THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

*PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,*
*Volume 21, Number 1/2020, pp. 69–76*

# EAGLESONG: AN ARX HASH WITH FAST DIFFUSION

Alan SZEPIENIEC[1], Tomer ASHUR[2,3]

[1] Nervos Foundation
[2] Cryptomeria
[3] imec-COSIC, KU Leuven
Corresponding author: Alan SZEPIENIEC, E-mail: alan@nervos.org

**Abstract**. We propose a hash function based on three design principles: the sponge construction, ARX operations, and the wide trail strategy. While the sponge construction applies generically to any sufficiently strong permutation, the wide trail strategy and the ARX operations are naturally somewhat incompatible. We show that while the ARX operations provide only very weakly nonlinear S-boxes, it is possible to build very strong linear diffusion layers with them. As a result, the wide trail argument, which bounds the attacker's success probability in terms of the minimum number of active S-boxes across two rounds, survives. The proposed hash function is one of a very select group of ARX ciphers featuring rigorous bounds against differential and linear cryptanalysis.

***Key words***: hash function, addition-rotate-xor, sponge, wide trail strategy, diffusion.

## 1. INTRODUCTION

In the design of symmetric ciphers, the *wide trail strategy* is a common technique offering a strong resilience against differential and linear cryptanalysis [6] and lies at the base of many popular and widely adopted ciphers such as the AES [5]. The technique entails alternating between a confusion layer, in which many highly nonlinear but localized functional blocks operate in parallel on a large state; and a diffusion layer, in which the localized effects of the previous layer are spread out across the entire state. The combined effect of both layers results in an upper bound on the differential and linear probabilities, thereby enabling the designer to estimate the number of rounds required to reach a target security level against the matching attacks.

An alternate design strategy increasingly popular particularly in the context of software-oriented ciphers is the minimalist reliance on only three instructions: Modular Addition, Cyclic Rotation, and Exclusive-or (`xor`), or ARX for short. While these operations do not generally offer the strong non-linearity required by the relatively hard-to-compute functional blocks in the wide trail strategy, ARX ciphers can afford to compensate with a larger number of rounds: compared to the highly non-linear functional blocks of the wide trail strategy, the addition, rotation, and `xor` functions are relatively fast.

The ARX and wide trail design strategies seem to be fundamentally in opposition to each other. Popular ARX ciphers like Salsa [3], Blake [1], and Speck [2] drop rigorous bounds on the differential and linear probabilities altogether in favor of heuristic arguments. A notable exception to this list but still in support of the opposition of strategies, is the SPARX and LAX family of ciphers [8]. The designers of these families introduce an alternative to the wide trail strategy called the *long trail strategy*, which advocates using large and expensive functional blocks derived from many simple operations (such as ARX operations), coupled with a cheaper and weaker diffusion layer. The long trail strategy admits provable bounds on the linear and differential success probabilities.

*Our contribution*. We present Eaglesong, a hash function whose design successfully unifies the ARX and wide trail strategies. The key to this fusion is the algebraic interpretation of the addition, rotation, and `xor` operations. This algebraic perspective in turn enables the construction of a strong diffusion layer, a rigorous analysis of its properties, and the derivation of bounds on the linear and differential probabilities in accordance with which the number of rounds is set.

## 2. SPECIFICATION

Eaglesong is a sponge function. The innovation is in the construction of the permutation $F$. We revisit the standard construction of a hash function from a sponge function in Section 2.1 for the sake of a self-contained presentation. For other modes of operation derived from the sponge construction we refer the reader to the exposition by Bertoni et al. [4].

The Eaglesong permutation $F$ operates on a state of 16 words of 32 bits by applying the round function $N=43$ times for a 128 bit security level. The round function consists of four steps:

1. *Bit matrix*, with which words are mixed across the entire state. This diffusion is accomplished using an invertible $16\times16$ matrix over $\mathbb{F}_2$. The action of this matrix is realized with only the `xor` operation.
2. *Circulant multiplication*, in which bits are mixed within each word. This operation is realized with `xor`s and rotations, compatible with multiplication-by-constant in the quotient ring $\mathbb{F}_2[x]/\langle x^{32}+1\rangle$.
3. *Injection of constants*, in which a predetermined list of random constants are `xor`ed into the state. The constants are different each round.
4. *Nonlinear map*, in which nonlinear operations are applied to the state elements. Here we use integer addition modulo $2^{32}$, which is nonlinear with respect to vector spaces over $\mathbb{F}_2$. Modular addition is applied twice, before and after a rotation of the words by 8 bits in opposite directions.

