a diffusion layer. The drawback of the strategy is that the layer becomes quite complicated and usually has to be implemented through a table lookup. For comparison, the diffusion layer of an XS-circuit can be made very simple. However, S operations of the circuit cannot be applied in parallel although this is allowed by the Wide trail strategy.

# 2 Preliminaries

Let $(a, B, c)$ be a regular XS-circuit of order $n$ (see [1] for definitions and further details). We assume that the circuit is in the first canonical form, that is, $a$ is a nonzero column vector, $B$ is a Frobenius cell, $c = (0, \ldots, 0, 1)$. Denote by $b$ the last column of $B$. All the vectors $a$, $b$, $c$ are binary of dimension $n$.

Instantiating the circuit over a field $\mathbb{F}$ of characteristic 2 and substituting an oracle $S\colon \mathbb{F} \to \mathbb{F}$

for the operation $\mathsf{S}$, we get the mapping

$$(a,B,c)[S]\colon \mathbb{F}^n \to \mathbb{F}^n,\ (x_1,x_2,\ldots,x_n)\mapsto (x_2,x_3,\ldots,x_n,x_{n+1}),$$
$$x_{n+1}=(x_1,x_2,\ldots,x_n)b+S((x_1,x_2,\ldots,x_n)a).$$

Let $(a,B,c)^t$ be the $t$-round cascade built by connecting $t$ instances of $(a,B,c)$. The cascade utilizes $t$ operations $\mathsf{S}$. Instantiating these operations by oracles $S_1,\ldots,S_t$, we obtain the mapping $(a,B,c)^t[S_1,\ldots,S_t]$. It may be described algorithmically as follows: having received an input $(x_1,x_2,\ldots,x_n)$, the sequence

$$x_{\tau+n}=(x_\tau,\ldots,x_{\tau+n-1})b+S_\tau((x_\tau,\ldots,x_{\tau+n-1})a),\quad \tau=1,2,\ldots,t,$$

is calculated and the vector $(x_{t+1},\ldots,x_{t+n})$ is returned as the output.

**Example 1 (GFN1).** The GFN1 family of XS-circuits was introduced in [16]. The circuit of dimension $n\geq 2$ has the second canonical form: $a=(1,0,\ldots,0)^T$, $B$ is a Frobenius cell with $b=a$, $c=a^T$. Replacing $(a,B,c)$ with

$$(B^{-1}a,B^{-1}BB,cB)=((0,\ldots,0,1)^T,B,(0,\ldots,0,1)),$$

we obtain the first canonical form, for which

$$x_{\tau+n}=x_\tau+S_\tau(x_{\tau+n-1}),\quad \tau=1,2,\ldots. \qquad \square$$

Let us suppose now that the cascade processes not one but two inputs simultaneously. From there, $(x_1,x_2,\ldots,x_n)$ is the X-difference of input vectors and $(x_{\tau+1},\ldots,x_{\tau+n})$ is the difference of the $\tau$th round outputs. See [1, Section 4] for further details. There differences are denoted using the symbol $\Delta$ but here we simplify the notation.

The difference $u_\tau$ at the input of $S_\tau$ has the form $(x_\tau,\ldots,x_{\tau+n-1})a$. The corresponding output difference $v_\tau$ can be written as $(x_\tau,\ldots,x_{\tau+n-1})b+x_{\tau+n}$. Due to the bijectivity of $S_\tau$, the equality $u_\tau=0$ holds if and only if $v_\tau=0$. In other words,

$$(x_\tau,\ldots,x_{\tau+n-1})a=0\Leftrightarrow(x_\tau,\ldots,x_{\tau+n-1})b+x_{\tau+n}=0.$$

Let us construct a matrix $G=G(n,a,b,t)$ of dimensions $(t+n)\times 2t$. Its columns go in pairs, the $\tau$th pair has the form:

$$\begin{matrix}0&0\\ \vdots&\vdots\\ 0&0\end{matrix}\Bigg\}\,\tau-1$$
$$\begin{matrix}a&b\\ 0&1\end{matrix}\Big\}\,n+1$$
$$\begin{matrix}0&0\\ \vdots&\vdots\\ 0&0\end{matrix}\Bigg\}\,t-\tau$$

With this,

$$(x_1,x_2,\ldots,x_{t+n})G=(u_1,v_1,u_2,v_2,\ldots,u_t,v_t).$$

We require that in each pair $(u_\tau,v_\tau)$ both elements are either zero or nonzero together. Denote by $\mathcal{W}$ the set of all vectors

$$w=(u_1,v_1,\ldots,u_t,v_t)=xG,\quad x\in\mathbb{F}^{t+n},$$

for which the requirement holds. The set $\mathcal{W}$ is completely determined by the base field $\mathbb{F}$, the vectors $a$, $b$ and the number of rounds $t$. The zero vector obviously belongs to $\mathcal{W}$.

