CRYPTOGRAPHIC PROPERTIES OF ADDITION MODULO 2^n

S. M. DEHNAVI*, A. MAHMOODI RISHAKANI, M. R. MIRZAEE SHAMSABAD, HAMIDREZA MAIMANI, EINOLLAH PASHA

Abstract. The operation of modular addition modulo a power of two is one of the most applied operations in symmetric cryptography. For example, modular addition is used in RC6, MARS and Twofish block ciphers and RC4, Bluetooth and Rabbit stream ciphers. In this paper, we study statistical and algebraic properties of modular addition modulo a power of two. We obtain the probability distribution of modular addition carry bits along with conditional probability distribution of these carry bits. Using these probability distributions and Markovity of modular addition carry bits, we compute the joint probability distribution of arbitrary number of modular addition carry bits. Then, we examine algebraic properties of modular addition with a constant and obtain the number of terms as well as the algebraic degrees of the component Boolean functions of the output of modular addition with a constant. Finally, we present another formula for the ANF of component Boolean functions of modular addition modulo a power of two. This formula contains more information than representations which are presented in cryptographic literature, up to now.

Keywords: Modular addition, Boolean function, Component Boolean function, Carry bit, Algebraic degree

* std_dehnavism@khu.ac.ir

1. **Introduction**

The operation of modular addition modulo a power of two is one of the most applied operations in symmetric cryptography. For instance, modular addition is used in RC6[8], MARS[4] and Twofish[9] block ciphers and RC4[7], Bluetooth[1] and Rabbit[2] stream ciphers. In this paper, we investigate statistical and algebraic properties of modular addition modulo a power of two. Firstly, we obtain the probability distribution of modular addition carry bits and conditional probability distribution of these carry bits. Then, using these probability distributions and Markovity of modular addition carry bits, we compute the joint probability distribution of arbitrary number of modular addition carry bits.

Algebraic properties of modular addition modulo a power of two is studied in [3]. We examine algebraic properties of modular addition with a constant and obtain the number of terms and algebraic degrees of the component Boolean functions of the output of modular addition with a constant.

Finally, we present another formula for the ANF of component Boolean functions of modular addition modulo a power of two. This formula contains more information than previous representations which have been presented in cryptographic literature, up to now.

Section 2 presents preliminary notations and definitions. In Section 3 we investigate the probability distribution of modular addition carry bits. Section 4 studies algebraic properties of modular addition with a constant and Section 5 is the conclusion.

2. **Preliminaries**

In this paper, the number of elements (cardinal) of a finite set A is denoted by $|A|$. Hamming weight of a natural number or a binary vector x is denoted by $\mathbf{w}(x)$. For two numbers or binary vectors x, y, $\mathbf{d}(x, y)$ denotes their Hamming distance. The i-th bit of a natural number or a binary vector x is denoted by x_i. The notation $\wedge$ is used for AND operation, $\ggg$ is used for cyclic right shift or rotation operation and $\oplus$ is used for XOR operation.

Let $\mathbb{Z}_{2^n}$ be the ring of integers modulo 2^n. For each $a \in \mathbb{Z}_{2^n}$, the unique integer $\bar{a} \in \mathbb{Z}_{2^n}$ with $\bar{a} + a = 2^n - 1$ is called the complement of a.

Let $\mathbb{F}_2$ be the field with two elements. There is a one-to-one correspondence between $\mathbb{Z}_{2^n}$, the ring of integers modulo 2^n and $\mathbb{F}_2^n$, Cartesian product of n copies of $\mathbb{F}_2$:

$$\varphi : \mathbb{F}_2^n \to \mathbb{Z}_{2^n},$$

$$x = (x_{n-1}, ..., x_0) \mapsto \varphi(x) = \sum_{i=0}^{n-1} x_i 2^i.$$

Now let $\prec$ be the partial order on $\mathbb{F}_2^n$ defined as

$$x \prec a \Leftrightarrow x_i \leq a_i, \quad 0 \leq i < n.$$

For $x, u \in \mathbb{F}_2^n$, we define $x^u = x_{n-1}^{u_{n-1}} ... x_0^{u_0}$ where x_i and u_i, $0 \leq i < n$, are the i-th components or bits of x and u, respectively.

