Des. Codes Cryptogr. (2013) 66:111–127
DOI 10.1007/s10623-012-9668-4

On CCZ-equivalence of addition mod 2^n

Ernst Schulte-Geers

Received: 2 August 2011 / Revised: 21 March 2012 / Accepted: 26 March 2012 /
Published online: 25 May 2012

Abstract We show that addition mod 2^n is CCZ-equivalent to a quadratic vectorial Boolean function. We use this to reduce the solution of systems of differential equations of addition to the solution of an equivalent system of linear equations and to derive a fully explicit formula for the correlation coefficients, which leads to enhanced results about the Walsh transform of addition mod 2^n. The results have direct applications in the cryptanalysis of cryptographic primitives which use addition mod 2^n.

Keywords CCZ-equivalence · Addition mod 2^n · Walsh transform · Differential Equations of Addition

Mathematics Subject Classification 94C10 · 06E30 · 11T71

1 Introduction

Since they are "hard-wired" on modern microprocessors, arithmetic operations are attractive candidates for the use in fast cryptographic primitives. Especially addition mod 2^n is frequently used as one source of Boolean nonlinearity in modern block ciphers and hash functions. The first public proposal to use addition mod 2^n in modern cryptography appears to be Rueppel's summation generator [9]. In the sequel, addition mod 2^n has found widespread use. It is e.g. used in the ciphers FEAL, IDEA, SAFER, RC5, MARS, Twofish, the hash functions SHA-2 and in ARX-functions like the SHA-3 finalists Skein and BLAKE.

We revisit here the compatibility of addition mod 2^n ($\boxplus$) with XOR.

The different aspects of compatibility of XOR with modular additions have been the subject of extensive research, mostly along the following lines:

This is one of several papers published in *Designs, Codes and Cryptography* comprising the "Special Issue on Coding and Cryptography".

E. Schulte-Geers (✉)
Bundesamt für Sicherheit in der Informationstechnik, Bonn, Germany
e-mail: ernst.schulte-geers@bsi.bund.de

(1) The Markov chain of carries:

One way to treat the approximation of addition by XOR probabilistically is to consider the Markov chain generated by the carry bits [1,4,10,11]. Remarkably, the final result on the bias of the carry-bits was obtained only in 2010, when Alquié [1] gave an explicit formula for the bias of the i-th carry bit for any fixed number n of summands.

(2) Differential cryptanalysis (DC):

Beginning with the DC of FEAL by Biham and Shamir [2] partial results begin to appear, but a complete (fully explicit) formula for the differential probabilities of addition mod 2^n was apparently only given in 2001 by Lipmaa and Moriai ([6]). These authors also give efficient algorithms to compute certain maximal differential probabilities.

Closely related to DC of addition are "differential equations of addition" (DEA), where a DEA is an equation of the form $(\mathbf{x} \boxplus \mathbf{y}) \oplus ((\mathbf{x} \oplus \mathbf{a}) \boxplus (\mathbf{y} \oplus \mathbf{b})) = \mathbf{c}$ (where $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are given and $\mathbf{x}$, $\mathbf{y}$ are to be found). Paul and Preneel [8] studied the decision problem for systems of DEAs, showed that it is in $\mathscr{P}$, and moreover gave efficient "combinatorial" algorithms to solve such systems completely if solutions exists.

(3) Linear cryptanalysis (LC):

Wallén [12] gave a formula for the linear correlations of addition mod 2^n and undertook an in depth study of the algorithmic aspects of computing them. However, the formula is not explicit in the sense that one of its ingredients has to be computed recursively (details are given after Theorem 4 below).

Another method for the computation of linear correlations of addition, applicable also in the case of more than two summands, was given by Nyberg and Wallén in [7].

2 Our results

We observe that addition mod 2^n is CCZ-equivalent to a quadratic vectorial Boolean function. For $n \geq 3$ this is a new example of "true" CCZ-equivalence (and the first example for a function which is widely used). We apply this to the circle of questions above.

DC: We show that the solution of systems of DEA can be reduced to the solution of systems of very simple linear equations.

In the course we give a simple re-derivation for the Lipmaa-Moriai explicit formula for the differential probabilities of addition.

LC: We give a completely explicit formula for the Walsh transform of addition. We use this formula to determine for the first time maximal correlations for the case where one or two masks are fixed.

Further we show (in Appendix B) that all our formulæ can be computed efficiently.

The simplicity of our results and the easiness of obtaining them indicate that CCZ-equivalence is the natural framework to treat these aspects of addition.

3 Prelimimaries

3.1 Notations and conventions, reminder

3.1.1 Bits and bit vectors, modular addition

Throughout n is a fixed natural number.

$\mathbb{F}_2$ denotes the field with the two elements 0, 1 ("bits"), $\mathbb{F}_2^n$ denotes the $\mathbb{F}_2$-vector space of bit vectors of length n. We write the elements of $\mathbb{F}_2^n$ as rows (i.e. $1 \times n$ matrices) and

denote them by small bold face letters. Subscripted indices denote coordinates. The left most coordinate of a bit vector $\mathbf{x}$ has the index 0 and corresponds to the least significant bit of $\overline{\mathbf{x}}$. Bit vectors are always written in the form $\mathbf{x} = (x_0, \ldots, x_{n-1})$. (Note that this notation is "little-endian", but that bits in the bytes/integers etc. on a computer are stored "big-endian").

$\oplus$ denotes binary addition (of bits as well as of bit vectors).

For $\mathbf{x} = (x_0, \ldots, x_{n-1}) \in \mathbb{F}_2^n$ we let $\overline{\mathbf{x}} := \sum_{i=0}^{n-1} x_i 2^i \in \mathbb{N}$ denote the natural number represented by $\mathbf{x}$. Identifying bit vectors with natural numbers in this way, the addition mod 2^n of elements $\overline{\mathbf{x}}, \overline{\mathbf{y}} \in \mathbb{Z}_{2^n}$ gives an operation on $\mathbb{F}_2^n$. We denote this "modular addition" of bit vectors by $\boxplus$ (i.e. $\mathbf{x} \boxplus \mathbf{y} \cong (\overline{\mathbf{x}} + \overline{\mathbf{y}}) \bmod 2^n$).

For bit vectors $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$ we denote their bitwise AND with $\mathbf{x} \star \mathbf{y}$ (i.e. $\mathbf{x} \star \mathbf{y} = (x_0 y_0, \ldots, x_{n-1} y_{n-1})$), and their bitwise OR with $\mathbf{x}|\mathbf{y}$ (i.e. $\mathbf{x}|\mathbf{y} = (x_0|y_0, \ldots, x_{n-1}|y_{n-1})$).

The all one vector is denoted by $\mathbf{e}$, the bitwise complement of $\mathbf{x}$ then is $\mathbf{x} \oplus \mathbf{e}$.

$\preceq$ denotes the partial order on $\mathbb{F}_2^n$ defined by $\mathbf{x} \preceq \mathbf{y} \Leftrightarrow x_i \leq y_i \ \forall i \in \{0, \ldots, n-1\}$

The Hamming weight (also called "length") of $\mathbf{x} \in \mathbb{F}_2^n$ is denoted by $|\mathbf{x}|$.

A one-block of length $l \geq 1$ in a row vector $\mathbf{u} = (u_0, \ldots, u_{n-1}) \in \mathbb{F}_2^n$ is an all-one section of $\mathbf{u}$ which cannot be extended, i.e. a subsequence $(u_k, \ldots, u_{k+l-1})$ consisting solely of ones such that $u_{k-1} = 0$ if $k > 0$ and $u_{k+l} = 0$ if $k + l < n$. Zero-blocks are defined analogously. The set of one-blocks in $\mathbf{u}$ is denoted by $\mathscr{B}_1(\mathbf{u})$.

A mapping $\mathbf{g} : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^m$ is called a vectorial Boolean function.

For any $\mathbf{z} \in \mathbb{F}_2^n$ the set $U_{\mathbf{z}} := \{\mathbf{x} \in \mathbb{F}_2^n \ : \ \mathbf{x} \preceq \mathbf{z}\}$ is an $\mathbb{F}_2$-vectorspace of dimension $|\mathbf{z}|$.

The algebraic normal form (ANF) of a vectorial Boolean function $\mathbf{f} : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^m$ is given by the 2^n coefficients $\mathbf{c_f}(\mathbf{z})$, $\mathbf{z} \in \mathbb{F}_2^n$ where $\mathbf{c_f}(\mathbf{z}) := \bigoplus_{\mathbf{x} \preceq \mathbf{z}} \mathbf{f}(\mathbf{x}) \in \mathbb{F}_2^m$. The degree of a vectorial Boolean function is $\deg \mathbf{f} := \max\{|\mathbf{z}| \ : \ \mathbf{c_f}(\mathbf{z}) \neq \mathbf{0}\}$.

A vectorial Boolean function of degree 2 is called quadratic.

The convolution $f * g$ of two functions $f, g : \mathbb{F}_2^n \longrightarrow \mathbb{R}$ is the function $f * g : \mathbb{F}_2^n \longrightarrow \mathbb{R}$ defined by $\mathbf{x} \mapsto (f * g)(\mathbf{x}) := \sum_{\mathbf{y} \in \mathbb{F}_2^n} f(\mathbf{x} \oplus \mathbf{y}) g(\mathbf{y})$

The $k-$th vector of the standard basis of $\mathbb{F}_2^n$ is denoted by $\mathbf{e}_k$, $0 \leq k \leq n-1$.

The superscript t indicates transposition of a matrix.

If $S \in \mathbb{F}_2^{m \times n}$ is a matrix and $\mathbf{x} \in \mathbb{F}_2^n$ we let $S(\mathbf{x}) := \mathbf{x} S^t \in \mathbb{F}_2^m$, i.e. we don't distinguish in notation between S and the linear mapping defined by S w.r. to the standard bases in domain and codomain.

$L : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^n$ denotes the "left shift" $\mathbf{x} = (x_0, \ldots, x_{n-1}) \mapsto L(\mathbf{x}) = (x_1, \ldots, x_{n-1}, 0)$.
$R : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^n$ denotes the "right shift" $\mathbf{x} = (x_0, \ldots, x_{n-1}) \mapsto R(\mathbf{x}) = (0, x_0, \ldots, x_{n-2})$.

Finally $M : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^n$ denotes the (right shifted) "partial sums mapping"

$$\mathbf{x} = (x_0, \ldots, x_{n-1}) \mapsto M(\mathbf{x}) = (0, x_0, x_0 \oplus x_1, \ldots, x_0 \oplus \ldots \oplus x_{n-2})$$

I denotes the identity matrix. Clearly $M = R(I \oplus R)^{-1} = (I \oplus R)^{-1} R$ and $L = R^t$, $R = L^t$ With the convention above M is the matrix with all ones below the main diagonal and zeroes elsewhere.