A diagram overview of the round function is presented in Fig. 1.

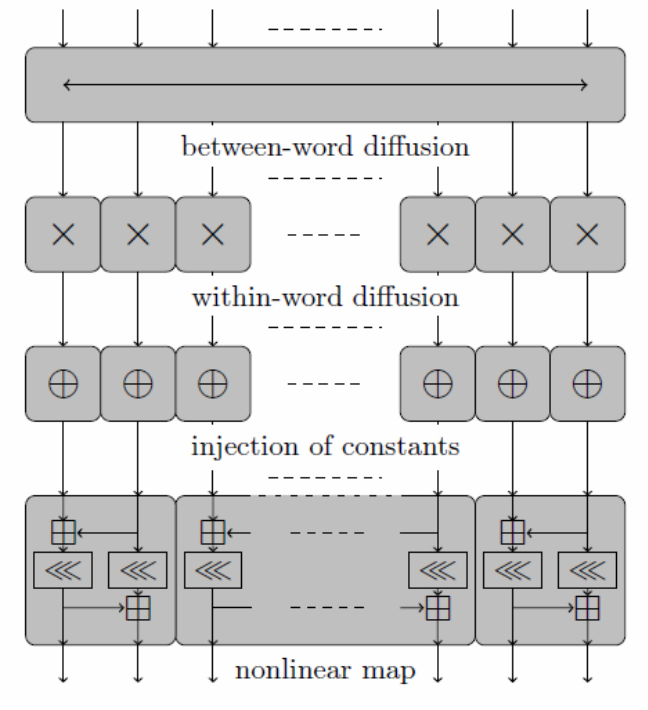


Fig. 1 – Eaglesong round function $F$.

**Bit matrix.** The between-word diffusion layer applies a linear operation to the vector of state elements. Specifically, the operation on the state vector e is given by a matrix $\mathbf{M}\in\mathbb{F}_2^{16\times16}$ via $\mathbf{e}^{\mathrm{T}}\mapsto\mathbf{e}^{\mathrm{T}}\mathbf{M}$. As this matrix consists of known elements in the coefficient ring $\mathbb{F}_2$, the multiplications involved in this linear transformation correspond to simply `xor`ing or ignoring the elements of the state vector depending on whether the matching element of the matrix is a 1 or a 0.

The matrix $\mathbf{M}$ is determined via the binary BCH code with $m=5$, $\delta=6$, $n=31$, extension field $\mathbb{E}=\mathbb{F}_2[x]/\langle x^5+x^2+1\rangle$, and generator polynomial

$$g(X)=X^{15}+X^{11}+X^{10}+X^9+X^8+X^7+X^5+X^3+X^2+X+1.$$

As a linear code, this code is defined by a generator matrix $\mathbf{G} \in \mathbb{F}_2^{16\times 31}$. Without loss of generality we may consider the echelon form of $\mathbf{G}$, or in the lingo of coding theory, its systematic form: $\mathbf{G}_s = \left(\mathbf{I} \mid \bar{\mathbf{M}}\right)$, where $\bar{\mathbf{M}} \in \mathbb{F}_2^{16\times 15}$. Then we find $\mathbf{M}$ by adjoining to $\bar{\mathbf{M}}$ the unique column vector $\mathbf{k} \in \mathbb{F}_2^{16}$ satisfying $\mathbf{k}^{\mathrm{T}}\bar{\mathbf{M}} = 0$. In particular, $\mathbf{M}$ is given by:

$$\mathbf{M} = \left(\begin{array}{cccccccccccccccc|c}
1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1\\
0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1\\
0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1\\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1\\
1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0\\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1\\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0\\
1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1\\
0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1\\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1\\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1\\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1\\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0\\
0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0\\
0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0\\
1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1
\end{array}\right) \tag{1}$$