We call the situation when $(u_\tau, v_\tau) \neq (0,0)$ the *activation* of $S_\tau$. Let $\mathrm{wt}_2(w)$ be the total number of activations, that is, nonzero pairs $(u_\tau, v_\tau)$, in the vector $w$.

For $t \geq n$ we are interested in the quantity

$$d(\mathcal{W}) = \min_{w \in \mathcal{W}, w \neq 0} \mathrm{wt}_2(w).$$

It is the minimum number of activations when applying the mappings $(a, B, c)^t[S_1, \ldots, S_t]$ to pairs of different vectors from $\mathbb{F}^n$. Note that the minimization covers all admissible tuples $(S_1, \ldots, S_t)$ and all admissible input pairs.

For $t < n$ we set $d(\mathcal{W}) = 0$. This reflects the fact that as long as the number of rounds is less than the dimension of the circuit, it is possible to avoid activations by manipulating the initial diffirence $(x_1, \ldots, x_n) \neq 0$ (see [1, Section 8]).

The quantity $d(\mathcal{W})$ can also be denoted as $d(\mathbb{F}, n, a, b, t)$ implying that $\mathcal{W}$ is uniquely determined by the parameters $(\mathbb{F}, n, a, b, t)$. Let

$$d(n, a, b, t) = \min_{\mathbb{F}} d(\mathbb{F}, n, a, b, t),$$

where the minimum is taken over all fields of characteristic 2. Any such field is an extension of $\mathbb{F}_2$ and therefore

$$d(n, a, b, t) \leq d(\mathbb{F}, n, a, b, t) \leq d(\mathbb{F}_2, n, a, b, t).$$

The cascade $(a, B, c)^t$ guarantees at least $d(n, a, b, t)$ activations regardless of the choice of $\mathbb{F}$, round oracles and input pairs. We call $d(n, a, b, t)$ the *guaranteed* number of activations.

# 3 Connection to the linear codes

The set $\mathcal{W}$ is a subset of the vector space

$$\mathcal{C} = \{xG \colon x \in \mathbb{F}^{t+n}\} \subseteq \mathbb{F}^{2t}.$$

The following lemma means that for $t \geq n$ the space $\mathcal{C}$ has dimension $t + n$ and, therefore, it is a linear code with the parameters $[2t, t + n]$.

**Lemma 1.** *Let vectors $a$ and $b$ define a regular* $\mathsf{XS}$*-circuit of the first canonical form of dimension $n$. If $t \geq n$, then the matrix $G = G(n, a, b, t)$ has full rank:* $\operatorname{rank} G = t + n$.

*Proof.* Let us associate with the first two columns of $G$ the polynomials $a(\lambda) = \sum_{i=0}^{n-1} a_i \lambda^i$ and $f_B(\lambda) = \lambda^n + \sum_{i=0}^{n-1} b_i \lambda^i$. Here $a_i$ and $b_i$ are coordinates of $a$ and $b$ respectively. We follow the notation introduced in [1, Section 7]. Note that Theorem 9 of the cited paper states that for a regular $\mathsf{XS}$-circuit the polynomials $a(\lambda)$ and $f_B(\lambda)$ are coprime.

The monomial $\lambda^i$ in $a(\lambda)$ marks the position in the first column in which the coefficient $a_i$ is located. The same holds for $f_B(\lambda)$ and the second column. In general, the $\tau$th pair of columns is described by the polynomials $\lambda^{\tau-1} a(\lambda)$ and $\lambda^{\tau-1} f_B(\lambda)$.

The first $2t$ columns of $G$ are linearly dependent if there exist nonzero polynomials $p(\lambda)$ and $q(\lambda)$ whose degrees are less than $t$ and which satisfy

$$p(\lambda) a(\lambda) + q(\lambda) f_B(\lambda) = 0.$$

For $t = n$, since $a(\lambda)$ and $f_B(\lambda)$ are coprime, there are no suitable polynomials $p(\lambda)$, $q(\lambda)$ and the matrix $G$ has full rank.