3.1.2 Fourier transform

We use the Fourier transform on the abelian group $(\mathbb{F}_2^n, \oplus)$ in the following (standard) form:

We write the standard character χ of $(\mathbb{F}_2, \oplus)$ (i.e. $\chi(0) = 1$, $\chi(1) = -1$) in the suggestive form $\chi(x) = (-1)^x$.

For two binary vectors $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$ we let $\mathbf{x} \cdot \mathbf{y} := \mathbf{x}\mathbf{y}^t = \bigoplus_{i=0}^{n-1} x_i \, y_i$.

Each $\mathbf{y} \in \mathbb{F}_2^n$ defines a character $\chi_{\mathbf{y}}$ of $(\mathbb{F}_2^n, \oplus)$ by $\chi_{\mathbf{y}} : \mathbb{F}_2^n \longrightarrow \mathbb{C}$, $\chi_{\mathbf{y}}(x) := \chi(\mathbf{y} \cdot \mathbf{x}) = (-1)^{\mathbf{y} \cdot \mathbf{x}}$.

We identify $\mathbb{F}_2^n$ with its character group via $\mathbf{y} \cong \chi_{\mathbf{y}}$.

For a real function $f : \mathbb{F}_2^n \longrightarrow \mathbb{R}$ and $\mathbf{u} \in \mathbb{F}_2^n$ we define the Fourier transform $\hat{f}$ by

$$\hat{f}(\mathbf{u}) := \sum_{\mathbf{x} \in \mathbb{F}_2^n} f(\mathbf{x})(-1)^{\mathbf{u}\cdot\mathbf{x}}$$

For a vectorial Boolean function $\mathbf{g} : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^m$ and $\mathbf{t} \in \mathbb{F}_2^m, \mathbf{u} \in \mathbb{F}_2^n$ we define the Walsh transform $W_{\mathbf{g}}$ by

$$W_{\mathbf{g}}(\mathbf{t}, \mathbf{u}) := \sum_{\mathbf{x} \in \mathbb{F}_2^n} (-1)^{\mathbf{t}\cdot\mathbf{g}(\mathbf{x}) \oplus \mathbf{u}\cdot\mathbf{x}}.$$

We call the value $\phi_{\mathbf{g}}(\mathbf{t}, \mathbf{u}) := \frac{W_{\mathbf{g}}(\mathbf{t},\mathbf{u})}{2^n}$ the "linear correlation" of $\mathbf{g}$ at $(\mathbf{t}, \mathbf{u})$.

Clearly $W_{\mathbf{g}} = \hat{1}_{G_{\mathbf{g}}}$, that is, the Walsh transform of the vectorial Boolean function $\mathbf{g}$ is simply the ordinary Fourier transform of the indicator function $1_{G_{\mathbf{g}}}$ of the graph

$$G_{\mathbf{g}} := \{(\mathbf{x}, \mathbf{g}(\mathbf{x})) \ : \ \mathbf{x} \in \mathbb{F}_2^n \} \subset \mathbb{F}_2^{n+m}.$$

3.1.3 CCZ-equivalence of vectorial Boolean functions

Vectorial Boolean functions $\mathbf{h}, \mathbf{g} : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^m$ are called CCZ-equivalent ([3]) iff there is an affine-linear bijection A of $\mathbb{F}_2^{n+m}$ such that

$$G_{\mathbf{g}} = A(G_{\mathbf{h}}) := \{ A(\mathbf{y}) \ : \ \mathbf{y} \in G_{\mathbf{h}} \}$$

that is, $\mathbf{h}$ and $\mathbf{g}$ are CCZ-equivalent iff the graph $G_{\mathbf{h}}$ of $\mathbf{h}$ can be affine-linearly transformed into the graph $G_{\mathbf{g}}$ of $\mathbf{g}$.

4 CCZ-equivalence

4.1 Recursive carry computation

The "grade school method" for adding integers $\overline{\mathbf{x}}, \overline{\mathbf{y}}$ in the binary number system is well known. We recall it here since we need it in the sequel. Let $c_j = c_j(\mathbf{x}, \mathbf{y})$ denote the carry-bit which is produced in position $j-1$ and added in position j and let $\mathbf{c} = \mathbf{c}(\mathbf{x}, \mathbf{y})$ be the carry vector produced in the modular addition of $\mathbf{x}$ and $\mathbf{y}$.

The grade school addition algorithm performs the following recursion ([9]):

$$\begin{aligned} c_0 &= 0 \\ (\mathbf{x} \boxplus \mathbf{y})_j &= x_j \oplus y_j \oplus c_j \qquad \text{and} \\ c_{j+1} &= x_j y_j \oplus (x_j \oplus y_j) c_j \end{aligned}$$

(since a carry is produced at position j iff the majority of x_j, y_j, c_j is 1).

Clearly this algorithm computes also the carry for the addition mod 2^n if only performed up to $j+1 = n-1$, so that $\mathbf{x} \boxplus \mathbf{y} = \mathbf{x} \oplus \mathbf{y} \oplus \mathbf{c}$, where $\mathbf{c}$ is computed as above.

4.2 Linear equivalence of the modular addition graph

For cryptographic purposes one is interested in the Walsh spectrum and in the differential spectrum of modular addition. These are properties of the graph

$$G_{\boxplus} := \{(\mathbf{x}, \mathbf{y}, \mathbf{x} \boxplus \mathbf{y}) \,:\, \mathbf{x} \in \mathbb{F}_2^n \,, \mathbf{y} \in \mathbb{F}_2^n \,\} \subset \mathbb{F}_2^{3n}.$$

In general, if for sets $A, B \subset \mathbb{F}_2^m$ and a linear permutation T the relation $A = T(B)$ holds, then the Walsh resp. differential properties are transformed by the relations

$$\hat{1}_A(\mathbf{u}) = \hat{1}_B(T^t(\mathbf{u}))$$

resp.

$$(1_A * 1_A)(\mathbf{x}) = (1_B * 1_B)(T^{-1}(\mathbf{x})) \quad \text{(here } * \text{ denotes convolution (see 3.1.1))}$$

It is therefore of great value to observe that the graph of modular addition can be linearly transformed into the graph of a much simpler vectorial Boolean function.

To describe this let $\mathbf{q} : \mathbb{F}_2^n \times \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^n$ be the mapping $(\mathbf{x}, \mathbf{y}) \mapsto M(\mathbf{x} \star \mathbf{y})$, and let

$$G_{\mathbf{q}} := \{(\mathbf{x}, \mathbf{y}, \mathbf{q}(\mathbf{x}, \mathbf{y})) \,:\, \mathbf{x} \in \mathbb{F}_2^n \,,\, \mathbf{y} \in \mathbb{F}_2^n\}$$

be the graph of $\mathbf{q}$.

The following lemma is the basis for our results.

Lemma 1 *The mapping*

$$\beta : \mathbb{F}_2^n \times \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^n \times \mathbb{F}_2^n \,, \quad (\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x} \oplus \mathbf{c}(\mathbf{x}, \mathbf{y}), \mathbf{y} \oplus \mathbf{c}(\mathbf{x}, \mathbf{y}))$$

is a bijection with inverse

$$\alpha : \mathbb{F}_2^n \times \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^n \times \mathbb{F}_2^n \,, \quad (\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y}), \mathbf{y} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y}))$$

Proof Let $\mathbf{c} = \mathbf{c}(\mathbf{x}, \mathbf{y})$ be the carry vector produced in the modular addition of $\mathbf{x}$ and $\mathbf{y}$.

By the grade school algorithm (see 4.1) we have $c_0 = 0$ and $c_{j+1} \oplus c_j = (x_j \oplus c_j) \cdot (y_j \oplus c_j)$ for $j = 0, \ldots, n-2$. Thus $(I \oplus R)(\mathbf{c}) = R((\mathbf{x} \oplus \mathbf{c}) \star (\mathbf{y} \oplus \mathbf{c}))$ and hence

$\mathbf{c} = M((\mathbf{x} \oplus \mathbf{c}) \star (\mathbf{y} \oplus \mathbf{c})) = \mathbf{q}(\mathbf{x} \oplus \mathbf{c}, \mathbf{y} \oplus \mathbf{c})$.

Since $(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \oplus \mathbf{c}, \mathbf{y} \oplus \mathbf{c}) \oplus (\mathbf{c}, \mathbf{c})$ and since $\mathbf{c}$ can be computed from $(\mathbf{x} \oplus \mathbf{c})$, $(\mathbf{y} \oplus \mathbf{c})$ the mapping $\mathbb{F}_2^{2n} \ni (\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x} \oplus \mathbf{c}(\mathbf{x}, \mathbf{y}), \mathbf{y} \oplus \mathbf{c}(\mathbf{x}, \mathbf{y})) \in \mathbb{F}_2^{2n}$ is injective (hence bijective). Since $\mathbf{c} = \mathbf{q}(\mathbf{x} \oplus \mathbf{c}, \mathbf{y} \oplus \mathbf{c})$ the inverse mapping is $(\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y}), \mathbf{y} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y}))$ □

Theorem 1 *(CZZ-equivalence of $\boxplus$ and $\mathbf{q}$)*
The linear mapping $T : (\mathbb{F}_2^n)^3 \longrightarrow (\mathbb{F}_2^n)^3$, $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \mapsto T(\mathbf{x}, \mathbf{y}, \mathbf{z}) = (\mathbf{x} \oplus \mathbf{z}, \mathbf{y} \oplus \mathbf{z}, \mathbf{x} \oplus \mathbf{y} \oplus \mathbf{z})$ maps the graph of $\mathbf{q}$ bijectively onto the graph of $\boxplus$.