An exhaustive search in the primal code $\left\{\left(\mathbf{x}^{\mathrm{T}}, \mathbf{x}^{\mathrm{T}}\mathbf{M}\right) \mid \mathbf{x} \in \mathbb{F}_2^{16}\right\}$ and the dual code $\left\{\left(\mathbf{x}^{\mathrm{T}}, \mathbf{x}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\right) \mid \mathbf{x} \in \mathbb{F}_2^{16}\right\}$ shows that neither have nonzero codewords of Hamming weight less than 8. Moreover, the minimum weight of codewords in the primal and dual code is 5 with respect to the following alternative "pairwise Hamming weight" metric: partition the codeword into adjacent pairs of bits, and then count the number of pairs different from 00.

**Circulant multiplication.** The within-word diffusion interprets each word as an element in the ring $\mathbb{F}_2[x]/\langle x^{32}+1\rangle$ and maps it to its product by a trinomial from the same ring. The known coefficient of this multiplication is given by $1 + x^{a_i} + x^{b_i}$ where all $a_i \neq 0$ and $b_i \notin \{0, a_i\}$ are determined from SHAKE256 seeded with the ASCII string *"I thought of calling it 'information', but the word was overly used, so I decided to call it 'uncertainty'. When I discussed it with John von Neumann, he had a better idea"*. We use rejection sampling to ensure that no coefficient is its own inverse, and that the Hamming weight of any product of coefficients is at least 7.

Trinomials are invertible under multiplication in the ring $\mathbb{F}_2[x]/\langle x^{32}+1\rangle$, meaning that this operation maps uniform inputs to uniform outputs. Moreover, given that multiplication by $x^{a_i}$ represents rotation by $a_i$ positions, the multiplication can be computed with two rotations (denoted by $\lll$) and two `xors` (denoted by $\oplus$):

$$e_i \mapsto e_i \times \left(1 + x^{a_i} + x^{b_i}\right) = e_i \oplus (e_i \lll a_i) \oplus (e_i \lll b_i).$$

This operation has a characterization in the language of coding theory: the set $\left\{\left(e_i, e_i \times (1 + x^{a_i} + x^{b_i})\right) \mid e_i \in \mathbb{F}_2[x]/\langle x^{32}+1\rangle\right\}$ represents a quasi-cyclic code of length $n = 64$ and dimension $k = 32$. Its minimal distance can be verified by way of brute force to be at least four.

**Injection of constants.** In this layer, the constant $c_i$ is xored into state element $e_i$. The $c_i$ are determined from SHAKE256 seeded with the ASCII string "I have always been on the machines' side". They are provided in the full version of the paper, which is available from the authors upon request.

**Nonlinear map.** The nonlinear map consists of three steps: odd-to-even addition, followed by rotation, followed by even-to-odd addition[1]. Specifically and in sequence:

$$e_i \mapsto \begin{cases} e_i \boxplus e_{i+1} & \text{if } i \text{ even} \\ e_i & \text{otherwise;} \end{cases} \qquad e_i \mapsto \begin{cases} e_i \lll 8 & \text{if } i \text{ even} \\ e_i \lll 24 & \text{otherwise;} \end{cases} \qquad e_i \mapsto \begin{cases} e_i \boxplus e_{i-1} & \text{if } i \text{ odd} \\ e_i & \text{otherwise.} \end{cases}$$

Here $\boxplus$ represents addition modulo $2^{32}$. These operations are the only component of the Eaglesong round function that are nonlinear for $\mathbb{F}_2[x]/\langle x^{32}+1\rangle$.

This nonlinear map departs from traditional substitution-permutation networks, in which the substitution layer is typically a bricklayer function of S-boxes, applying to each word without mixing them with others. In contrast, the present map mixes pairs of words. Alternatively, the Addition-Rotation-Addition block can be viewed as an S-box applying to pairs of words rather than individual ones.

### 2.2. The hash function

The sequence of four steps described above constitutes a single round. Repeating this round $N = 43$ times computes the Eaglesong permutation. Using this permutation in a sponge construction is what gives the hash function.