Every function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is called a Boolean function. Any Boolean function has a unique algebraic representation called its Algebraic Normal Form or ANF[5]. In fact,

$$f(x) = \bigoplus_{u \in \mathbb{Z}_{2^n}} h_u x^u, \quad h_u \in \mathbb{F}_2.$$

where,

$$h_u = \bigoplus_{x \prec u} f(x).$$

Algebraic degree of a Boolean function f is denoted by $d(f)$ and is equal to the number of variables in the longest term in the ANF of f, or equivalently, it is equal to the greatest value of $\mathbf{w}(u)$ among all terms with $h_u \neq 0$. The number of terms in the ANF of the Boolean function f is denoted by $n(f)$.

Every function $f : \mathbb{F}_2^m \to \mathbb{F}_2^n$ with $n > 1$ is called a vectorial Boolean function or a Boolean map. Such a function can be represented as a vector of Boolean functions $(f_{n-1}, ..., f_0)$, where each f_i, $0 \leq i < n$, is a Boolean function $f_i : \mathbb{F}_2^m \to \mathbb{F}_2$ and is called the i-th component Boolean function. Moreover, similar to the case of Boolean functions, any vectorial Boolean function $f : \mathbb{F}_2^m \to \mathbb{F}_2^n$ has a unique representation called its vectorial ANF. In fact,

$$f(x) = \bigoplus_{u \in \mathbb{Z}_{2^m}} h_u x^u, \quad h_u \in \mathbb{F}_2^n.$$

The inner product of two vectors $a, b \in \mathbb{F}_2^m$ in $\mathbb{F}_2^m$ is defined by

$$a.b = \bigoplus_{i=0}^{m-1} a_i b_i,$$

and the inner product of these vectors in the set of real numbers is defined by

$$a \circ b = \sum_{i=0}^{m-1} a_i b_i.$$

For any nonzero element $a \in \mathbb{Z}_{2^n}$, the greatest power of 2 that divides a is denoted by $p(a)$. We define $p(0) := n$. For two functions $f, g : \mathbb{F}_2^m \to \mathbb{F}_2^n$, we define $P_{f,g}$ as,

$$P_{f,g} = \frac{|\{x \in \mathbb{F}_2^m | f(x) = g(x)\}|}{2^m}.$$

and $E_{f,g}$ as,

$$E_{f,g} = n - \frac{\sum_{x \in \mathbb{F}_2^m} \mathbf{d}(f(x), g(x))}{2^m} = \frac{\sum_{i=0}^{n-1} |\{x \in \mathbb{F}_2^m | f_i(x) = g_i(x)\}|}{2^m}.$$

where f_i's and g_i's, $0 \leq i < n$, are the i-th component Boolean functions of f and g.

3. Probability Distribution of Carry Bits

There is a well-known recursive formula for the carry bits of modular addition modulo a power of two: Define

$$f : \mathbb{F}_2^{2n} \to \mathbb{F}_2^n,$$

$$z = f(x, y) = x + y \pmod{2^n}.$$

Then we have,

$$\begin{aligned} z_i &= x_i \oplus y_i \oplus c_i, \quad 0 \leq i < n, \\ c_0 &= 0, \\ c_i &= x_{i-1} y_{i-1} \oplus c_{i-1}(x_{i-1} \oplus y_{i-1}), \quad 1 \leq i < n. \end{aligned} \tag{3.1}$$

We use this formula in the proof of the following theorems.

Theorem 3.1. *Let $x, y \in \mathbb{Z}_{2^n}$. We have*

$$x + y = (x \oplus y) + 2(x \wedge y) \pmod{2^n}.$$

Proof. Suppose that $z = x + y \pmod{2^n}$, $a = x \oplus y$, $b = 2(x \wedge y)$ and $c = a + b \pmod{2^n}$. From (3.1), we have

$$z_i = x_i \oplus y_i \oplus d_i, \quad 0 \le i < n,$$
$$d_0 = 0,$$
$$d_i = x_{i-1}y_{i-1} \oplus d_{i-1}(x_{i-1} \oplus y_{i-1}), \quad 1 \le i < n.$$

Also,

$$b_0 = 0,$$
$$b_i = x_{i-1}y_{i-1}, \quad 1 \le i < n,$$
$$a_i = x_i \oplus y_i, \quad 0 \le i < n;$$

and,

$$c_i = a_i \oplus b_i \oplus e_i, \quad 0 \le i < n,$$
$$e_0 = 0,$$
$$e_i = a_{i-1}b_{i-1} \oplus e_{i-1}(a_{i-1} \oplus b_{i-1}), \quad 1 \le i < n.$$