Proof Let $(\mathbf{x}', \mathbf{y}') := \beta(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \oplus \mathbf{c}(\mathbf{x}, \mathbf{y}), \mathbf{y} \oplus \mathbf{c}(\mathbf{x}, \mathbf{y}))$ We have

$$\begin{aligned} G_{\boxplus} &= \{(\mathbf{x}, \mathbf{y}, \mathbf{x} \oplus \mathbf{y} \oplus \mathbf{c}(\mathbf{x}, \mathbf{y})) \,:\, \mathbf{x} \in \mathbb{F}_2^n \,, \mathbf{y} \in \mathbb{F}_2^n\} \\ &= \{(\mathbf{x} \oplus \mathbf{c}(\mathbf{x}, \mathbf{y}) \oplus \mathbf{c}(\mathbf{x}, \mathbf{y}), \mathbf{y} \oplus \mathbf{c}(\mathbf{x}, \mathbf{y}) \oplus \mathbf{c}(\mathbf{x}, \mathbf{y}), \mathbf{x} \oplus \mathbf{y} \oplus \mathbf{c}(\mathbf{x}, \mathbf{y})) \,:\, \mathbf{x} \in \mathbb{F}_2^n \,, \mathbf{y} \in \mathbb{F}_2^n\} \\ &= \{(\mathbf{x}' \oplus \mathbf{q}(\mathbf{x}', \mathbf{y}'), \mathbf{y}' \oplus \mathbf{q}(\mathbf{x}', \mathbf{y}'), \mathbf{x}' \oplus \mathbf{y}' \oplus \mathbf{q}(\mathbf{x}', \mathbf{y}')) \,:\, \mathbf{x}' \in \mathbb{F}_2^n \,, \mathbf{y}' \in \mathbb{F}_2^n\} \end{aligned}$$

where the final equality follows from Lemma 1. The result is now obvious. □

Some remarks:

1. Thus for each $n \geq 1$ the mappings $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \boxplus \mathbf{y}$ and $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{q}(\mathbf{x}, \mathbf{y})$ give an example of a pair of vectorial Boolean functions $F, G : \mathbb{F}_2^{2n} \longrightarrow \mathbb{F}_2^n$ which are CCZ-equivalent.
2. Written in coordinates, $\mathbf{q}$ is the vectorial Boolean function

$$(\mathbf{x}, \mathbf{y}) \mapsto (0, x_0 y_0, x_0 y_0 \oplus x_1 y_1, \ldots, x_0 y_0 \oplus \ldots \oplus x_{n-2} y_{n-2})$$

   For $j \geq 1$ the ANF of the coordinate function $q_j(\mathbf{x}, \mathbf{y})$ is of degree 2 in the variables $x_0, \ldots, x_{n-1}, y_0, \ldots, y_{n-1}$. On the other hand it is easy to see that the ANF of the coordinate function $(\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x} \boxplus \mathbf{y})_j$ is of degree $j + 1$ in these variables. For $n \geq 3$ therefore $\boxplus$ and $\mathbf{q}$ can not possibly be (extended) affine equivalent.
3. Clearly addition mod 2^n is (by a linear transformation) then also CCZ-equivalent to

$$(\mathbf{x}, \mathbf{y}) \mapsto (0, x_0 y_0, x_1 y_1, \ldots, x_{n-2} y_{n-2})$$

   (and others). This may be viewed as the simplest shape that addition mod 2^n takes when considered up to CCZ-equivalence.
4. The recursion for the carries is also the starting point of all other published investigations of the Boolean properties of addition mod 2^n. New in the approach here is the use of the bijections α, β to exploit CCZ-equivalence.
5. The graph of addition mod 2^n is thus "only" the graph of a quadratic vectorial function (and the reason for this is that the recursion for the carry-bits is quadratic). From this perspective addition mod 2^n is a very simple Boolean operation. Here we use this fact only to revisit the linear and differential properties of addition mod 2^n, but it should be clear that it can potentially also be useful for algebraic attacks.

5 Differential properties

The differential properties of quadratic vectorial Boolean functions are easy to describe, since their difference functions are affine-linear. To take advantage of this fact here, we observe:

Theorem 2 (*Differential equations of addition* (*DEA*) *are equivalent to systems of linear equations*)
Let $\mathbf{a}, \mathbf{b}, \mathbf{d} \in \mathbb{F}_2^n$ *and for* $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$ *let* $(\mathbf{x}', \mathbf{y}') := \alpha(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y}), \mathbf{y} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y}))$. *Then* $\mathbf{q}(\mathbf{x} \oplus \mathbf{a}, \mathbf{y} \oplus \mathbf{b}) = \mathbf{q}(\mathbf{x}, \mathbf{y}) \oplus \mathbf{d} \Leftrightarrow$

$$(\mathbf{x}' \boxplus \mathbf{y}') \oplus \big((\mathbf{x}' \oplus \mathbf{a} \oplus \mathbf{d}) \boxplus (\mathbf{y}' \oplus \mathbf{b} \oplus \mathbf{d})\big) = \mathbf{a} \oplus \mathbf{b} \oplus \mathbf{d}$$

Proof From Lemma 1 we have for $\mathbf{q} = \mathbf{q}(\mathbf{x}, \mathbf{y})$ that $\mathbf{q} = \mathbf{c}(\mathbf{x} \oplus \mathbf{q}, \mathbf{y} \oplus \mathbf{q})$ for all $\mathbf{x}, \mathbf{y}$ and therefore $\mathbf{x} \oplus \mathbf{y} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y}) = \mathbf{x}' \oplus \mathbf{y}' \oplus \mathbf{c}(\mathbf{x}', \mathbf{y}') = \mathbf{x}' \boxplus \mathbf{y}'$.
For the "if"-direction assume that $\mathbf{q}(\mathbf{x} \oplus \mathbf{a}, \mathbf{y} \oplus \mathbf{b}) = \mathbf{q}(\mathbf{x}, \mathbf{y}) \oplus \mathbf{d}$. Then

$$(\mathbf{x} \oplus \mathbf{a})' = \mathbf{x} \oplus \mathbf{a} \oplus \mathbf{q}(\mathbf{x} \oplus \mathbf{a}, \mathbf{y} \oplus \mathbf{b}) = \mathbf{x} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y}) \oplus \mathbf{a} \oplus \mathbf{d} = \mathbf{x}' \oplus \mathbf{a} \oplus \mathbf{d}$$
$$(\mathbf{y} \oplus \mathbf{b})' = \mathbf{y} \oplus \mathbf{b} \oplus \mathbf{q}(\mathbf{x} \oplus \mathbf{a}, \mathbf{y} \oplus \mathbf{b}) = \mathbf{y} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y}) \oplus \mathbf{b} \oplus \mathbf{d} = \mathbf{y}' \oplus \mathbf{b} \oplus \mathbf{d}$$

and $(\mathbf{x} \oplus \mathbf{y} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y})) \oplus (\mathbf{x} \oplus \mathbf{a} \oplus \mathbf{y} \oplus \mathbf{b} \oplus \mathbf{q}(\mathbf{x} \oplus \mathbf{a}, \mathbf{y} \oplus \mathbf{b})) = \mathbf{a} \oplus \mathbf{b} \oplus \mathbf{d}$, that is
$(\mathbf{x}' \boxplus \mathbf{y}') \oplus \big((\mathbf{x}' \oplus \mathbf{a} \oplus \mathbf{d}) \boxplus (\mathbf{y}' \oplus \mathbf{b} \oplus \mathbf{d})\big) = \mathbf{a} \oplus \mathbf{b} \oplus \mathbf{d}$
For the "only-if" direction assume that $\mathbf{c}(\mathbf{x}' \oplus \mathbf{a} \oplus \mathbf{d}, \mathbf{y}' \oplus \mathbf{b} \oplus \mathbf{d}) = \mathbf{c}(\mathbf{x}', \mathbf{y}') \oplus \mathbf{d}$. Then

$$\mathbf{x}' \oplus \mathbf{a} \oplus \mathbf{d} \oplus \mathbf{c}(\mathbf{x}' \oplus \mathbf{a} \oplus \mathbf{d}, \mathbf{y}' \oplus \mathbf{b} \oplus \mathbf{d}) = \mathbf{x}' \oplus \mathbf{c}(\mathbf{x}', \mathbf{y}') \oplus \mathbf{a} = \mathbf{x} \oplus \mathbf{a}$$
$$\mathbf{y}' \oplus \mathbf{b} \oplus \mathbf{d} \oplus \mathbf{c}(\mathbf{x}' \oplus \mathbf{a} \oplus \mathbf{d}, \mathbf{y}' \oplus \mathbf{b} \oplus \mathbf{d}) = \mathbf{y}' \oplus \mathbf{c}(\mathbf{x}', \mathbf{y}') \oplus \mathbf{b} = \mathbf{y} \oplus \mathbf{b}$$

and $(\mathbf{x}' \oplus \mathbf{y}' \oplus \mathbf{c}(\mathbf{x}', \mathbf{y}')) \oplus \big(\mathbf{x}' \oplus \mathbf{a} \oplus \mathbf{y}' \oplus \mathbf{b} \oplus \mathbf{c}(\mathbf{x}' \oplus \mathbf{a} \oplus \mathbf{d}, \mathbf{y}' \oplus \mathbf{b} \oplus \mathbf{d})\big) = \mathbf{a} \oplus \mathbf{b} \oplus \mathbf{d}$, that is

$$(\mathbf{x} \oplus \mathbf{y} \oplus \mathbf{q}(\mathbf{x}, \mathbf{y})) \oplus ((\mathbf{x} \oplus \mathbf{a} \oplus \mathbf{y} \oplus \mathbf{b} \oplus \mathbf{q}(\mathbf{x} \oplus \mathbf{a}, \mathbf{y} \oplus \mathbf{b}))) = \mathbf{a} \oplus \mathbf{b} \oplus \mathbf{d} \qquad \square$$

Equivalently

$$\mathbf{q}(\mathbf{x} \oplus \mathbf{b} \oplus \mathbf{d}, \mathbf{y} \oplus \mathbf{a} \oplus \mathbf{d}) = \mathbf{q}(\mathbf{x}, \mathbf{y}) \oplus \mathbf{a} \oplus \mathbf{b} \oplus \mathbf{d} \;\Leftrightarrow\; (\mathbf{x}' \boxplus \mathbf{y}') \oplus \big((\mathbf{x}' \oplus \mathbf{a}) \boxplus (\mathbf{y}' \oplus \mathbf{b})\big) = \mathbf{d}$$

But the left equation is the same as

$$R(\mathbf{x} \star (\mathbf{a} \oplus \mathbf{d})) \oplus R(\mathbf{y} \star (\mathbf{b} \oplus \mathbf{d})) = R((\mathbf{a} \oplus \mathbf{d}) \star (\mathbf{b} \oplus \mathbf{d})) \oplus (I \oplus R)(\mathbf{a} \oplus \mathbf{b} \oplus \mathbf{d}) \quad (1)$$

i.e. a (very simple) system of linear equations.

Thus the set of solutions of a DEA is simply an affine linear subspace of $(\mathbb{F}_2^n)^2$, "deformed" by α. This observation reduces the question of solvability resp. the computation of all solutions of a system of DEAs to linear algebra: just solve the equivalent linear system first, and then compute the image of the solution set under α. (In Appendix B it is shown that $\mathbf{q}$ resp. α can be computed very efficiently). We remark that single equations of this type can easily be solved explicitly:

Lemma 2 *Let* $\mathbf{a}, \mathbf{b}, \mathbf{d} \in \mathbb{F}_2^n$ *and consider the equation*

$$(\mathbf{a} \star \mathbf{x}) \oplus (\mathbf{b} \star \mathbf{y}) = \mathbf{d} \quad (2)$$

Then (2) has a solution iff $\mathbf{d} \preceq \mathbf{a}|\mathbf{b}$. *In this case* $(\mathbf{x}_0, \mathbf{y}_0) := (\mathbf{a} \star \mathbf{d}, (\mathbf{b} \oplus \mathbf{a} \star \mathbf{b}) \star \mathbf{d})$ *is a solution. All other solutions can be obtained from* $(\mathbf{x}_0, \mathbf{y}_0)$ *by modifying* $\mathbf{y}_0$ *arbitrarily at positions* i *were* $a_i = 1, b_i = 0$, *modifying* $\mathbf{x}_0$ *arbitrarily at the positions where* $a_i = 0, b_i = 1$, *modifying* $\mathbf{x}_0, \mathbf{y}_0$ *in the same way (i.e. leaving* $x_i \oplus y_i$ *unchanged) where* $a_i = b_i = 1$ *and choosing* x_i, y_i *arbitrarily if* $a_i = b_i = 0$. *Thus there are* $2^{2n-|\mathbf{a}|\mathbf{b}|}$ *solutions, if solutions exist.*

Proof Since the left hand side of (2) clearly is $\preceq \mathbf{a}|\mathbf{b}$, solutions of (2) can only exist if $\mathbf{d} \preceq \mathbf{a}|\mathbf{b}$. That $(\mathbf{x}_0, \mathbf{y}_0)$ is in this case a solution is easily seen when bit vectors are interpreted as indicator functions of sets. The next statement is simply the basic fact that every solution of an inhomogenous system of linear equations may be written as the sum of one special solution of the inhomogeneous system and a solution of the homogenous system, formulated for the special kind of linear equations above. The number of solutions is then obvious. $\square$

This leads immediately to the solutions of (1).