The first linear dependence appears in $G$ at $t = n + 1$ when choosing $p(\lambda) = f_B(\lambda)$ and $q(\lambda) = a(\lambda)$. The penultimate column becomes dependent on the previous ones. But the last column remains independent, since it is the only one containing 1 in the last row. Thus, $\operatorname{rank} G = 2n + 1 = t + n$ and $G$ is again full-ranked.

The argument can be repeated: each new pair of columns adds 1 to the rank of $G$. Full rank is preserved, which was to be proven. □

The minimum distance of $\mathcal{C}$ is the quantity

$$d(\mathcal{C}) = \min_{w \in \mathcal{C}, w \neq 0} \operatorname{wt}(w),$$

where $\operatorname{wt}(w)$ is the Hamming weight of $w$. According to the Singleton bound (see, for example, [11]),

$$d(\mathcal{C}) \leq 2t + 1 - (t + n) = t - n + 1.$$

Since $\operatorname{wt}(w)/2 \leq \operatorname{wt}_2(w) \leq \operatorname{wt}(w)$ and $\mathcal{W} \subseteq \mathcal{C}$, it holds that

$$d(\mathcal{C})/2 \leq d(\mathcal{W}) \leq d(\mathcal{C}).$$

In particular, $d(\mathcal{W}) \leq t - n + 1$. This estimate means that over $t \geq n$ rounds we cannot guarantee more than $t - n + 1$ activations. Further we are interested in lower bounds for $d(\mathcal{W})$.

Let $t \geq n$ and, therefore, $\operatorname{rank} G = t + n$ by Lemma 1. Suppose that when processing a nonzero input difference using some round oracles, activations occur only in rounds whose numbers belong to a set $\mathcal{T} \subseteq \{1, 2, \ldots, t\}$. We call $\mathcal{T}$ the *activation profile*. Following this profile, let us divide $G$ into two parts: $G_0$ and $G_1$. The matrix $G_1$ consists of pairs of columns whose numbers are in $\mathcal{T}$ and $G_0$ consists of the remaining columns. By construction, there exists a nonzero vector $x \in \mathbb{F}^{t+n}$ such that $xG_0 = 0$ and $xG_1$ does not contain zeros. This means that the partition $(G_0, G_1)$ is feasible in the sense of the following definition.

**Definition.** Let $G_0$ and $G_1$ be matrices composed of different pairs of columns of $G$. The partition $(G_0, G_1)$ is *feasible* if

1) $\operatorname{rank} G_0 < t + n$;

2) $\operatorname{rank}(G_0 \mid g) > \operatorname{rank} G_0$ for each column $g$ of $G_1$.

Indeed, if $\operatorname{rank} G_0 = t + n$, then from $xG_0 = 0$ it follows that $x = 0$ which contradicts the construction. And if $\operatorname{rank}(G_0 \mid g) = \operatorname{rank} G_0$, then from $xG_0 = 0$ it follows that $xg = 0$. The latter means that $xG_1$ contains zero, again a contradiction.

In the following lemma, we show that feasibility of a partition $(G_0, G_1)$ is not only necessary but also a sufficient condition for the feasibility of the underlying activation profile.

**Lemma 2.** *Let vectors $a$ and $b$ define a regular* XS*-circuit of the first canonical form of dimension $n$. Let $t \geq n$ and $k$ be the maximum number of pairs of columns in the matrix $G_0$ where the maximum is taken over all feasible partitions $(G_0, G_1)$ of $G = G(n, a, b, t)$. Then*

$$d(n, a, b, t) = t - k.$$

*Proof.* Let $(G_0, G_1)$ be a feasible partition of $G$. It is necessary to prove that there exists an extension $\mathbb{F}$ of the field $\mathbb{F}_2$ and a vector $x \in \mathbb{F}^{t+n}$ such that $xG_0 = 0$ and $xG_1$ does not contain zeros.

The set
$$L = \{xG_1 \colon x \in \mathbb{F}^{t+n}, xG_0 = 0\} \subseteq \mathbb{F}^{2(t-k)}$$
is a vector space of dimension $r = t + n - \operatorname{rank} G_0$. It can be written as
$$L = \{yP \colon y \in \mathbb{F}^r\},$$
where $P$ is a binary matrix of dimensions $r \times 2(t-k)$. The matrix $P$ does not contain a zero column due to the second restriction on the feasibility of the partition $(G_0, G_1)$.