At first we prove:

$$e_i = d_i \oplus x_{i-1}y_{i-1}, \quad 1 \le i < n. \tag{3.2}$$

We use induction on i. For $i = 1$, we have

$$d_1 = x_0y_0 \oplus d_0(x_0 \oplus y_0) = x_0y_0,$$
$$e_1 = a_0b_0 \oplus e_0(a_0 \oplus b_0) = a_0b_0 = 0.$$

So, $e_1 = d_1 \oplus x_0y_0$. Now assume that (3.2) is valid for $i - 1$. Then,

$$\begin{aligned}
&e_i \oplus x_{i-1}y_{i-1} = a_{i-1}b_{i-1} \oplus e_{i-1}(a_{i-1} \oplus b_{i-1}) \oplus x_{i-1}y_{i-1} \\
&= (x_{i-1} \oplus y_{i-1})x_{i-2}y_{i-2} \oplus (d_{i-1} \oplus x_{i-2}y_{i-2})(x_{i-1} \oplus y_{i-1} \oplus x_{i-2}y_{i-2}) \oplus \\
&x_{i-1}y_{i-1} \\
&= (x_{i-1} \oplus y_{i-1})x_{i-2}y_{i-2} \oplus d_{i-1}(x_{i-1} \oplus y_{i-1}) \oplus d_{i-1}x_{i-2}y_{i-2} \oplus x_{i-2}y_{i-2}(x_{i-1} \oplus \\
&y_{i-1}) \oplus x_{i-2}y_{i-2} \oplus x_{i-1}y_{i-1} \\
&= (x_{i-1}y_{i-1} \oplus d_{i-1}(x_{i-1} \oplus y_{i-1})) \oplus d_{i-1}x_{i-2}y_{i-2} \oplus x_{i-2}y_{i-2} \\
&= d_i \oplus d_{i-1}x_{i-2}y_{i-2} \oplus x_{i-2}y_{i-2} = d_i.
\end{aligned}$$

The last equation is obtained by multiplying two sides of the equation

$$d_{i-1} = x_{i-2}y_{i-2} \oplus d_{i-2}(x_{i-2} \oplus y_{i-2})$$

by $x_{i-2}y_{i-2}$. In fact, we have

$$d_{i-1}x_{i-2}y_{i-2} = x_{i-2}y_{i-2} \oplus d_{i-2}x_{i-2}y_{i-2} \oplus d_{i-2}x_{i-2}y_{i-2} = x_{i-2}y_{i-2}.$$

This proves validity of (3.2). Now we are ready to prove $c = z$:

$$c_0 = a_0 \oplus b_0 \oplus e_0 = a_0 \oplus b_0 = x_0 \oplus y_0 = x_0 \oplus y_0 \oplus d_0 = z_0,$$

and for $i > 0$,

$$c_i = a_i \oplus b_i \oplus e_i = (x_i \oplus y_i) \oplus x_{i-1}y_{i-1} \oplus d_i \oplus x_{i-1}y_{i-1} = x_i \oplus y_i \oplus d_i = z_i.$$

□

As a result of Theorem 3.1, we can find $P_{f,g}$, where f is the map of modular addition modulo 2^n and g is the bitwise XOR map: for f and g to be equal, it suffices to have $2(x \wedge y) = 0$. By a simple combinatorial enumeration, we get

$$P_{f,g} = \left(\frac{3}{4}\right)^{n-1}. \tag{3.3}$$

In other words, (3.3) gives an approximation of modular addition with bitwise XOR. There is a well-known fact about the distribution of carry bits of modular addition:

Theorem 3.2. *If $x, y \in \mathbb{Z}_{2^n}$ and $z = x + y \pmod{2^n}$, then,*

$$P\left(c_i = 0\right) = \frac{1}{2} + \frac{1}{2^{i+1}}, \quad 0 \leq i < n.$$

Here, according to (3.1), c_i, $0 \leq i < n$, is the i-th carry bit of modular addition.