Corollary 1 *The DEA* $(\mathbf{x}' \boxplus \mathbf{y}') \oplus \big((\mathbf{x}' \oplus \mathbf{a}) \boxplus (\mathbf{y}' \oplus \mathbf{b})\big) = \mathbf{d}$ *(or equivalently (1)) has solutions iff*

$$(I \oplus R)(\mathbf{a} \oplus \mathbf{b} \oplus \mathbf{d}) \preceq R(\mathbf{a} \oplus \mathbf{d})|R(\mathbf{b} \oplus \mathbf{d})$$

In this case, with $\hat{\mathbf{d}} := R((\mathbf{a} \oplus \mathbf{d}) * (\mathbf{b} \oplus \mathbf{d})) \oplus (I \oplus R)(\mathbf{a} \oplus \mathbf{b} \oplus \mathbf{d})$, $\hat{\mathbf{a}} := R(\mathbf{a} \oplus \mathbf{d})$, $\hat{\mathbf{b}} := R(\mathbf{b} \oplus \mathbf{d})$ *a special solution of* $(**)$ *is given by* $(\mathbf{x}_0, \mathbf{y}_0) := (L(\hat{\mathbf{a}} \star \hat{\mathbf{d}}), L((\hat{\mathbf{b}} \oplus \hat{\mathbf{a}} \star \hat{\mathbf{b}}) \star \hat{\mathbf{d}}))$. *Since* $\mathbf{x}_0 \star \mathbf{y}_0 = \mathbf{0}$ *for this special solution* $\mathbf{x}_0' = \mathbf{x}_0$, $\mathbf{y}_0' = \mathbf{y}_0$.

Combining the two preceding results we find:

Theorem 3 (Lipmaa-Moriai) *Let* (X, Y) *be uniformly distributed on* $\mathbb{F}_2^n \times \mathbb{F}_2^n$. *Then*

$$P\left((X \boxplus Y) \oplus ((X \oplus \mathbf{a}) \boxplus (Y \oplus \mathbf{b})) = \mathbf{d}\right) = 1_{\{(I \oplus R)(\mathbf{a} \oplus \mathbf{b} \oplus \mathbf{d}) \preceq R((\mathbf{a} \oplus \mathbf{d})|(\mathbf{b} \oplus \mathbf{d}))\}} 2^{-|(R((\mathbf{a} \oplus \mathbf{d})|(\mathbf{b} \oplus \mathbf{d}))|}$$

This is Theorem 1 from [6] (up to notation), where it forms the basis of an extensive discussion of the differential spectrum of addition mod 2^n.

Lemma 3 *Let* $\mathbf{a}, \mathbf{d} \in \mathbb{F}_2^n$ *and consider the equation*

$$\mathbf{x} \oplus \mathbf{d} = \mathbf{x} \boxplus \mathbf{a} \tag{3}$$

Then (3) has a solution iff $(I \oplus R)(\mathbf{a} \oplus \mathbf{d}) \preceq R(\mathbf{d})$. *In this case* $\mathbf{x}_0 := \mathbf{d} \star L(\mathbf{a} \oplus \mathbf{d})$ *is a solution. All other solutions can be obtained from* $\mathbf{x}_0$ *by modifying* $\mathbf{x}_0$ *arbitrarily at positions* i *were* $d_i = 0$, *or in the highest position. Thus there are* $2^{n-|R(\mathbf{d})|}$ *solutions if solutions exist.*

Proof This is the special case $\mathbf{y}' = 0, \mathbf{b} = 0$ of the above. Thus $\mathbf{y} = \mathbf{0}$, $\mathbf{x}' = \mathbf{x}$ and $(\mathbf{x}, \mathbf{0})$ is a solution of (1) iff

$$R(\mathbf{x} \star \mathbf{d}) = R(\mathbf{d} \star (\mathbf{a} \oplus \mathbf{d})) \oplus (I \oplus R)(\mathbf{a} \oplus \mathbf{d}) =: \mathbf{t}$$

Clearly this equation has a solution iff $\mathbf{t} \preceq R(\mathbf{d})$, i.e. iff $\mathbf{t} \star R(\mathbf{d}) = \mathbf{t}$ and in this case especially both $\mathbf{x}$ s.th. $R(\mathbf{x}) = \mathbf{t} \star R(\mathbf{d})$ are solutions. Since $\mathbf{t} \star R(\mathbf{d}) = (\mathbf{a} \oplus \mathbf{d}) \star R(\mathbf{d})$ the first claim follows. Since exactly the bits of $\mathbf{x}$ at the positions $i \in \{0, \ldots, n-2\}$ where $d_i = 1$ are fixed, the remaining bits may be arbitrarily chosen. □

Thus in this special case solutions exist iff $(\mathbf{a} \oplus R(\mathbf{a}) \oplus \mathbf{d}) \star (R(\mathbf{d}) \oplus \mathbf{e}) = \mathbf{0}$. This was (in an equivalent formulation) already remarked in [5]. There it was proposed to solve the equation by a probabilistic automaton.

Finally we note that solving systems of DEAs is much easier than solving general linear equations. In fact, by Theorem 2 a system of k DEAs is equivalent to a system of equations of the form

$$(\mathbf{a}_i \star \mathbf{x}) \oplus (\mathbf{b}_i \star \mathbf{y}) = \mathbf{d}_i \quad i = 1, \ldots, k$$

(plus trivial equations for the least significant bits). The linear equations on the left hand side of each such equation are described by a coefficient matrix $\begin{pmatrix} A_i & B_i \end{pmatrix}$ where A_i, B_i are diagonal matrices. Therefore such a system of equations can be solved extremely efficiently.

6 Linear correlations of addition mod 2^n

In this section we revisit linear correlations of addition. We show that CCZ-equivalence leads to a simple explicit formula for the correlation coefficients and use this formula to extend the results of [12].

Let $\mathbf{u} := (u_0, \ldots, u_{n-1})$, $\mathbf{v} := (v_0, \ldots, v_{n-1})$, $\mathbf{w} := (w_0, \ldots, w_{n-1}) \in \mathbb{F}_2^n$
($\mathbf{u}$ is the "output mask", and $\mathbf{v}, \mathbf{w}$ are the "input masks"), and let $\mathbf{z} = M^t(\mathbf{u})$.
We are interested in the linear correlation coefficients of addition mod 2^n

$$\phi_2(\mathbf{u}, \mathbf{v}, \mathbf{w}) := \frac{1}{2^{2n}} \sum_{\mathbf{x} \in \mathbb{F}_2^n, \mathbf{y} \in \mathbb{F}_2^n} (-1)^{\mathbf{u} \cdot (\mathbf{x} \boxplus \mathbf{y}) \oplus \mathbf{v} \cdot \mathbf{x} \oplus \mathbf{w} \cdot \mathbf{y}} = \frac{1}{2^{2n}} \hat{1}_{G_\boxplus}(\mathbf{v}, \mathbf{w}, \mathbf{u})$$

By the above we need only compute the Walsh transform of $\mathbf{q}$, as (by Theorem 1) the relation

$$\phi_2(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \phi_{\mathbf{q}}(\mathbf{u} \oplus \mathbf{v} \oplus \mathbf{w}, \mathbf{u} \oplus \mathbf{v}, \mathbf{u} \oplus \mathbf{w})$$

holds. But $\phi_{\mathbf{q}}(\mathbf{u}, \mathbf{v}, \mathbf{w})$ is easy to compute:

$$\begin{aligned} \phi_{\mathbf{q}}(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \tfrac{1}{2^{2n}} \textstyle\sum_{\mathbf{x} \in \mathbb{F}_2^n, \mathbf{y} \in \mathbb{F}_2^n} (-1)^{M^t(\mathbf{u}) \cdot (\mathbf{x} \star \mathbf{y}) \oplus \mathbf{v} \cdot \mathbf{x} \oplus \mathbf{w} \cdot \mathbf{y}} \\ &= \delta_{0, v_{n-1}} \delta_{0, w_{n-1}} \textstyle\prod_{i=0}^{n-2} \left(\frac{1}{4} \sum_{x_i \in \mathbb{F}_2, y_i \in \mathbb{F}_2} (-1)^{z_i x_i y_i + v_i x_i + w_i y_i} \right) \\ &= 1_{\{\mathbf{v} \preceq \mathbf{z}\}} 1_{\{\mathbf{w} \preceq \mathbf{z}\}} (-1)^{\mathbf{v} \cdot \mathbf{w}} 2^{-|\mathbf{z}|} \end{aligned}$$

(Here we have used the easily verified fact that for bits a, b, c the relation $\frac{1}{4}\sum_{x\in\mathbb{F}_2, y\in\mathbb{F}_2}(-1)^{axy\oplus bx\oplus cy} = 1_{\{b\preceq a\}}1_{\{c\preceq a\}}(-1)^{bc}2^{-a}$ holds.)
We thus have:

Theorem 4 (Walsh transform of addition mod 2^n)
Let $\mathbf{z} := M^t(\mathbf{u}\oplus\mathbf{v}\oplus\mathbf{w})$. *Then*

$$\phi_2(\mathbf{u},\mathbf{v},\mathbf{w}) = 1_{\{\mathbf{u}\oplus\mathbf{v}\preceq\mathbf{z}\}}1_{\{\mathbf{u}\oplus\mathbf{w}\preceq\mathbf{z}\}}(-1)^{(\mathbf{u}\oplus\mathbf{w})\cdot(\mathbf{u}\oplus\mathbf{v})}2^{-|\mathbf{z}|} \quad \textit{resp.}$$
$$\phi_2(\mathbf{u},\mathbf{u}\oplus\mathbf{v},\mathbf{u}\oplus\mathbf{w}) = 1_{\{\mathbf{v}\preceq\mathbf{z}\}}1_{\{\mathbf{w}\preceq\mathbf{z}\}}(-1)^{\mathbf{v}\cdot\mathbf{w}}2^{-|\mathbf{z}|}$$