Suppose that $L$ does not contain a vector without zero coordinates. Then we choose an arbitrary vector $yP \in L$, build an extension $\mathbb{F}'$ of the field $\mathbb{F}$ and extend $y$ to a vector $y'$ of the same dimension but over $\mathbb{F}'$. We construct $y'$ in such way that a particular zero coordinate of $yP$ becomes nonzero in $y'P$ while nonzero coordinates of $yP$ remain nonzero in $y'P$. After constructing the pair $(\mathbb{F}', y')$ we interpret it as $(\mathbb{F}, y)$ and repeat the extension until we get the vector $yP$ without zeros. It remains to show how to extend $y$ to $y'$.

Define $\mathbb{F}'$ as an extension of $\mathbb{F}$ of degree 2. Without loss of generality, let elements of $\mathbb{F}'$ be $(m+1)$-bit words $\alpha = \alpha_1 \ldots \alpha_m \alpha_{m+1}$ and $\alpha \in \mathbb{F}$ if and only if $\alpha_{m+1} = 0$. Let the addition in $\mathbb{F}'$ be the usual XOR. The extension of $y$ consists in setting the last (zero) bits of its coordinates. Let $\beta$ be a vector composed of these bits. Since $P$ does not contain zero columns, it is possible to choose $\beta$ so that a particular coordinate of $\beta P$ is nonzero. The corresponding coordinate of $y'P$ is also nonzero. Moreover, if a certain coordinate $yP$ is nonzero, then the corresponding coordinate $y'P$ remains nonzero. That was to be proven. □

**Remark 1.** *The minimum distance of the code* $\mathcal{C} = \{xG\}$ *can also be defined as* $d(\mathcal{C}) = t - k$, *where* $k$ *is the maximum number of columns in* $G_0$ *and the maximum is taken over all feasible partitions* $(G_0, G_1)$ *of* $G$ *(see, for example, [9, Theorem 1.4.5]). The difference is in changing the partitioning restrictions. Now* $G_i$ *not necessarily consists of pairs of related columns, the requirement* $\operatorname{rank} G_0 < t + n$ *is preserved, but the requirement* $\operatorname{rank}(G_0 \mid g) > \operatorname{rank} G_0$ *becomes redundant.*

**Remark 2.** *Let* $\operatorname{rank} G_0 = t+n-1$. *Then in the proof above, the vector space* $L$ *has dimension* 1 *and the matrix* $P$ *becomes the row vector* $(1, 1, \ldots, 1)$. *This means that with* $\mathbb{F} = \mathbb{F}_2$ *there exists a nonzero* $x \in \mathbb{F}^{t+n}$ *such that* $xG_0 = 0$ *and* $xG_1 = (1, 1, \ldots, 1)$. *In other words, the activation profile associated with the partition* $(G_0, G_1)$ *is feasible over* $\mathbb{F}_2$. *Moreover, as we see below, this profile is a segment of a linear recurrence sequence over* $\mathbb{F}_2$.

# 4 The case $\mathbb{F} = \mathbb{F}_2$

In the case $\mathbb{F} = \mathbb{F}_2$, the condition $u_\tau = 0 \Leftrightarrow v_\tau = 0$ is eqiuvalent to $u_\tau = v_\tau$. With this,
$$x_{\tau+n} = (x_\tau, \ldots, x_{\tau+n-1})(a+b), \quad \tau = 1, 2, \ldots, t,$$
that is, the sequence $(x_1, \ldots, x_{t+n})$ is a segment of a nonzero linear recurrence sequence (LRS) over $\mathbb{F}_2$. The characteristic polynomial of the sequence is
$$f(\lambda) = \lambda^n + f_{n-1}\lambda^{n-1} + \ldots + f_1\lambda + f_0, \quad (f_0, f_1, \ldots, f_{n-1}) = a + b.$$

The vectors $(x_\tau, \ldots, x_{\tau+n-1})$, $\tau = 1, 2, \ldots$, stand as states of the linear feedback shift register (LFSR) associated with $f(\lambda)$. When choosing the first bits of LFSR states, we get the sequence $(x_\tau)$, and when choosing the linear combinations $(x_\tau, \ldots, x_{\tau+n-1})a$, we get the sequence $(u_\tau)$. The latter sequence is also a LRS with the same characteristic polynomial $f(\lambda)$.