Regarding the maps f and g in (3.3) and using Theorem 3.2, we have

$$E_{f,g} = \sum_{i=0}^{n-1} \left(\frac{1}{2} + \frac{1}{2^{i+1}}\right) = \frac{n}{2} + \frac{2^n - 1}{2^n}.$$

Theorem 3.3. *Regarding the notations of Theorem 3.2, if $(a_{n-1}, ..., a_0) \in \mathbb{F}_2^n$ is given, then,*

$$P(c_{n-1} = a_{n-1}, ..., c_0 = a_0) = \begin{cases} \left(\frac{3}{4}\right)^{n-1} 3^{-\mathbf{w}(b)} & a_0 = 0, \\ 0 & a_0 \neq 0. \end{cases}$$

Here, c_i's, $0 \leq i < n$, are modular addition carry bits and the vector b is

$$b = (b_{n-1}, ..., b_0) = (a_{n-1} \oplus a_{n-2}, ..., a_1 \oplus a_0, 0).$$

Proof. In the case $a_0 = 1$, there is nothing to prove, because $c_0 = 0$. If $a_0 \neq 1$, it is not hard to see that

$$P(c_i = 0 | c_{i-1} = 0) = P(c_i = 1 | c_{i-1} = 1) = \frac{3}{4},$$

and,

$$P(c_i = 0|c_{i-1} = 1) = P(c_i = 1|c_{i-1} = 0) = \frac{1}{4}.$$

Relation (3.1) and a simple calculation show that the sequence or stochastic process $\{c_i\}_{i\geq 0}$ is a Markov chain. More precisely, we have

$$P(c_i = a_i|c_{i-1} = a_{i-1}, ..., c_0 = a_0) = P(c_i = a_i|c_{i-1} = a_{i-1}).$$

Therefore,
$P(c_{n-1} = a_{n-1}, ..., c_0 = 0) = P(c_0 = 0)P(c_1 = a_1|c_0 = 0) \cdots P(c_{n-1} = a_{n-1}|c_{n-2} = a_{n-2}) = \left(\frac{3}{4}\right)^{d_1} \left(\frac{1}{4}\right)^{d_2} = \frac{3^{d_1}}{4^{d_1+d_2}}$,
where,

$$d_1 = |\{j|0 \leq j < n-1, a_j = a_{j+1}\}|,$$
$$d_2 = |\{j|0 \leq j < n-1, a_j \neq a_{j+1}\}|.$$

It is clear that $d_1 = n - \mathbf{w}(b) - 1$ and $d_1 + d_2 = n - 1$. This ends the proof. $\square$

We note that (3.3) can also be derived from Theorem 3.3.

Theorem 3.4. *According to the notations of Theorem 3.3, for every $i \geq 2$ and $1 \leq j < i$, and any two bits a, b, we have*

$$P(c_i = a|c_{i-j} = b) = \frac{1}{2} + \frac{(-1)^{a+b}}{2^{j+1}}. \tag{3.4}$$

Proof. The case $i = 2$ for every $1 \leq j < i$ or $j = 1$ has been proved in Theorem 3.3. By induction, suppose that (3.4) is valid for i and every $1 \leq j < i$. We will prove the validity of (3.4) for $i + 1$ and every $1 \leq j < i + 1$:

$$P(c_{i+1} = a|c_{i+1-j} = b) = \frac{P(c_{i+1}=a, c_{i+1-j}=b)}{P(c_{i+1-j}=b)}$$

$$= \frac{P(c_{i+1}=a, c_i=0, c_{i+1-j}=b)}{P(c_{i+1-j}=b)} + \frac{P(c_{i+1}=a, c_i=1, c_{i+1-j}=b)}{P(c_{i+1-j}=b)}$$

$$= \frac{P(c_{i+1-j}=b)P(c_i=0|c_{i+1-j}=b)P(c_{i+1}=a|c_i=0, c_{i+1-j}=b)}{P(c_{i+1-j}=b)}$$

$$+ \frac{P(c_{i+1-j}=b)P(c_i=1|c_{i+1-j}=b)P(c_{i+1}=a|c_i=1, c_{i+1-j}=b)}{P(c_{i+1-j}=b)}$$

$$= P(c_i = 0|c_{i+1-j} = b)P(c_{i+1} = a|c_i = 0) + P(c_i = 1|c_{i+1-j} = b)P(c_{i+1} = a|c_i = 1)$$

$$= \left(\frac{1}{2} + \frac{(-1)^b}{2^j}\right)\left(\frac{1}{2} + \frac{(-1)^a}{2^2}\right) + \left(\frac{1}{2} + \frac{(-1)^{b+1}}{2^j}\right)\left(\frac{1}{2} + \frac{(-1)^{a+1}}{2^2}\right) = \frac{1}{2} + \frac{(-1)^{a+b}}{2^{j+1}}.$$