Some remarks

1. This theorem is equivalent to Theorem 1 of [12], however, there the simple explicit description of $\mathbf{z}$ was not obtained. Instead $\mathbf{z}$ had to be computed recursively from the vectors $\mathbf{u}$ and $\mathbf{v}\oplus\mathbf{w}\oplus\mathbf{e}$. It is this explicit description of $\mathbf{z}$ as $\mathbf{z} = M^t(\mathbf{u}\oplus\mathbf{v}\oplus\mathbf{w})$ which gives additional insight into the Walsh spectrum of addition.
2. If the absolute value of $\phi_2(\mathbf{u},\mathbf{v},\mathbf{w})$ is non-zero it e.g. only depends on the Hamming weight of the binary vector $\mathbf{z} = M^t(\mathbf{u}\oplus\mathbf{v}\oplus\mathbf{w})$. With $\mathbf{s} := \mathbf{u}\oplus\mathbf{v}\oplus\mathbf{w}$ we have $\mathbf{z} = (s_1\oplus\ldots\oplus s_{n-1}, s_2\oplus\ldots\oplus s_{n-1}, \ldots, s_{n-2}\oplus s_{n-1}, s_{n-1}, 0)$.
3. By the above we have the following "pencil and paper" method to compute $\phi_2(\mathbf{u},\mathbf{v},\mathbf{w})$:
   (a) compute $\mathbf{s} := \mathbf{u}\oplus\mathbf{v}\oplus\mathbf{w}$.
   (b) compute recursively $z_{n-1} = 0$, $z_{i-1} = z_i \oplus s_i$ for $i \geq 1$.
   (c) $\phi_2(\mathbf{u},\mathbf{v},\mathbf{w})$ is non-zero, iff for all i with $z_i = 0$ the equality $u_i = v_i = w_i$ holds. In this case the absolute value of $\phi_2(\mathbf{u},\mathbf{v},\mathbf{w})$ is $2^{-|\mathbf{z}|}$ and the sign is -1 iff the Hamming weight $|(\mathbf{u}\oplus\mathbf{v})\star(\mathbf{u}\oplus\mathbf{w})|$ is odd.
4. Clearly we have in the above computed the Walsh transform of special quadratic Boolean functions. How do these results compare to the general case?
   Let $A\in\mathbb{F}_2^{n\times n}$ and consider the quadratic Boolean function $\mathbb{F}_2^n \ni \mathbf{x}\mapsto f(\mathbf{x}) := (\mathbf{x}A)\cdot\mathbf{x}$. Then it is known that the Walsh transform $w_f(\mathbf{u}) := W_f(1,\mathbf{u})$ takes only the values $0, 2^{n-r/2}, -2^{n-r/2}$, where r is the (necessarily even) rank of the matrix $S := A\oplus A^t$. Further, the support of w_f is an affine subspace of the form $\mathbf{k}_a\oplus S(\mathbb{F}_2^n)$. The exact determination of the affine subspace and of the sign of the Walsh transform can be difficult.
   If Z denotes the $(n\times n)$ diagonal matrix with entries $Z_{i,i} = z_i$ and if $\mathbf{0}$ denotes the $(n\times n)$ zero-matrixwe may take for the quadratic function $(\mathbf{x},\mathbf{y})\mapsto\mathbf{z}\cdot(\mathbf{x}\star\mathbf{y})$:

$$A = \begin{pmatrix}\mathbf{0} & Z\\ \mathbf{0} & \mathbf{0}\end{pmatrix} \text{ and } \quad S = \begin{pmatrix}\mathbf{0} & Z\\ Z & \mathbf{0}\end{pmatrix}$$

   and the relation to the general case becomes obvious.

We collect some properties of the correlation matrix:

Corollary 2 The correlation matrix of addition mod 2^n has the symmetry properties*:*

1. $\phi_2(\mathbf{u},\mathbf{v},\mathbf{w}) = (-1)^{(\mathbf{u}\oplus\mathbf{v})\cdot(\mathbf{u}\oplus\mathbf{v})}\phi_2(\mathbf{v},\mathbf{u},\mathbf{w}) = (-1)^{(\mathbf{u}\oplus\mathbf{w})\cdot(\mathbf{u}\oplus\mathbf{w})}\phi_2(\mathbf{w},\mathbf{v},\mathbf{u})$
2. $\phi_2(\mathbf{u},\mathbf{v},\mathbf{w}) = \phi_2(\mathbf{u},\mathbf{w},\mathbf{v})$

Corollary 3 The correlation matrix of addition mod 2^n has the properties*:*

1. $\phi_2(\mathbf{u},\mathbf{u},\mathbf{u}) = 2^{-|M^t(\mathbf{u})|}$
2. $\phi_2(\mathbf{u},\mathbf{u}\oplus L(\mathbf{u}),\mathbf{u}) = 2^{-|L(\mathbf{u})|}$

3. *If* $\mathbf{e}_k \preceq M^t(\mathbf{u})$ *for a standard basis vector* $\mathbf{e}_k$ *then* $\phi_2(\mathbf{u}, \mathbf{u} \oplus \mathbf{e}_k, \mathbf{u}) \neq 0$. *The corresponding vector* $\mathbf{z} = M^t(\mathbf{u} \oplus \mathbf{e}_k) = M^t(\mathbf{u}) \oplus \bigoplus_{i=0}^{k-1} \mathbf{e}_i$, *i.e. it is computed from the corresponding* $\mathbf{z} = M^t(\mathbf{u})$ *for* $(\mathbf{u}, \mathbf{u}, \mathbf{u})$ *by complementation below* k.

Proof The formulæ in Corollary 2 and (1) and (3) of Corollary 3 follow directly from Theorem 4. (2) of Corollary 3 follows from the observation that $M^t(I \oplus L)(\mathbf{u}) = L(\mathbf{u})$ and the second formula in Theorem 4. □

In cryptanalysis one is mainly interested in correlations of high absolute value.

Of special interest is the situation where one or two masks are given, and the remaining masks are to be determined so as to maximize the absolute value of the correlation.

In the sequel we deal with the problem of determining "best" approximations when one or two of the masks are fixed. We first consider the case of one mask being fixed.

From Corollary 3. (3) we have especially:

if $j \geq 1$ and u_j is the leading bit of $\mathbf{u} \neq 0$ (i.e. $u_j = 1$, and $u_i = 0$ for $i > j$), then

$$|\phi_2(\mathbf{u}, \mathbf{u} \oplus \mathbf{e}_{j-1}, \mathbf{u})| = 2^{|M^t(\mathbf{u})|-j-1}.$$

Thus in "rows" $\mathbf{u}$ with leading bit $j \geq 1$ of the correlation matrix there exist correlations of absolute value

$$2^{\max\{-|M^t(\mathbf{u})|, |M^t(\mathbf{u})|-j-1\}} \geq 2^{-\lfloor (j+1)/2 \rfloor}.$$

But what is the true value of the highest absolute bias? We need some combinatorial preparations to answer this question precisely.

6.1 A combinatorial problem

Consider the following problem: given $\mathbf{u} \in \mathbb{F}_2^n$, let

$$\mathscr{A}_L(\mathbf{u}) := \{\mathbf{x} \in \mathbb{F}_2^n \ : \ \mathbf{x} \oplus \mathbf{u} \preceq L(\mathbf{x})\}$$

Problem: find a $L-$minimal vector for $\mathbf{u}$, i.e. find a vector $\mathbf{a} = \mathbf{a}(\mathbf{u}) \in \mathscr{A}_L(\mathbf{u})$ s.th. the Hamming weight $|L(\mathbf{a})|$ is minimal.

We need the following assertion (the proof is deferred to the Appendix):

Theorem 5 (= Theorem A.3) *Let* $\mathbf{u} \in \mathbb{F}_2^n$.
Let $\mathbf{a}_L(\mathbf{u}) := \mathbf{u} \oplus (\mathbf{u} \star L(\mathbf{u})) \oplus \ldots \oplus (\mathbf{u} \star L(\mathbf{u}) \star L^2(\mathbf{u}) \star \cdots \star L^{n-1}(\mathbf{u}))$.
Then $\mathbf{a}_L(\mathbf{u})$ *is a* L*-minimal vector for* $\mathbf{u}$.

The function $\mathbf{a}_L(\mathbf{u})$ was introduced in [6] under the name of "all one parity". The relation to the combinatorial problem considered here was not made explicit there.

6.2 Maximal correlations

We return to the problem of determining maximal correlations of addition mod 2^n.

In order to have a convenient way to formulate the results we call the elements $\phi_2(\mathbf{u}, ., .)$ the "row" $\mathbf{u}$ of the correlation matrix, and the elements $\phi_2(., \mathbf{v}, \mathbf{w})$ the "column" $(\mathbf{v}, \mathbf{w})$ of the correlation matrix. (By the symmetry properties (Corollary 2) only these cases need to be considered.) We first treat the row case.

6.2.1 Row maxima

Theorem 6 (row maxima) *Let* $\mathbf{u} \in \mathbb{F}_2^n$ *and let* $\mathbf{b}(\mathbf{u}) := (I \oplus L)(\mathbf{a}_L(\mathbf{u}))$

1. *In row* $\mathbf{u}$ *of* ϕ_2 *the element* $\phi_2(\mathbf{u}, \mathbf{b}(\mathbf{u}), \mathbf{u})$ *is an element of maximal modulus*
2. *Let* $d(\mathbf{u}) := |L(\mathbf{a}_L(\mathbf{u}))|$ *be the "minimal exponent" for* $\mathbf{u}$. *Then* $d(\mathbf{u})$ *can be computed as* $d(\mathbf{u}) = \sum_{B \in \mathscr{B}_1(L(\mathbf{u}))} \lceil |B|/2 \rceil$
3. $\phi_2(\mathbf{u}, \mathbf{b}(\mathbf{u}), \mathbf{u}) = 2^{-d(\mathbf{u})}$

We first observe:

Lemma 4 *Let* $\mathbf{u}, \mathbf{s} \in \mathbb{F}_2^n$ *and let* $\mathbf{s} = (I \oplus L)(\mathbf{a})$. *Then*

$$\mathbf{s} \oplus \mathbf{u} \preceq M^t(\mathbf{s}) \Leftrightarrow \mathbf{a} \oplus \mathbf{u} \preceq L(\mathbf{a})$$

Proof For the "if"-direction assume $\mathbf{s} \oplus \mathbf{u} \preceq M^t(\mathbf{s})$. Since $\mathbf{s} = (I \oplus L)(\mathbf{a})$ then $\mathbf{a} \oplus L(\mathbf{a}) \oplus \mathbf{u} \preceq L(\mathbf{a})$, and hence $\mathbf{a} \oplus \mathbf{u} \preceq L(\mathbf{a})$
For the "only-if" direction assume $\mathbf{a} \oplus \mathbf{u} \preceq L(\mathbf{a})$. Then $\mathbf{s} \oplus M^t(\mathbf{s}) \oplus \mathbf{u} \preceq M^t(\mathbf{s})$ and hence $\mathbf{s} \oplus \mathbf{u} \preceq M^t(\mathbf{s})$. □

We now make the connection to the combinatorial problem of the last subsection.