The sequence $(u_\tau)$ is nonzero. Indeed, the underlying XS-circuit is regular and for any nonzero input difference $(x_1, \ldots, x_n)$ at least one activation must occur over the first $n$ rounds (see the discussion before Theorem 10 in [1]) which means that $(u_1, \ldots, u_n) \neq (0, \ldots, 0)$.

For the same reason, $f_0 = 1$ and the sequences $(x_\tau)$, $(u_\tau)$ are purely periodic. Indeed, otherwise, a nonzero input difference $(x_1, \ldots, x_n) = (1, 0, \ldots, 0)$ induces a zero difference after $n - 1$ rounds, which is impossible due to the regularity.

The number of activations over $t$ rounds is the number of nonzero elements (units) in the segment $(u_1, \ldots, u_t)$. We can use known results on the number of occurrences of particular elements in segments of LRS. Let $r$ be the least period of $(u_\tau)$, $R$ be the order of $f$ (the maximum least period of nonzero LRS with the characteristic polynomial $f$). Then according to Theorems 8.82 and 8.85 from [10], the number of activation is at least

$$\frac{t}{2} - 2^{n/2-1}\left(\frac{r}{R}\right)^{1/2}\left(t_0 + \frac{2}{\pi}\log r + \frac{2}{5} + \frac{t_1}{r}\right).$$

Here $t_0$ and $t_1$ are respectively the quotient and remainder when dividing $t$ by $r$. If $t_1 = 0$, then only the term $t_0$ can be left in the last brackets.

It makes sense to apply the estimate above only for large $n$, $t$ and $r$. In practice, these parameters are small and the minimum number of activations can be found by exhaustive search over all LRS profiles $(u_1, \ldots, u_t)$ in time of order $2^n t$.

**Example 2 (SM4).** The SM4 circuit is used in the block cipher of the same name (formerly known as SMS4). See [8] for details of the cipher and [1] for details of the circuit.

The circuit is already in the first canonical form, its dimension is 4, the characteristic polynomial $f(\lambda) = \lambda^4 + \lambda^3 + \lambda^2 + \lambda + 1$. The polynomial $f(\lambda)$ is irreducible of order 5. Therefore, the least period of $(u_\tau)$ equals 5.

The minimum number of activations is achieved on the start segments of the following LRS:

$$0, 0, 0, 1, 1,\ 0, 0, 0, 1, 1,\ \ldots$$

If $t = 5t_0 + t_1$, $0 \leq t_1 < 5$, then this number is

$$\begin{cases} 2t_0, & t_1 = 0, 1, 2, 3, \\ 2t_0 + 1, & t_1 = 4. \end{cases}$$

□

**Example 3 (GFN1, continued).** Let us continue Example 1 and consider the GFN1 circuit of dimension $n$ in the first canonical form. For this circuit, $a = (0, 0, \ldots, 0, 1)^T$, $b = (1, 0, \ldots, 0, 0)^T$ and $f(\lambda) = \lambda^n + \lambda^{n-1} + 1$.

For $n = 2, 3, 4$, the polynomial $f(\lambda)$ is primitive. The least period of $(u_\tau)$ equals $r = 2^n - 1$ and every full period $(u_1, \ldots, u_r)$ contains exactly $2^{n-1}$ units. Therefore,

$$d(\mathbb{F}_2, n, a, b, 2^n - 1) = 2^{n-1}$$

and the activation rate over $2^n - 1$ rounds can potentially achieve the value $2^{n-1}/(2^n - 1) > 1/2$. This value is indeed achieved for $n = 2, 3$ but, as we show later, not for $n = 4$. □

# 5 The algorithm

The following algorithm summarizes our constructions and reasoning.

**Algorithm** GNA (THE GUARANTEED NUMBER OF ACTIVATIONS)

**Input**: $(n, a, b, t)$, where $a$ and $b$ are binary vectors of dimension $n$ that define a regular XS-circuit in the first canonical form, $t$ is a number of rounds.