$\square$

Theorem 3.5. *For every $2 \leq r \leq n$ and $a \in \mathbb{F}_2^r$ with $a = (a_{r-1}, ..., a_0)$, and every $(j_{r-1}, ..., j_0)$ satisfying $0 < j_0 < ... < j_{r-1} < n$, we have*

$$P\left(c_{j_{r-1}} = a_{r-1}, ..., c_{j_0} = a_0\right) = \prod_{k=0}^{r-1} \left(\frac{1}{2} + \frac{(-1)^{a_k + a_{k-1}}}{2^{j_k - j_{k-1} + 1}}\right).$$

Here $a_{-1} = j_{-1} := 0$.

Proof. The proof is the same as Theorem 3.4. Using equations (3.1) directly or equivalently using Markovity of $\{c_i\}_{i \geq 0}$, we have

$$P\left(c_{j_{r-1}} = a_{r-1}, ..., c_{j_0} = a_0\right)$$
$$= P\left(c_{j_0} = a_0\right) \prod_{k=1}^{r-1} P\left(c_{j_k} = a_k | c_{j_{k-1}} = a_{k-1}, ..., c_{j_0} = a_0\right)$$
$$= P\left(c_{j_0} = a_0\right) \prod_{k=1}^{r-1} P\left(c_{j_k} = a_k | c_{j_{k-1}} = a_{k-1}\right) = \prod_{k=0}^{r-1} \left(\frac{1}{2} + \frac{(-1)^{a_k + a_{k-1}}}{2^{j_k - j_{k-1} + 1}}\right).$$

□

4. Algebraic Properties of Modular Addition with a Constant

In this section, we study algebraic properties of modular addition with a constant. We obtain the algebraic degree and the number of terms in the ANF of the component Boolean functions of modular addition with a constant. The proof of the following theorem can be found in [3].

Theorem 4.1. *Let*

$$f : \mathbb{F}_2^{2n} \rightarrow \mathbb{F}_2^n,$$
$$f(x, y) = x + y \pmod{2^n}.$$

Then for each i, $0 \leq i < n$, we have

$$d(f_i) = i + 1,$$
$$n(f_i) = 2^i + 1.$$

Theorem 4.2. *Let $a \in \mathbb{Z}_{2^n}$ be given and,*

$$f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n,$$
$$z = f(x) = x + a \pmod{2^n}.$$

Then for each i, $0 \le i < n$,

$$n(f_i) = \begin{cases} 1 & i < p(a), \\ 2 & i = p(a), \\ 2^{w(a,i)} + a_i & i > p(a). \end{cases}$$

Here, we define $w(a,i) := \mathbf{w}(0,...,0,a_{i-1},...,a_0)$.

Proof. Suppose that $p(a) = k$. If $i < k$, then according to (3.1), we have $a_i = c_i = 0$. So $z_i = x_i$ and $n(f_i) = 1$. If $i = k$, then by (3.1), we have $a_i = 1$ and $c_i = 0$; so $z_i = x_i \oplus 1$ and $n(f_i) = 2$.

Now if $i > k$, then according to (3.1), we have $z_i = x_i \oplus a_i \oplus c_i$ and $n(f_i) = 1 + a_i + n(c_i)$. By induction on i, we prove that for any $i > k$, $n(c_i) = 2^{w(a,i)} - 1$: for $i = k+1$ we have

$$c_{k+1} = x_k, \quad w(a,k+1) = \mathbf{w}(0,...,0,1,0,...,0) = 1.$$

So, $n(c_{k+1}) = 1 = 2^{w(a,k+1)} - 1$. Suppose that for $i > k+1$, $n(c_i) = 2^{w(a,i)} - 1$. According to (3.1), for $i+1$ we have

$$c_{i+1} = x_i a_i \oplus c_i(x_i \oplus a_i).$$

Now if $a_i = 0$, then $w(a,i+1) = w(a,i)$ and $n(c_{i+1}) = n(c_i)$. So,

$$n(c_{i+1}) = 2^{w(a,i+1)} - 1,$$

and if $a_i = 1$, then $w(a,i+1) = w(a,i) + 1$ and $n(c_{i+1}) = 2n(c_i) + 1$. Thus,

$$n(c_{i+1}) = 2\left(2^{w(a,i)} - 1\right) + 1 = 2^{w(a,i+1)} - 1.$$

Threfore, for all $i > k$, $n(c_i) = 2^{w(a,i)} - 1$ and $n(f_i) = 2^{w(a,i)} + a_i$. $\square$

Theorem 4.2 shows that, if a constant a has more ones in its binary representation, then z_{n-1} would have more terms in its ANF.