Lemma 5 *Let* $\phi_2(\mathbf{u}, \mathbf{v}, \mathbf{w}) \neq 0$, $\mathbf{s} := \mathbf{u} \oplus \mathbf{v} \oplus \mathbf{w}$. *Then*

1. $\mathbf{s} \oplus \mathbf{u} \preceq M^t(\mathbf{s})$
2. $|\phi_2(\mathbf{u}, \mathbf{v}, \mathbf{w})| \leq 2^{-d(\mathbf{u})}$

Proof (1) if $\mathbf{u} \oplus \mathbf{v} \preceq M^t(\mathbf{s})$ and $\mathbf{u} \oplus \mathbf{w} \preceq M^t(\mathbf{s})$, then also $\mathbf{s} \oplus \mathbf{u} = \mathbf{v} \oplus \mathbf{w} \preceq M^t(\mathbf{s})$.
(2) follows by Theorem 5 since $\mathbf{a} := (I \oplus M^t)(\mathbf{s}) \in \mathscr{A}_L(\mathbf{u})$ (by (1) and Lemma 4).
□

Proof of Theorem 6 Claims (1) and (3): by the preceding lemma $2^{-d(\mathbf{u})}$ is an upper bound for the absolute values of the correlations in row $\mathbf{u}$. We show that there is a correlation in row $\mathbf{u}$ of this modulus. Let $\mathbf{v} := \mathbf{b}(\mathbf{u})$ and $\mathbf{w} := \mathbf{u}$. Then $\mathbf{u} \oplus \mathbf{v} \oplus \mathbf{w} = \mathbf{b}(\mathbf{u})$.

Clearly $\mathbf{0} = \mathbf{u} \oplus \mathbf{w} \preceq M^t(\mathbf{b}(\mathbf{u}))$. Further $\mathbf{u} \oplus \mathbf{v} = \mathbf{u} \oplus \mathbf{b}(\mathbf{u}) \preceq M^t(\mathbf{b}(\mathbf{u}))$ by Lemma 4), since $\mathbf{a}_L(\mathbf{u}) \oplus \mathbf{u} \preceq L(\mathbf{a}_L(\mathbf{u}))$. Thus $\phi_2(\mathbf{u}, \mathbf{b}(\mathbf{u}), \mathbf{u})$ is non-zero. Finally $M^t(\mathbf{b}(\mathbf{u})) = L(\mathbf{a}_L(\mathbf{u}))$ and therefore $|\phi_2(\mathbf{u}, \mathbf{b}(\mathbf{u}), \mathbf{u})| = 2^{-|L(\mathbf{a}_L(\mathbf{u}))|}$. Since $\mathbf{a}_L(\mathbf{u})$ is L-minimal (by Theorem 5) both assertions follow.

(2): follows iteratively using Lemma A 2, 5 from Appendix A. □

Finally from the formula for $\mathbf{a}(\mathbf{u})$ a simple lower bound for $d(\mathbf{u})$ can be obtained:

Corollary 4 *Let* $d(\mathbf{u})$ *be the minimal exponent for* $\mathbf{u}$. *Then* $d(\mathbf{u}) \geq |L(\mathbf{u})|/2$. *If each 1-block in* $L(\mathbf{u})$ *has even length there is equality.*

By Theorem 6.3 the structure and number of ones in $L(\mathbf{u})$ determine the linear approximability of $\mathbf{u} \cdot (\mathbf{x} \boxplus \mathbf{y})$ in a simple way: the less ones, and the better the subdivision in blocks of even length, the higher is the linear approximability.

Example 1 Let $n = 13$ and let $\mathbf{u} = (1110\,1101\,0011\,1) = (7351)_2$ (little-endian!). Then $L(\mathbf{u}) = (1101\,1010\,0111\,0)$ has 1 one-block of length 1, 2 one-blocks of length 2, and one one-block of length 3. Thus $d(\mathbf{u}) = 1 + 2 + 2 = 5$. Further $\mathbf{a}_L(\mathbf{u}) = (1010\,0101\,0010\,1) = (5285)_2$ and $\mathbf{b}(\mathbf{u}) = (1110\,1111\,0111\,1) = (7927)_2$. Hence in row 7351 of the correlation matrix the element $\phi_2(7351, 7927, 7351) = 2^{-5}$ is maximal.

6.2.2 Column maxima

It is also evident how a correlation of maximal modulus can be found when two masks, say $\mathbf{v}$ and $\mathbf{w}$, are fixed. We have to determine $\mathbf{s} = \mathbf{u} \oplus \mathbf{v} \oplus \mathbf{w}$ s.th. $M^t(\mathbf{s})$ has minimal Hamming weight under the side conditions

$$\mathbf{s} \oplus \mathbf{w} = \mathbf{u} \oplus \mathbf{v} \preceq M^t(\mathbf{s}) \quad \text{and} \quad \mathbf{s} \oplus \mathbf{v} = \mathbf{u} \oplus \mathbf{w} \preceq M^t(\mathbf{s}) \ .$$

With $\mathbf{s} = \mathbf{a} \oplus L(\mathbf{a})$ the task then is (equivalence is shown as above):

Determine $\mathbf{a}$ s.th. $L(\mathbf{a})$ has minimal weight under the side conditions

$$\mathbf{a} \oplus \mathbf{v} \preceq L(\mathbf{a}), \quad \mathbf{a} \oplus \mathbf{w} \preceq L(\mathbf{a}) \tag{4}$$

Especially the correlation can only be non-zero if the leading bits of $\mathbf{v}$ and $\mathbf{w}$ are equal (say $v_{n-1} = w_{n-1} = 1$). The side conditions are equivalent to the single condition

$$(\mathbf{v}|\mathbf{w}) \ \oplus (\mathbf{a} \star (\mathbf{v} \oplus \mathbf{w} \oplus \mathbf{e})) \preceq L(\mathbf{a}) \tag{5}$$

Especially each solution $\mathbf{a}$ of (4) must fulfil :

$$\mathbf{v} \oplus \mathbf{w} \preceq L(\mathbf{a}); \quad \text{i.e.} \ \ R(\mathbf{v} \oplus \mathbf{w}) \preceq \mathbf{a}$$

Additionally we must have: if $v_i = w_i = 0$ and $a_i = 1$ then it must be that $a_{i+1} = 1$. If a section 10^k in $\mathbf{v}|\mathbf{w}$ corresponds to a section 0^{k+1} in $\mathbf{w} \star \mathbf{v}$, then the corresponding section in $L(\mathbf{a})$ must (for each solution $\mathbf{a}$ of (5) be 1^{k+1}. Let $\mathbf{m}(\mathbf{v}, \mathbf{w})$ denote the vector which is constructed by these modifications starting from $R(\mathbf{v} \oplus \mathbf{w})$. (A mathematical precise definition is given in Lemma B 1 in Appendix B.)

Theorem 7 (column maxima) *Let* $\mathbf{v}, \mathbf{w} \in \mathbb{F}_2^n$ *be given s.th.* $v_{n-1} = w_{n-1} = 1$, *let* $\mathbf{a}(\mathbf{v}, \mathbf{w}) := \mathbf{a}_L(\mathbf{v} \star \mathbf{w})|\mathbf{m}(\mathbf{v}, \mathbf{w})$ *and let* $\mathbf{u}(\mathbf{v}, \mathbf{w}) := (\mathbf{v} \oplus \mathbf{w}) \oplus (I \oplus L)(\mathbf{a}(\mathbf{v}, \mathbf{w}))$. *Then*

1. $\mathbf{a}(\mathbf{v}, \mathbf{w})$ *is a solution of (4) with minimal length and minimal* $|L(\mathbf{a})|$
2. $|\phi_2(\mathbf{u}(\mathbf{v}, \mathbf{w}), \mathbf{v}, \mathbf{w})| = \max\{|\phi_2(\mathbf{u}, \mathbf{v}, \mathbf{w})| \ : \ \mathbf{u} \in \mathbb{F}_2^n \}$

Proof Since $\mathbf{e}$ is a solution of (4) the set of solutions is not empty. (1): Let $\mathbf{a}$ be a solution of (4). We know from the discussion above that we must have $\mathbf{m}(\mathbf{v}, \mathbf{w}) \preceq \mathbf{a}$. Let $\mathbf{b}$ be a zero-block in $\mathbf{m}(\mathbf{v}, \mathbf{w})$. Then $\mathbf{v}$ and $\mathbf{w}$ differ at most on the leading bit of $\mathbf{b}$, and the corresponding block of $\mathbf{a}$ can be optimised independently of the other parts of $\mathbf{a}$. It is then clear (by Theorem 5) that $\mathbf{a}_L(\mathbf{v} \star \mathbf{w})$ is optimal on these sections and (1) follows. (2) is now easy. □

Example 2 Let $n = 11$ and $\mathbf{v} = (1101\,0001\,011) = (1675)_2, \mathbf{w} = (0100\,0101\,111) = (1954)_2$, (little-endian!). Then $R(\mathbf{v} \oplus \mathbf{w}) = (0100\,1010\,010)$, $\mathbf{m}(\mathbf{v}, \mathbf{w}) = (0100\,1111\,010)$ and $\mathbf{a}_L(\mathbf{v} \star \mathbf{w}) = (0100\,0001\,001)$. Finally $\mathbf{a}(\mathbf{v}, \mathbf{w}) = (0100\,1111\,011), \mathbf{u}(\mathbf{v}, \mathbf{w}) = (0100\,0101\,001) = (1186)_2$ (little-endian) and the weight of $(\mathbf{u}(\mathbf{v}, \mathbf{w}) \oplus \mathbf{w}) \star (\mathbf{u}(\mathbf{v}, \mathbf{w}) \oplus \mathbf{v})$ is odd. Hence in column (1675, 1954) of the correlation matrix the element $\phi_2(1186, 1675, 1954) = -2^{-7}$ is of maximal absolute value.

7 Summary

We have shown that addition mod 2^n is CCZ-equivalent to a quadratic vectorial Boolean function. This fact is not only theoretically interesting (it makes e.g. clear why only powers of 2 appear as non-zero absolute values of Walsh-coefficients/ differential probabilities of

addition), it also leads to practical results: it makes the solution of differential equations of addition extremely easy (Sect. 4.1, improving on [8]), and it leads to advanced results on the Walsh transform of addition mod 2^n (identifying for the first time "row"- resp. "column"-maxima and making the formula of [12] explicit).