The sponge construction consists of two phases, absorbing and squeezing, and is associated with two parameters, the *rate* $r$ and the *capacity* $c$ such that the full state consists of $r+c$ bits. First, the state is initialized to the all zero string. Then a delimiter is appended to the input before the entire input is cut into chunks of $r$ bits, with the last chunk being possibly smaller. Every chunk is xored into the state, after which the permutation is applied. This describes the absorbing phase. In each iteration of the squeezing phase, $r$ bits are read out from the state before the permutation is applied. To obtain a hash function one truncates the output to whatever output length is specified. For Eaglesong, we set $c = r = 256$ bits and $0\times06$ as the delimiter byte.

Out of space considerations, we omit pseudocode for the permutation and sponge construction. We refer the interested reader to the full version of this paper for a standalone specification of the complete hash function.

## 3. SECURITY ANALYSIS

We assess the security of Eaglesong (as a block cipher in the PRP game) from the perspective of two branches of statistical cryptanalysis: differential cryptanalysis, and linear cryptanalysis; and provide bounds on the distinguishing advantage for one query – or conversely, on the requisite number of queries for a reasonably successful distinguisher – as a function of the number of rounds $N$. Of these branches, linear cryptanalysis yields the weakest bound, and the number of rounds is set accordingly. For the sake of brevity we omit an exhaustive list of alternative statistical cryptanalyses and rely on the overwhelmingly accurate right-hand rule that differential and linear cryptanalysis are the best performers anyway.

### 3.1. Differential cryptanalysis

Differential cryptanalysis studies the propagation of differences through various stages of the cipher. Specifically, let $F:\{0,1\}^n \to \{0,1\}^n$ be a function, then for fixed *input difference* $\Delta x \in \{0,1\}^n$ and *output difference* $\Delta y \in \{0,1\}^n$ we define the *differential probability* as

[1] Here and elsewhere, indexation starts, as it should, from zero.

$$\mathrm{DP}_F(\Delta x,\Delta y)\overset{\Delta}{=}\Pr_X\left[F(X)\oplus F(X\oplus\Delta x)=\Delta y\right]. \tag{2}$$

When the cipher is parameterized by an unknown key $K$, we prefer the *expected differential probability* instead:

$$\mathrm{EDP}_F(\Delta x,\Delta y)\overset{\Delta}{=}\mathop{\mathrm{E}}_K\left[\Pr_X\left[F(X;K)\oplus F(X\oplus\Delta x;K)=\Delta y\right]\right]. \tag{3}$$

The utility of these notions stems from their capacity to be outliers. If the attacker knows a suitable differential pair $(\Delta x,\Delta y)$ for which $\mathrm{EDP}_F(\Delta x,\Delta y)$ is much larger than it would be if $F$ were random (see [8] for the expected behaviour of random permutations), then by computing this value from a large enough number of plaintext and ciphertext pairs, the attacker can determine which function he is interfacing with – the cipher, or a random permutation. From the designer's standpoint, it is imperative to make the EDP indistinguishable from that of a random permutation, for all differential pairs. From a security analysis perspective, a common method is to bound the *maximum expected differential probability (MEDP)*:

$$\mathrm{MEDP}_F\overset{\Delta}{=}\max_{\Delta x,\Delta y\in\{0,1\}^n\setminus\{0\}}\mathrm{EDP}_F(\Delta x,\Delta y). \tag{4}$$

If the function $F$ decomposes as a sequence of stages, i.e., $F=f_0\circ f_1\circ\cdots\circ f_{r-1}$, then one can consider a vector of differences called a *differential characteristic* $\boldsymbol{\Delta x}=\left(\Delta x_0,\Delta x_1,\ldots,\Delta x_r\right)$ and the associated *(expected) differential characteristic probability ((E)DCP)*:

$$\mathrm{DCP}_F(\boldsymbol{\Delta x})\overset{\Delta}{=}\prod_{i=0}^{r-1}\mathrm{DP}_{f_i}(\Delta x_i,\Delta x_{i+1}), \tag{5}$$

$$\mathrm{EDCP}_F(\boldsymbol{\Delta x})\overset{\Delta}{=}\mathop{E}_K\left[\prod_{i=0}^{r-1}\Pr_X\left[f_i(X;K)\oplus f_i(X\oplus\Delta x_i;K)=\Delta x_{i+1}\right]\right]. \tag{6}$$