**Output**: $d(n, a, b, t)$, the guaranteed number of activations over $t$ rounds of the input circuit.

**Steps**:

1. If $t < n$, then return 0. If $t = n$, return 1.
2. Construct the matrix $G = G(n, a, b, t)$ as explained in Section 3. The dimensions of $G$ are $(t + n) \times 2t$, $\operatorname{rank} G = t + n$. The columns of $G$ are grouped in pairs.
3. Calculate $d(\mathbb{F}_2, n, a, b, t)$ as described in Section 4 and set $k \leftarrow t - d(\mathbb{F}_2, n, a, b, t)$.
4. Make a list of all possible partitions of $G$ into submatrices $G_0$ and $G_1$ such that $G_0$ contains exactly $k + 1$ pairs of columns of $G$.
5. For each partition $(G_0, G_1)$:
   (a) if $\operatorname{rank} G_0 \geq t + n - 1$, then continue (go to the end of the loop);
   (b) if there is a column $g$ in $G_1$ such that $\operatorname{rank}(G_0 \mid g) = \operatorname{rank} G_0$, then continue;
   (c) set $k \leftarrow k + 1$ and go to Step 4.
6. Return $t - k$.

**Theorem.** *The algorithm GNA is correct.*

*Proof.* A direct consequence of Lemma 2 and Remark 2.

In Step 5a of the algorithm, we skip the case $\operatorname{rank} G_0 = t + n - 1$ because in this case the activation profile associated with the partition $(G_0, G_1)$ is feasible over $\mathbb{F}_2$ and the initial bound $d(\mathbb{F}_2, n, a, b, t)$ for $d(n, a, b, t)$ cannot be strengthened. □

Let us discuss the complexity of the algorithm. Step 3 runs in time of order $2^n t$. Then, for each $k = t - d(\mathbb{F}_2, n, a, b, t), \ldots, t - d(n, a, b, t)$, GNA processes $\binom{t}{k+1}$ partitions $(G_0, G_1)$ using linear algebra on submatrices of $G$. The total number of partitions is exponential in $t$. Thus, GNA is exponential in both $t$ and $n$ and can only be used for small and moderate input dimensions. Fortunately, these are the dimensions that are interesting in practice. Moreover, in many cases (one of which is discussed in Example 4), iterating over partitions can be significantly simplified.

The algorithm GNA gives us the guaranteed number of *differential* activations. We can easily adapt the algorithm to deal with *linear* activations (see [1, Section 9]). To do this, we pass from $(a, B, c)$ to the dual circuit $(c^T, B^T, a^T)$ and determine vectors $a'$ and $b'$ that define its first canonical form (we only need to determine $a'$, since $b' = b$). The quantity $\mathrm{GNA}(n, a', b', t)$ is the guaranteed number of linear activations.

**Example 4 (SM4, continued).** For SM4 its dual has the same first canonical form. So the guaranteed numbers of differential and linear activations are the same. The outputs of GNA against SM4 for $t \leq 32$ coincide with the estimates of Example 2. Thus, the activation rate close to 2/5 is achieved. In particular, the guaranteed number of activations over 32 rounds (exactly the case of the block cipher SM4) is 12.

Note that here we are processing the abstract SM4 circuit, not its instantiation in SM4. In this instantiation, S operations are constructed using round keys, table $S$-boxes, rotations and XORs of binary words. Lower bounds on the number of active $S$-*boxes* (not activations / active *rounds*) in SM4 can be found in [13, 14, 15].

Iterating over $\binom{t}{k+1}$ partitions in Step 4 of GNA can be simplified. For example, in the case of SM4, if any 4 of 5 consecutive pairs of columns fall into $G_0$, then the corresponding partition is not feasible and can be immediately rejected. Indeed, the 5 consecutive pairs of columns are linearly dependent while 4 pairs are not (it follows from the same reasoning as in the proof of Lemma 1). Therefore, a pair not included in $G_0$ contains a column $g$ which is linearly expressed through the columns of $G_0$ and, therefore, the second condition of feasibility is violated. □

**Example 5 (GFN1, continued).** An GFN1 circuit of arbitrary dimension is self-dual: $(a, B, c) = (c^T, B^T, a^T)$. Therefore, a bound on differential activations is also a bound on linear activations.

For the circuit of dimension $n = 4$, GNA gives 7 activations over 15 rounds. It is one less than estimated in Example 3 through LRS profiles. The optimal activation profile found by GNA looks as follows:

000111101**100100**.

It differs from the related LRS profile

000111101**011001**

starting from the 10th round. The LRS profile gives 3 activations after the fork while the optimal profile gives only 2 activations. □

**Example 6 (activation times).** The $i$th activation time, $\rho_i$, is the minimum number of rounds that guarantees $i$ activations (see [1, Section 8]). In the next table, we present the values $\rho_i$ for GFN1 of dimension $n = 4$ and for SM4. We calculate $\rho_i$ using the GNA algorithm.