On the practical side the results are helpful for the analysis of cryptographic primitives which use addition mod 2^n. (E.g. since the correlation matrix F_r of the $MIX(r)$ operation in Threefish is $F_r(\mathbf{t}_1, \mathbf{t}_2; \mathbf{u}_1, \mathbf{u}_2) = \phi_2(\mathbf{t}_1 \oplus \mathbf{t}_2; \mathbf{u}_1, \mathbf{u}_2 \oplus (\mathbf{t}_2 \ggg r))$ the results of Sect. 5 can be applied almost directly in the linear cryptanalysis of Threefish).

On the theoretical side the results lead to a better understanding of the cryptographic properties of addition mod 2^n. They show that (in the sense of CCZ-equivalence) addition mod 2^n is a very simple vectorial Boolean function.

Acknowledgments I would like to thank the anonymous reviewers of WCC 2011 and of this journal for valuable suggestions.

Appendix A: On L-minimal vectors

In this Appendix we prove Theorem 5. Recall that L denotes the left shift, that

$$\mathscr{A}_L(\mathbf{u}) := \{\mathbf{x} \in \mathbb{F}_2^n \ : \ \mathbf{x} \oplus \mathbf{u} \preceq L(\mathbf{x})\}$$

and that we want to find $\mathbf{a} = \mathbf{a}(\mathbf{u}) \in \mathscr{A}_L(\mathbf{u})$ s.th. the Hamming weight $|L(\mathbf{a})|$ is minimal.

In order to have a convenient way to discuss this problem we introduce some definitions and notations.

A 1 Definition *Let* $\mathbf{u} \in \mathbb{F}_2^n$.

1. $\mathbf{x} \in \mathbb{F}_2^n$ *is called "L-admissible" for* $\mathbf{u}$ *iff* $\mathbf{x} \in \mathscr{A}_L(\mathbf{u})$
2. $d(\mathbf{u}) := \min\{|L(\mathbf{x})| \ : \ \mathbf{x} \in \mathscr{A}_L(\mathbf{u})\}$
3. $\mathbf{x} \in \mathscr{A}_L(\mathbf{u})$ *s.th.* $|L(\mathbf{x})| = d(\mathbf{u})$ *is called "L-minimal".*
4. $\mathbf{x} \in \mathscr{A}_L(\mathbf{u})$ *s.th.* $|\mathbf{x}| = \min\{|\mathbf{y}| \ : \ \mathbf{y} \in \mathcal{A}(\mathbf{u})\}$ *is called "shortest L-admissible".*
5. $\mathbf{a}_L(\mathbf{u}) := \mathbf{u} \oplus (\mathbf{u} \star L(\mathbf{u})) \oplus \ldots \oplus (\mathbf{u} \star L(\mathbf{u}) \star L^2(\mathbf{u}) \cdots \star L^{n-1}(\mathbf{u}))$
6. *for* $k \in \{1, \ldots, n-1\}$ *let* $\ell^k(\mathbf{u}) := (u_0, \ldots, u_{k-1})$ *denote the k lowest bits of* $\mathbf{u}$ *and* $\mathbf{h}^k(\mathbf{u}) := (u_{n-k}, \ldots, u_{n-1})$ *denote the k highest bits of* $\mathbf{u}$. *Further* $\ell^n(\mathbf{u}) = \mathbf{h}^0(\mathbf{u}) := \mathbf{u}$ *and* $u_n := 0$.

A descriptive construction of $\mathbf{a}_L(\mathbf{u})$ is as follows:

$\mathbf{a}_L(\mathbf{u})$ is computed from $\mathbf{u}$ by replacing in the 1-blocks of $\mathbf{u}$ every second "1" by a "0" (counting down from $n-1$), leaving the zeroes unchanged. (Thus the computation of $\mathbf{a}_L(\mathbf{u})$ starts anew after each zero-bit in $\mathbf{u}$).

We remark that $\mathbf{a}_L(\mathbf{u})$ is the unique element in $\mathscr{A}_L(\mathbf{u})$ without adjacent ones.

Remark 1 Let $\mathbf{u} \in \mathbb{F}_2^n$.

Then $\mathbf{a}_L(\mathbf{u})$ is the only element $\mathbf{a} \in \mathscr{A}_L(\mathbf{u})$ such that $\mathbf{a} \star L(\mathbf{a}) = \mathbf{0}$.

We omit the simple proof. Intuitively $\mathbf{a}_L(\mathbf{u})$ should be a L-minimal vector for $\mathbf{u}$. We want to prove that this intuition is true. In general there will be several L-minimal vectors for a given $\mathbf{u}$.

Example 3 Let $n = 6$ and let $\mathbf{u} = (111101) = (47)_2$ (little-endian!). Then each of $\{010101, 110101, 101101, 101011\}$ is L-minimal for $\mathbf{u}$. Here $\mathbf{a}_L(\mathbf{u}) = (010101)$

A 2 Lemma *Let* $\mathbf{u} \in \mathbb{F}_2^n$

1. *if* $\mathbf{x} \in \mathbb{F}_2^n$ *is L-admissible for* $\mathbf{u}$*, then for each* $k \in \{0, \ldots, n-1\}$ *the vector* $\mathbf{h}^k(\mathbf{x})$ *is L-admissible for* $\mathbf{h}^k(\mathbf{u})$
2. *if* $\mathbf{x} \in \mathbb{F}_2^n$ *is L-admissible for* $\mathbf{u}$ *and if* $x_k = 0$ *for a* $k \in \{1, \ldots, n\}$*, then* $\ell^k(\mathbf{x})$ *is L-admissible for* $\ell^k(\mathbf{u})$
3. *if* $\mathbf{x}$ *is L-minimal (resp. shortest L-admissible) for* $\mathbf{u}$ *and* $x_k = 0$ *for a* $k \in \{1, \ldots, n-1\}$ *then* $\ell^k(\mathbf{x})$ *is L-minimal (resp. shortest L-admissible) for* $\ell^k(\mathbf{u})$*, und* $\mathbf{h}^{n-k}(\mathbf{x})$ *is shortest L-admissible for* $\mathbf{h}^{n-k}(\mathbf{u})$*. Furher we have*
   (a) *if* $\mathbf{w}^{n-k}$ *is another shortest L-admissible vector for* $\mathbf{h}^{n-k}(\mathbf{u})$*, then* $(\ell^k(\mathbf{x}), \mathbf{w}^{n-k})$ *is L-minimal (shortest L-admissible) for* $\mathbf{u}$
   (b) *if* $\mathbf{w}^k$ *is another L-minimal (resp. shortest L-admissible) vector for* $\mathbf{l}^k(\mathbf{u})$*, then* $(\mathbf{w}^k, \mathbf{h}^{n-k}(\mathbf{x}))$ *is L-minimal (shortest L-admissible) for* $\mathbf{u}$
4. $\mathbf{x} \in \mathbb{F}_2^n$ *is L-minimal (shortest L-admissible) for* $\mathbf{u}$*, iff* $(\mathbf{x}, 0) \in \mathbb{F}_2^{n+1}$ *is L-minimal (shortest L-admissible) for* $(\mathbf{u}, 0) \in \mathbb{F}_2^{n+1}$
5. $\mathbf{a}_L(\mathbf{e}) = (\ldots, 0, 1, 0, 1)$ *is L-minimal and shortest L-admissible for* $\mathbf{e}$.
6. $\mathbf{a}_L(\mathbf{u})$ *is L-admissible for* $\mathbf{u}$

Proof (1), (2) and (4) follow immediately from the definition of "L-admissible".

(3): let $\mathbf{x}$ be L-minimal (resp. shortest L-admissible) for $\mathbf{u}$ with $x_k = 0$. From (1) and (2) we know that $\ell^k(\mathbf{x})$ resp. $\mathbf{h}^{n-k}(\mathbf{x})$ are then L-admissible for $\ell^k(\mathbf{u})$ bzw. $\mathbf{h}^{n-k}(\mathbf{u})$. If one or both parts are replaced by other L-admissible parts the resulting vector will remain L-admissible for $\mathbf{u}$. Therefore both parts must have the stated optimality properties (else the resulting vector could be shortened by using a shorter part). (a) and (b) follow, since the newly constructed vectors (resp. their left shift) have the same length as the orignal $\mathbf{x}$ (resp. $L(\mathbf{x})$)

(6): for $\mathbf{a} := \mathbf{a}_L(\mathbf{u})$ it holds that $\mathbf{a} \oplus \mathbf{u} = \mathbf{u} \star L(\mathbf{a})$. Especially thus: $\mathbf{a}_L(\mathbf{u}) \oplus \mathbf{u} \preceq L(\mathbf{a}_L(\mathbf{u}))$.

(5): since $\mathbf{a} \oplus \mathbf{u} \preceq L(\mathbf{a})$ for each L-admissible $\mathbf{a}$ the inequality $|\mathbf{a}| + |\mathbf{u}| - 2|\mathbf{u} \star \mathbf{a}| \leq |\mathbf{a}|$ holds, i.e. $|\mathbf{a} \star \mathbf{u}| \geq |\mathbf{u}|/2$. If $\mathbf{a}$ is L-admissible for $\mathbf{e}$ we must therefore have that $|\mathbf{a}| \geq n/2$, i.e $|\mathbf{a}| \geq \lceil n/2 \rceil$ and $|L\mathbf{a}| \geq \lfloor n/2 \rfloor$. Since for $\mathbf{a}_L(\mathbf{e})$ equality holds in both cases, both minimality properties follow. □

We now prove the main result of this Appendix.

A 3 Theorem *Let* $\mathbf{u} \in \mathbb{F}_2^n$ *. Then*

1. $\mathbf{a}_L(\mathbf{u})$ *is a L-minimal vector for* $\mathbf{u}$
2. $\mathbf{a}_L(\mathbf{u})$ *is shortest L-admissible for* $\mathbf{u}$.

Proof If $\mathbf{u} = 0$ the assertions are true. Let $\mathbf{u} \neq 0$ and let (by Lemma A 2,4.) w.l.o.g. $u_{n-1} = 1$.

We use induction on n. For $n = 1, 2, 3$ the assertions are true (just check all cases). Let $n \geq 4$ and $\mathbf{w}$ be L-minimal (resp. shortest L-admissible) for $\mathbf{u}$. Then there is a $k \in \{1, \ldots, n-2\}$ s.th. $w_k = 0$ (else we would have $|L(\mathbf{w})| = n-1$ but $|L(\mathbf{a}_L(\mathbf{u}))| < n-1$ (resp. $|\mathbf{w}| \geq n-1$ but $|\mathbf{a}_L(\mathbf{u})| \leq n-2$).) The vector $\mathbf{v} := (\ell^k(\mathbf{w}), \mathbf{h}^{n-k}(\mathbf{a}_L(\mathbf{u}))$ is by induction hypothesis and (Lemma A 2,3.(a)) L-minimal (resp. shortest L-admissible).