We hope to bound the maximum of this value, the *maximum expected differential characteristic probability (MEDCP)*:

$$\mathrm{MEDCP}_F\overset{\Delta}{=}\max_{\boldsymbol{\Delta x}\in(\{0,1\}^n)^{r+1}}\mathrm{EDCP}_F(\boldsymbol{\Delta x}). \tag{7}$$

We say a bit in a differential characteristic is *active* if it is set, and components in the cipher's circuit are *active* for the characteristic if it has at least one active input bit.

We apply two heuristics. First, we ignore the positive reinforcement of characteristic probabilities due to different characteristics that have the same first and last value, i.e. $\boldsymbol{\Delta x}\neq\boldsymbol{\Delta y}$ but $\Delta x_0=\Delta y_0$ and $\Delta x_{r-1}=\Delta y_{r-1}$. Second, we assume that the differential probabilities are independent in each stage. Although these assumptions are strictly speaking false, they are common and have proved to be "good enough" over the years.

We decompose the Eaglesong permutation $F$ as a sequence of $N$ rounds $f_i$, $i\in\{0,\ldots,N-1\}$. With the above heuristics we can then identify $\mathrm{MEDCP}_F$ with the $N$th power of $\mathrm{MEDCP}_{f_i}$. Furthermore, after decomposing a single round function into its four layers, one observes that the bit matrix, circulant multiplication, and injection of constants (and key), do not decrease the EDP because they are all affine. All EDP decrease must therefore come from the nonlinear map, and more specifically from the modular additions contained therein.

Modular addition takes two outputs $X$ and $Y$, and produces one output, $Z=X+Y \bmod 2^{32}$. With respect to nonzero input differences, we distinguish two cases. First, $\Delta x,\Delta y\in\{0,2^{31}\}$ giving rise to a $\Delta z\in\{0,2^{31}\}$ with probability 1. Second, and more interestingly, $\Delta x$ or $\Delta y$ or both take values from outside the set $\{0,2^{31}\}$, and whatever value the output difference $\Delta z$ takes, it takes it with probability at most one half.

The Addition-Rotation-Addition (ARA) block is constructed such that there is always at least one addition incurring an MEDP degradation of 0.5 if the block is active. In fact, the event of 0.5 MEDP degradation is rare. We count only four difference characteristics with this probability:

$$\left(2^{31},0\right) \xrightarrow[\mathrm{DP}=1]{\boxplus_1} \left(2^{31},0\right) \xrightarrow[\mathrm{DP}=1]{\lll} \left(2^{7},0\right) \xrightarrow[\mathrm{DP}=0.5]{\boxplus_2} \left(2^{7},2^{7}\right) \tag{8}$$

$$\left(2^{31},2^{31}\right) \xrightarrow[\mathrm{DP}=1]{\boxplus_1} \left(0,2^{31}\right) \xrightarrow[\mathrm{DP}=1]{\lll} \left(0,2^{23}\right) \xrightarrow[\mathrm{DP}=0.5]{\boxplus_2} \left(0,2^{23}\right) \tag{9}$$

$$\left(2^{23},0\right) \xrightarrow[\mathrm{DP}=0.5]{\boxplus_1} \left(2^{23},0\right) \xrightarrow[\mathrm{DP}=1]{\lll} \left(2^{31},0\right) \xrightarrow[\mathrm{DP}=1]{\boxplus_2} \left(2^{31},2^{31}\right) \tag{10}$$

$$\left(2^{7},2^{7}\right) \xrightarrow[\mathrm{DP}=0.5]{\boxplus_1} \left(0,2^{7}\right) \xrightarrow[\mathrm{DP}=1]{\lll} \left(0,2^{31}\right) \xrightarrow[\mathrm{DP}=1]{\boxplus_2} \left(0,2^{31}\right) \tag{11}$$

Aside from the difference $(0,0)$, which propagates with probability 1, any other characteristic occurs with probability at most 0.25.