Theorem 4.3. *With the notations of Theorem 4.2 we have*

$$d(f_i) = \begin{cases} 1 & i \le p(a) + 1, \\ i - p(a) & i > p(a) + 1. \end{cases}$$

Proof. Using equation (3.1), we have

$$d(f_i) = \begin{cases} 1 & d(c_i) \le 1, \\ d(c_i) & d(c_i) > 1. \end{cases}$$

Now, suppose that $p(a) = k$. Then,

$$a_0 = ... = a_{k-1} = 0, \quad a_k = 1,$$

and, if we replace the above values in equation (3.1), then we have

$$c_0 = ... = c_k = 0, \quad c_{k+1} = x_k.$$

So we have $d(c_i) \leq 1$ and $d(f_i) = 1$, for any $i \leq k+1 = p(a)+1$.
Now we prove by induction that for any $i > p(a)+1$, $d(z_i) = i - p(a)$: according to (3.1), for $i = p(a)+2 = k+2$, we have

$$c_i = c_{k+2} = x_{k+1}a_{k+1} \oplus c_{k+1}(x_{k+1} \oplus a_{k+1}),$$

and,

$$d(c_{k+2}) = 2.$$

Thus $d(f_{k+2}) = 2 = i - p(a)$. Suppose that for $i > k+2$, $d(z_i) = i - p(a)$. Then, for $i+1$ we have

$$d(c_{i+1}) = d(c_i) + 1 = i - k + 1 > 2.$$

Therefore, $d(f_{i+1}) = i + 1 - p(a)$ and this ends the proof. □

Example. Suppose that $z = x + 13 \pmod{2^8}$. We have $p(13) = 0$. So by Theorems 4.2 and 4.3, we have

$$n(z_0) = 2,$$

$$n(z_1) = n(z_2) = 3,$$

$$n(z_3) = 5,$$

$$n(z_4) = n(_5) = n(z_6) = n(z_7) = 7.$$

and,

$$d(z_0) = 1,$$

$$d(z_i) = i, \quad 1 \leq i \leq 7.$$

In the following theorems, we shall study other forms of the ANF of modular addition. The proof of following theorem can be found in [3].

Theorem 4.4. *Let*

$$f : \mathbb{F}_2^{2n} \to \mathbb{F}_2^n,$$

$$z = f(x, y) = x + y \pmod{2^n}.$$

For any i, $0 \leq i \leq n-1$, we have

$$z_i = \bigoplus_{t=0}^{2^i} x^t y^{2^i - t}.$$

Theorem 4.5. *Let*

$$f : \mathbb{F}_2^{2n} \to \mathbb{F}_2^n,$$
$$f(x,y) = x + y \pmod{2^n}.$$

Then any monomial appears exactly once in vectorial ANF of f.

Proof. According to Theorem 4.4, for all $0 \le i < n$, we have

$$z_i = \bigoplus_{t=0}^{2^i} x^t y^{2^i - t}.$$

For all $0 \le i < j \le n-1$, the monomial $x^u y^{2^i-u}$ with $0 \le u \le 2^i$ and the monomial $x^v y^{2^j-v}$ with $0 \le v \le 2^j$ can not be equal, because for $u \ne v$, they are different, obviously, and for $u = v$ we have $2^i - u \ne 2^j - v$. ☐

The following results are derived from Theorem 4.5.

Corollary 4.6. *Suppose that*

$$f : \mathbb{F}_2^{2n} \to \mathbb{F}_2^n,$$
$$f(x,y) = x + y \pmod{2^n}.$$

Then the number of all terms in vectorial ANF of f equals to

$$\sum_{i=0}^{n-1} \left(2^i + 1\right) = 2^n + n - 1.$$

Corollary 4.7. *Suppose that $z = x + y \pmod{2^n}$ and $a \in \mathbb{Z}_{2^n}$. Then the number of terms in Boolean function $a.z$ equals to*

$$a \circ \left(2^{n-1} + 1, ..., 3, 2\right).$$

Corollary 4.7. shows that, if we have a linear combination of the output of modular addition, how we can obtain the number of terms in the component Boolean functions of this linear combination.