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\rho_i$(GFN1) | 4 | 7 | 8 | 10 | 12 | 13 | 14 | 17 | 20 | 22 | 23 | 25 |
| $\rho_i$(SM4) | 4 | 5 | 9 | 10 | 14 | 15 | 19 | 20 | 24 | 25 | 29 | 30 |

The time $\rho_7$(GFN1) $= 14$ given in the table refines Proposition 5 of [5]. □

**Acknowledgments**: The author thanks Egor Lawrenov for fruitful discussions and computer experiments, the results of which are used in the above examples. The author also thanks the anonymous referees of CTCrypt'2020 and MVK for their valuable feedback.

# References


[1] S. Agievich. XS-circuits in block ciphers. *Mat. Vopr. Kriptogr.* **10** (2) (2019), pp. 7–30. URL: https://doi.org/10.4213/mvk281.

[2] M. Albrecht, L. Grassi, C. Rechberger, A. Roy, and T. Tiessen. MiMC: efficient encryption and cryptographic hashing with minimal multiplicative complexity. In: *Advances in Cryptology – ASIACRYPT 2016*. Ed. by J. Cheon and T. Takagi. Vol. 10031. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 191–219.

[3] R. Avanzi. *A Salad of Block Ciphers*. Cryptology ePrint Archive, Report 2016/1171. https://eprint.iacr.org/2016/1171. 2016.

[4] E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. In: *Advances in Cryptology — CRYPTO '90*. Ed. by A. Menezes and S. Vanstone. Vol. 537. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1991, pp. 2–21.

[5] A. Bogdanov and K. Shibutani. Generalized Feistel networks revisited. *Des. Codes Cryptogr.* **66** (2013), pp. 75–97. URL: https://doi.org/10.1007/s10623-012-9660-z.

[6] J. Daemen. Cipher and Hash Function Design Strategies Based on Linear and Differential Cryptanalysis. Doctoral Dissertation. K.U. Leuven, 1995. URL: https://cs.ru.nl/~joan/papers/JDA_Thesis_1995.pdf.

[7] J. Daemen and V. Rijmen. The wide trail design strategy. In: *Cryptography and Coding 2001*. Ed. by B. Honary. Vol. 2260. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2001, pp. 222–238.

[8] W. Diffie and G. Ledin. *SMS4 Encryption Algorithm for Wireless Networks*. Cryptology ePrint Archive, Report 2008/329. 2008. URL: https://eprint.iacr.org/2008/329.

[9] W. C. Huffman and V. Pless. *Fundamentals of error-correcting codes*. Cambridge: Cambridge Univ. Press, 2003. URL: https://cds.cern.ch/record/1139892.

[10] R. Lidl and H. Niederreiter. *Finite Fields*. New York, NY, USA: Cambridge University Press, 1997.

[11] F.J. MacWilliams and N.J.A. Sloane. *The Theory of Error-Correcting Codes*. Mathematical Studies. Elsevier Science, 1977.

[12] M. Matsui. Linear cryptanalysis method for DES cipher. In: *Advances in Cryptology — EUROCRYPT '93*. Ed. by T. Helleseth. Vol. 765. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1994, pp. 386–397.

[13] S. Wu and M. Wang. *Security Evaluation against Differential Cryptanalysis for Block Cipher Structures*. Cryptology ePrint Archive, Report 2011/551. https://eprint.iacr.org/2011/551. 2011.

[14] J. Zhang, W. Wu, and Y. Zheng. Security of SM4 against (related-key) differential cryptanalysis. In: *Information Security Practice and Experience. ISPEC 2016*. Ed. by F. Bao, L. Chen, R. Deng, and G. Wang. Vol. 10060. Lecture Notes in Computer Science. Cham: Springer, 2016, pp. 65–78.

[15] M. Zhang, J. Liu, and X. Wang. *The upper bounds on differential characteristics in block cipher SMS4*. Cryptology ePrint Archive, Report 2010/155. https://eprint.iacr.org/2010/155. 2010.

[16] Y. Zheng, T. Matsumoto, and H. Imai. On the construction of block ciphers provably secure and not relying on any unproved hypotheses. In: *Advances in Cryptology – CRYPTO'89 Proceedings*. Ed. by G. Brassard. Vol. 435. Lecture Notes in Computer Science. New York, NY: Springer, 1990, pp. 461–480.