The distinction between differential probability 0.5, and 0.25 or less, motivates a distinction between weakly active (0.5), and strongly active (0.25 or less), Addition-Rotation-Addition blocks. In particular, it enables an argument lower-bounding the minimum number of strongly active Addition-Rotation-Addition blocks across any two rounds. We know that the bitmatrix induces a code of minimum distance 5 under the pairwise Hamming weight metric. An exhaustive computer enumeration of all possible output patterns of 1, 2, 3, or 4 weakly active Addition-Rotation-Addition blocks (and no strongly active ones), indicates that in all cases 5 or more blocks are strongly active in the next layer. The same observation holds in reverse. The worst possible scenario not captured by our enumeration is still five active blocks across two rounds, but with at least two strongly active ones. Consequently, the EDCP of any characteristic spanning two rounds is at most $0.5^3 \cdot 0.25^2 = 2^{-7}$. Finally, the MEDCP of the full $N$ rounds is this number raised to the power $N/2$:

$$\mathrm{MEDCP}_F \leq 2^{-7N/2}. \tag{12}$$

### 3.2. Linear cryptanalysis

Where differential cryptanalysis studies the propagation of differences in pairs of inputs and outputs, linear cryptanalysis studies the propagation of masks and the inputs and outputs that they satisfy. Specifically, a mask $a,b \in \{0,1\}^n$ on an input-output pair $(X,Y)$ with $X \in \{0,1\}^n$ and $Y = F(X)$ for some function $F$ is *satisfying* if the sum of inner products is zero, i.e., $a \cdot X \oplus b \cdot Y = 0$. For a function $F$, we associate the *linear potential* (LP) to this event; the expectation of that potential, in the case of a keyed primitive; and the maximum of that expectation:

$$\mathrm{LP}_F(a,b) \stackrel{\Delta}{=} \left(2 \cdot \Pr_X\left[a \cdot X \oplus b \cdot F(X) = 0\right] - 1\right)^2; \tag{13}$$

$$\mathrm{ELP}_F(a,b) \stackrel{\Delta}{=} \mathop{\mathrm{E}}_K\left[\mathrm{LP}_{F(\cdot;K)}(a,b)\right]; \tag{14}$$

$$\mathrm{MELP}_F \stackrel{\Delta}{=} \max_{a,b \in \{0,1\}^n \setminus \{0\}} \mathrm{ELP}_F(a,b). \tag{15}$$

Instead of differential characteristics, we have *linear trails* that analogously represent sequences of masks applied to the inputs and outputs of a function that decomposes into various stages $F = f_0 \circ \cdots \circ f_{r-1}$. We associate the *((maximum) expected) linear trail potential (((M)E)LTP)* accordingly.

$$\mathrm{LTP}_F(a,b) \stackrel{\Delta}{=} \prod_{i=0}^{r-1} \mathrm{LP}_{f_i}(a_i,b_i); \tag{16}$$

$$\mathrm{ELTP}_F(a,b) \stackrel{\Delta}{=} \mathop{\mathrm{E}}_{K}\left[\prod_{i=0}^{r-1}\mathrm{LP}_{f_i(\cdot;K)}(a_i,b_i)^2\right]; \tag{17}$$

$$\mathrm{MELTP}_F \stackrel{\Delta}{=} \max_{a,b}\mathrm{ELTP}_F(a,b). \tag{18}$$

Like their differential counterparts, the utility of these notions stems from their capacity to be outliers. An attacker who computes a sufficiently close approximation to the actual linear probability can distinguish the cipher from a random permutation. The task of the designer is then to push the linear potential sufficiently close to zero so that it cannot be detected using the amount of data given.

Also like in differential cryptanalysis, the affine operations do not affect the ELP; all ELP decrease must come from the modular additions in the nonlinear layer. The key challenge is therefore to determine $\mathrm{MELP}_{\mathrm{ARA}}$.

First, we determine the MELP of modular addition, where $(a,b)$ is the mask applied to the input and $c$ is applied to the output. Observe that the mask $a=b=c=1$ is always satisfied, reflecting the fact that the least significant bit of the output is always the exclusive-or of the least significant bits of the inputs. In all other cases, $\mathrm{LP}_{\mathrm{ModAdd}}\left((a,b),c\right)\le\frac{1}{4}$.