Corollary 4.8. *Suppose that*

$$f : \mathbb{F}_2^{2n} \to \mathbb{F}_2^n,$$
$$f(x,y) = x + y \pmod{2^n}.$$

Then this function is sparse in relation to a random vectorial Boolean function $g : \mathbb{F}_2^{2n} \to \mathbb{F}_2^n$, i.e.

$$\lim_{n\to\infty} \frac{2^n + n - 1}{n 2^{2n-1}} = 0.$$

Here, $n2^{2n-1}$ is the expected number of monomials in the ANF of a Boolean function with domain $\mathbb{F}_2^{2n}$.

Example. Suppose that $x, y \in \mathbb{Z}_{2^8}$ and $w = x + y \pmod{2^8}$. If

$$z = w \oplus (w \ggg 3) \oplus (w \ggg 5),$$

then we have

$$n(z_0) = (2^0 + 1) + (2^3 + 1) + (2^5 + 1) = 44,$$
$$n(z_7) = (2^7 + 1) + (2^4 + 1) + (2^2 + 1) = 151.$$

Theorem 4.9. *Suppose that*

$$f : \mathbb{F}_2^{2n} \to \mathbb{F}_2^n,$$
$$z = f(x, y) = x + y \pmod{2^n}.$$

Then for all i, $0 \le i < n$, we have

$$z_i = \bigoplus_{t=0}^{2^i} x^t \left(\prod_{k=0}^{i-1} y_k^{\bar{t}_k \oplus \prod_{r=0}^{k-1} \bar{t}_r} \right).$$

Here, we define $(0, ..., 0)^{(0,...,0)} := 1$ and $\prod_{r=0}^{-1} \bar{t}_r := 1$.

Proof. According to Theorem 4.4, for all i, $0 \le i < n$, we have

$$z_i = \bigoplus_{t=0}^{2^i} x^t y^{2^i - t} = \bigoplus_{t=0}^{2^i} x^t y^{\bar{t}+1}.$$

Now, if we denote the j-th bit of $\bar{t} + 1$ by $[\bar{t} + 1]_j$, then we have

$$[\bar{t} + 1]_0 = \bar{t}_0 \oplus 1 = \bar{t}_0 \oplus \prod_{r=0}^{-1} \bar{t}_r,$$

and according to (3.1), for all j, $1 \le j \le i - 1$, we get

$$[\bar{t} + 1]_j = \bar{t}_j \oplus c_j, \quad c_j = \bar{t}_{j-1} 1_{j-1} \oplus c_{j-1} \left(\bar{t}_{j-1} \oplus 1_{j-1}\right).$$

Thus,

$$c_1 = \bar{t}_0, \quad c_j = c_{j-1} \bar{t}_{j-1} = \prod_{r=0}^{i-1} \bar{t}_r, \quad 3 \le j \le i - 1.$$

So for all j, $0 \le j \le i - 1$, we have

$$[\bar{t} + 1]_j = \bar{t}_j \oplus \prod_{r=0}^{j-1} \bar{t}_r.$$

Therefore,

$$y^{\bar{t}+1} = \prod_{k=0}^{i-1} y_k^{\bar{t}_k \oplus \prod_{r=0}^{k-1} \bar{t}_r}.$$

□

In the proof of the following theorem, we present a formula for the ANF of component Boolean functions of modular addition. This formula is somehow presented in [6], but we use another method to prove this formula.

Theorem 4.10. *Suppose that $n > 3$ and,*

$$f : \mathbb{F}_2^{2n} \to \mathbb{F}_2^n,$$
$$z = f(x, y) = x + y \pmod{2^n}.$$

Then we have

$$z_{n-1} = x_{n-1} \oplus y_{n-1} \oplus \left(\bigoplus_{\substack{0 \le i \le n-2 \\ T_i \in A_i}} (T_i x_i y_i) \right),$$

where $A_i = \{t_{n-2} \ldots t_{i+1} | t_j \in \{x_j, y_j\},\ i+1 \le j \le n-2\}$.