Since $v_{n-2} = 0$ then (by Lemma A 2, 3.)) $\ell_{n-2}(\mathbf{v})$ is minimal (resp.shortest L-admissible) for $\ell_{n-2}(\mathbf{u})$. Again by induction hypothesis (and Lemma A 2, 3. (b)) then the vector $(\mathbf{a}_L(\ell^{n-2}(\mathbf{u})), \mathbf{h}^2(\mathbf{v})) = \mathbf{a}_L(\mathbf{u})$ is L-minimal (resp. shortest L-admissible). □

Finally we remark without proof that all other (if any) L-minimal vectors for $\mathbf{u}$ can be found by repeatedly applying the following construction (from top (right) to bottom):

Remark 2 Let $\mathbf{a} = \mathbf{a}_L(\mathbf{u})$ and $\mathbf{b}$ be L-minimal for $\mathbf{u}$. If $a_j = 0, a_{j+1} = 1$ and $u_{j-1} = 1$ is the end of a one-block $(u_{j-\ell}, \ldots, u_{j-1})$ of $\mathbf{u}$ of odd (even) length ℓ and if $(b_{j-\ell}, \ldots, b_{j+1}) = (a_{j-\ell}, \ldots, a_{j+1}))$ the vector obtained from $\mathbf{b}$ by complementing $(b_{j-\ell}, \ldots, b_j)$ is (nearly) L-minimal.

Remark 3 Since $x \in \mathscr{A}_L(\mathbf{u}) \Leftrightarrow (\mathbf{x} \oplus \mathbf{u}) \star L(\mathbf{x}) = \mathbf{x} \oplus \mathbf{u}$ it holds that $\mathbf{a}_L(\mathbf{u})$ is a shortest vector in a set described by a special set of quadratic equations. Finding a shortest vector in a set described by a general system of quadratic equations is extremely hard.

In the same vein as above one can study the problem: given $\mathbf{u} \in \mathbb{F}_2^n$, let

$$\mathscr{A}_R(\mathbf{u}) := \{\mathbf{x} \in \mathbb{F}_2^n \ : \ \mathbf{x} \oplus \mathbf{u} \preceq R(\mathbf{x})\}$$

Problem: find a vector $\mathbf{a} = \mathbf{a}(\mathbf{u}) \in \mathscr{A}_R(\mathbf{u})$ s.th. the Hamming weight $|R(\mathbf{a})|$ is minimal.

Define $\mathbf{a}_R(\mathbf{u}) := \mathbf{u} \oplus (\mathbf{u} \star R(\mathbf{u})) \oplus \ldots \oplus (\mathbf{u} \star R(\mathbf{u}) \star R^2(\mathbf{u}) \star \cdots \star R^{n-1}(\mathbf{u}))$ and define "R-minimal" and "$R-$admissible" in analogy to the above. In a completely similar way one then can show:

A 4 Theorem *Let* $\mathbf{u} \in \mathbb{F}_2^n$. *Then*

1. $\mathbf{a}_R(\mathbf{u})$ *is a* R*-minimal vector for* $\mathbf{u}$
2. $\mathbf{a}_R(\mathbf{u})$ *is shortest* R*-admissible for* $\mathbf{u}$.

The functions $\mathbf{a}_L(\mathbf{u}), \mathbf{a}_R(\mathbf{u})$ were introduced in [6] under the name of "all one parities".

Appendix B: Efficient computation

One of the main objectives of [12,6] are efficient algorithms for the linear correlations resp. the differential probabilities of addition mod 2^n . It is the purpose of this Appendix show that all the functions above have fast algorithms and to provide an understanding why this is so. (Here a fast algorithm is one that uses of order log of wordlength operations on words. Note however that we don't give a full explanation in the sense of informatics, for which also RAM-models etc. would have to be specified.) It is apparent that it suffices to devise fast algorithms for the vectorial functions $\mathbf{z}$, $\mathbf{m}$, $\mathbf{a}_L$, $\mathbf{a}_R$ and $\mathbf{q}$. In fact, all these functions are of the same general type.

Let for $\mathbf{t}, \mathbf{x} \in \mathbb{F}_2^n$

$$\mathbf{u}_R(\mathbf{x}, \mathbf{t}) := (I \oplus XR)^{-1}(\mathbf{t}) \quad \text{resp.} \quad \mathbf{u}_L(\mathbf{x}, \mathbf{t}) := (I \oplus XL)^{-1}(\mathbf{t})$$

where X is the diagonal matrix $X = diag(\mathbf{x})$ with diagonal elements $X_{i,i} = x_i$. Clearly XR resp. XL are lower resp. upper triangular matrices with elements $x_1, \ldots, x_{n-1}$ resp. $x_0, \ldots, x_{n-2}$ in the first lower resp.upper secondary diagonal (and all other elements are 0). Especially $(XL)^n = (XR)^n = \mathbf{0}$. If $k = \lceil \log_2(n) \rceil$ then

$$(I \oplus XR)^{-1} = I \oplus (XR) \oplus \ldots \oplus (XR)^{n-1} = (I \oplus (XR)^{2^k})(I \oplus (XR)^{2^{k-1}}) \ldots (I \oplus XR)$$

$$(I \oplus XL)^{-1} = I \oplus (XL) \oplus \ldots \oplus (XL)^{n-1} = (I \oplus (XL)^{2^k})(I \oplus (XL)^{2^{k-1}}) \ldots (I \oplus XL)$$

Further $(XR)^m$ resp. $(XL)^m$ can have non-zero entries (if any) only in the m-th lower resp.upper secondary diagonal, and $(XR)^m_{m+j,j} = \prod_{i=j+1}^{m+j} x_i$ as well as $(XL)^m_{j,m+j} = \prod_{i=j}^{m+j-1} x_i$ are each product of resp. m consecutive entries of $\mathbf{x}$.

It is then clear that $\mathbf{u}_R$ resp. $\mathbf{u}_L$ can be computed in time k by repeated shifts and XORs. For the explanation we have kept the R (resp. L) notation from above, but it is clear that in the "big endian" format these are left (resp. right) shifts. In C-code (here for word=unsigned int) they are e.g. given by

```
unsigned int uR(unsigned int t, unsigned int x){
int j=1;
  unsigned int y,s;
    s=t;
    y=x;                          /*  y = x  */
   while (y!=0){
       s^=((y&(s<<j));          /*  s = s ⊕ (y · R^{2^{j-1}}(s))  */
       y&=(y<<j);               /*  y = y · R^{2^{j-1}}(y)  */
       j<<=1;}
   return s;
}
```

and similarly

```
unsigned int uL(unsigned int t, unsigned int x){
int j=1;
unsigned int y,s;
    s=t;
    y=x;                          /*  y = x  */
   while (y!=0){
       s^=((y&(s>>j));          /*  s = s ⊕ (y · L^{2^{j-1}}(s))  */
       y&=(y>>j);               /*  y = y · L^{2^{j-1}}(y)  */
       j<<=1;}
   return s;
}
```

If $\mathbf{t} \star \mathbf{x} = \mathbf{0}$ one can replace $\oplus$ (XOR) here with | (OR).
Finally, it is a routine matter to show

B 1 Lemma *Let* $\mathbf{u}, \mathbf{w}, \mathbf{v} \in \mathbb{F}_2^n$ *and let* $\mathbf{z}, \mathbf{m}, \mathbf{a}_L, \mathbf{a}_R$ *be as above. Then*

$$\begin{aligned}
\mathbf{z}(\mathbf{u}, \mathbf{w}, \mathbf{v}) &= \mathbf{u}_L(\mathbf{e}, L(\mathbf{u} \oplus \mathbf{w} \oplus \mathbf{v})) \\
\mathbf{a}_L(\mathbf{u}) &= \mathbf{u}_L(\mathbf{u}, \mathbf{u}) \\
\mathbf{q}(\mathbf{u}) &= \mathbf{u}_L(\mathbf{e}, L(\mathbf{u})) \\
\mathbf{a}_R(\mathbf{u}) &= \mathbf{u}_R(\mathbf{u}, \mathbf{u}) \\
\mathbf{m}(\mathbf{v}, \mathbf{w}) &= \mathbf{u}_R(R((\mathbf{v} \oplus \mathbf{e}) \star (\mathbf{w} \oplus \mathbf{e})), R(\mathbf{v} \oplus \mathbf{w}))
\end{aligned}$$

Thus all of the considered vectorial functions can be computed in log(wordlength) operations on words. A more complicated log(wordlength)-time algorithm for $\mathbf{z}$ was already given by Wallén ([12], Theorem 2). A different log(wordlength)-time algorithm for $\mathbf{a}_R$ was already given by Lipmaa and Moriai ([6], Theorem 2).

References

1. Alquié D.: Approximating Addition by XOR: How to Go All the Way. Tech. Rep. 072/2010, Cryptology ePrint Archive (2010). Available at http://eprint.iacr.org/2010/072.
2. Biham E., Shamir A.: Differential cryptanalysis of FEAL and n-Hash. In: Advances in Cryptology—EUROCRYPT 1991, no. 547 in Lecture Notes in Computer Science, pp. 1–16. Springer, Berlin (1991).

3. Carlet C., Charpin P., Zinoviev V.: Codes, bent functions and permutations suitable for DES-like crypto systems. Des. Codes Cryptogr. **15**(2), 125–156 (1998).
4. Holte J.: Carries, combinatorics and an amazing matrix. Am. Math. Mon. **104**(2), 138–149 (1997).
5. Leurent G., Thomsen S.: Practical partial collisions on the compression function of BMW. In: Fast Software Encryption 2011, no. 6733 in Lecture Notes in Computer Science. Springer, Berlin (2011).
6. Lipmaa H., Moriai S.: Efficient algorithms for computing differential properties of addition. In: Fast Software Encryption 2001, no. 2355 in Lecture Notes in Computer Science, pp. 336–350. Springer, Berlin (2002).
7. Nyberg C., Wallén J.: Improved linear distinguishers for SNOW 2.0. In: Fast Software Encryption 2006, no. 4047 in Lecture Notes in Computer Science, pp. 336–350. Springer, Berlin (2006).
8. Paul S., Preneel B.: Solving systems of differential equations of addition. In: ACISP 2005, no. 3574 in Lecture Notes in Computer Science, pp. 75–88. Springer, Berlin (2006). Extended Version available as Technical Report 294/2004 at http://eprint.iacr.org/2004/294.
9. Rueppel R.A.: Correlation immunity and the summation generator. In: Advances in Cryptology—CRYPT0 '85, no. 218 in Lecture Notes in Computer Science, pp. 260–272. Springer, Berlin (1986).
10. Sarkar P.: On Approximating addition by exclusive Or. Tech. Rep. 047/2009, Cryptology ePrint Archive (2009). Available at http://eprint.iacr.org/2009/047.
11. Staffelbach O., Meier W.: Cryptographic significance of the carry for ciphers based on integer addition. In: Advances in Cryptology—CRYPT0 '90, no. 537 in Lecture Notes in Computer Science, pp. 601–614. Springer, Berlin (1990).
12. Wallén J.: Linear approximations of addition mod 2^n. In: Fast Software Encryption 2003, no. 2887 in Lecture Notes in Computer Science, pp. 261–273. Springer, Berlin (2003).