For the ARA block, we observe four trails that are satisfied with potential 1, associated with all combinations of masks (0,0,0) or (1,1,1) for the two adders. The first and last elements of these trails should be considered inactive masks. The trails are as follows, with elements ordered according to the picture.

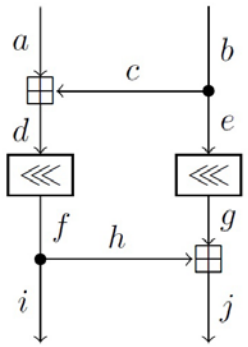


Fig. 2 – Trails elements ordering.

| ( $a$, | $b$, | $c$, | $d$, | $e$, | $f$, | $g$, | $h$, | $i$, | $j$ ) |
|---|---|---|---|---|---|---|---|---|---|
| ( 0, | 0, | 0, | 0, | 0, | 0, | 0, | 0, | 0, | 0 ) |
| ( 1, | 1, | 1, | 1, | 0, | 256, | 0, | 0, | 256, | 0 ) |
| ( 0, | 256, | 0, | 0, | 256, | 0, | 1, | 1, | 1, | 1 ) |
| ( 1, | 257, | 1, | 1, | 256, | 256, | 1, | 1, | 257, | 1 ) |

(19)

All other trails activate at least one adder and are hence satisfied with potential at most $2^{-2}$. This observation motivates an exhaustive enumeration of all masks for the full state in the input to the nonlinear layer in round $i+1$, in order to compute the matching activity pattern at the output of the nonlinear layer in round $i$, as well as the other way around. Our computer enumeration indicates that in all nonzero cases at least 3 pairs of words are active in the other round. Supplementing this exhaustive enumeration with a symbolic treatment of one active pair of words (with all the others being inactive) shows that on the other side of the linear layers at least two pairs of words are active. These enumerations establish that 3 is a lower-bound on the minimum number of active S-boxes in any trail spanning two rounds, and so

$$\mathrm{MELTP}_F \le \mathrm{MELP}_{f_i\circ f_{i+1}}^{N/2} \le \mathrm{MELP}_{ModAdd}^{3\times N/2} = 2^{-3N}. \tag{20}$$

## REFERENCES

1. J. AUMASSON, W. MEIER, R.C. PHAN, L. HENZEN, *The hash function BLAKE*, Information Security and Cryptography Book Series (ISC), Springer, 2014.
2. R. BEAULIEU, D. SHORS, J. SMITH, S. TREATMAN-CLARK, B. WEEKS, L. WINGERS*, The SIMON and SPECK families of lightweight block ciphers,* IACR Cryptology ePrint Archive, **1**, pp. 404-449, 2013.
3. D.J. BERNSTEIN, *The Salsa20 family of stream ciphers*, New Stream Cipher Designs – The eSTREAM Finalists, pp. 84-97, 2008.
4. G. BERTONI, J. DAEMEN, M. PEETERS, G. VAN ASSCHE, *Cryptographic sponge functions,* SHA-3 competition (round 3), 2011, available online: https://keccak.team/files/CSF-0.1.pdf.
5. J. DAEMEN, V. RIJMEN, *Rijndael for AES*, AES Candidate Conference, 2000, pp. 343–348.
6. J. DAEMEN, V. RIJMEN, The Design of Rijndael: AES – The Advanced Encryption Standard, Springer, 2002.
7. J. DAEMEN, V. RIJMEN, *Probability distributions of correlation and differentials in block ciphers*, J. Mathematical Cryptology **1**, *3*, pp. 221-242, 2007.
8. D. DINU, L. PERRIN, A. UDOVENKO, V. VELICHKOV, J. GROßSCHÄDL, A. BIRYUKOV, *Design strategies for ARX with provable bounds: Sparx and LAX*, in: "ASIACRYPT 2016 Part I" (eds. J.H. Cheon, T. Takagi), Lecture Notes in Computer Science, **10031**, pp. 484–513, 2016, Springer.

*Received June 27, 2019*