Proof. According to Theorem 4.9, we have

$$z_{n-1} = \bigoplus_{t=0}^{2^{n-1}} x^t y^{\bar{t}+1} = x_{n-1} \oplus y_{n-1} \oplus \bigoplus_{t=1}^{2^{n-1}-1} x^t y^{\bar{t}+1}.$$

We notice that for $1 \le t \le 2^{n-1} - 1$, we have $0 \le p(t) \le n-2$. So,

$$z_{n-1} = x_{n-1} \oplus y_{n-1} \oplus \bigoplus_{1 \le p(t) \le n-2} x^t y^{\bar{t}+1}.$$

By Theorem 4.9, if $p(t) = k$, then,

$$t = (t_{n-2}, \ldots, t_{k+1}, 1, 0, \ldots, 0),$$
$$\bar{t} + 1 = (\bar{t}_{n-2}, \ldots, \bar{t}_{k+1}, 1, 0, \ldots, 0).$$

This shows that $p(t) = p(\bar{t} + 1)$ and their remaining bits in binary representation are complement of each other. This ends the proof. □

As a result of the Theorem 4.10, we have:

Corollary 4.11. *Suppose that* $z = x + y \pmod{2^n}$*. Then the number of terms of degree* d*,* $1 < d \leq n$ *in the ANF of* z_{n-1} *equals to* 2^{d-2} *and there are two monomials of degree one.*

Theorem 4.10 presents the ANF of the output of the component Boolean functions of modular addition modulo a power of two. This relation contains more information than the representation which was presented in [3].

5. Conclusion

The operation of modular addition modulo a power of two is one of the most applied operations in symmetric cryptography. For example, modular addition is used in RC6, MARS and Twofish block ciphers and RC4, Bluetooth and Rabbit stream ciphers. In this paper, we examined statistical and algebraic properties of modular addition modulo a power of two. At first, we obtained the probability distribution of modular addition carry bits and conditional probability distribution of these carry bits. Then, using these probability distributions and Markovity of modular addition carry bits, we computed the joint probability distribution of arbitrary number of modular addition carry bits.

We investigated algebraic properties of modular addition with a constant and we obtained the number of terms along with the algebraic degrees of the component Boolean functions of the output of modular addition with a constant.

Lastly, we presented another formula for the ANF of component Boolean functions of modular addition modulo a power of two. This new formula contains more information than representations which have been presented in cryptographic literature, up to now.

References

[1] Bluetooth SIG, Specification of the Bluetooth System, Version 1.1, 1 Febraury 22, 2001, available at http://www.bluetooth.com.

[2] M. Boesgaard, M. Vesterager, T. Pedersen, J. Christiansen and O. Scavenius: Rabbit: A New High- Performance Stream Cipher, *Proceedings of Fast Software Encryption 2003*, Springer, Berlin, (2003).

[3] A. Braeken, I. Semaef, The ANF of Composition of Addition and Multiplication mod with a Boolean function , *FSE'05*, LNCS 2887, pp. 290-306, Springer Verlage, 2005.

[4] C. Burwick, D. Coppersmith, E. DAvingnon, R. Gennaro, Sh. Halevi, Ch. Julta, S. M. Matyas. Jr, L. OConnor, M. Peyravian, D. Safford, N. Zunnic, MARS

a Candidate Cipher for AES, *Proceeding of 1st Advanced Encryption Standard Candidate Conference*, Venture, California, Aug. 20-22 1998.
[5] Carlet, C., "Vectorial Boolean Functions for Cryptography", in *Boolean Models and Methods in Mathematics*, Computer Science, and Engineering.: Cambridge University Press, 2010, pp. 398-469, available at http://www.math.univ-paris13.fr/ carlet/chap-vectorial-fcts-corr.pdf
[6] Xu, K., Dai, Z., and Dai, Z., "The formulas of coefficients of sum and product of p-adic integers with applications to Witt vectors", available via http://arxiv.org/pdf/1007.0878v1.pdf
[7] R. L. Rivest, The RC4 encryption algorithm, RSA Data Security, Inc., Mar., 1992. *Proceeding of 1st Advanced Encryption Standard Candidate Conference*, Venture, California, Aug. 20-22 1998.
[8] R. L. Rivest, M. J. B. Robshaw, R. Sidney, Y. L. Yin, The RC6 Block Cipher, *Proceeding of 1st Advanced Encryption Standard Candidate Conference*, Venture, California, Aug. 20-22 1998.
[9] B. Schneier, J. Kelsey, D. Whiting, D. Wagner, C.Hall, N. Ferguson, Twofish: A 128-Bit Block cipher, 1998, Available via http://www.counterpane.com/twofish.html.