Mathematical aspects of the design and security of block ciphers

---

DISSERTATION

zur

Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität Rostock

vorgelegt von:
Lukas KÖLSCH, geb. am 19.05.1993 in Gifhorn
aus Rostock

Gutacher:

Prof.in Gohar Kyureghyan, Universität Rostock, Institut für Mathematik
Prof. Petr Lisoněk, Simon Fraser University Vancouver, Department of Mathematics
Prof. Ferruh Özbudak, Middle East Technical University Ankara, Department of Mathematics

Jahr der Einreichung: 2020
Jahr der Verteidigung: 2020

https://doi.org/10.18453/rosdok_id00002900

Selbstständigkeitserklärung

Hiermit versichere ich eidesstattlich, dass ich die vorliegende Arbeit mit dem Titel "Mathematical aspects of the design and security of block ciphers" selbstständig und ohne unerlaubte frende Hilfe angefertigt, keine anderen als die angegebenen Quellen und Hilfsmittel verwendet und die den verwendeten Quellen und Hilfsmitteln wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Ort, Datum:

Unterschrift:

Abstract

Block ciphers constitute a major part of modern symmetric cryptography. A mathematical analysis is necessary to ensure the security of the cipher. In this thesis, we develop several new contributions for the analysis of block ciphers. We determine cryptographic properties of several special cryptographically interesting mappings like almost perfect nonlinear functions. To do so, we consider the problem of inversion in $\mathbb{Z}_{2^n-1}$, and classify certain permutation polynomials over binary finite fields. We also give some new results both on the resistance of functions against differential-linear attacks as well as on the efficiency of implementation of certain block ciphers.

Acknowledgments

Firstly, and most importantly, I thank my supervisor Gohar Kyureghyan. Without you, I probably would not have started a career in academia. I thank you for always having an open ear and for creating a work atmosphere that was focused while also leaving me enough space and freedom to pursue some more extravagant ideas.

Parts of this thesis were written in collaboration with others: For that I want to thank my co-authors Anne Canteaut, Friedrich Wiemer, Faruk Göloğlu, Gohar Kyureghyan and Léo Perrin.

Lastly, I want to thank all of my friends and my family for all their support in many different forms in these last three years - and for reminding me time and time again that mathematics, contrary to popular opinion, is not the most important thing in life.

Contents

Selbstständigkeitserklärung **iii**

Abstract **v**

Acknowledgments **vii**

1 Introduction **1**

2 Preliminaries **5**
2.1 Notation and Definitions . . . 5
2.2 Cryptographic properties of vectorial Boolean functions . . . 7
2.3 Equivalences of vectorial Boolean functions . . . 11
2.4 APN functions . . . 12
2.5 A short introduction to characters . . . 14
2.5.1 Stickelberger's congruence . . . 17

3 Inversion in $\mathbb{Z}_{2^n-1}$ **21**
3.1 A new method for inversion in $\mathbb{Z}_{2^n-1}$: The modular add-with-carry approach . . . 21
3.2 The Gold exponents . . . 23
3.2.1 The APN Gold exponents . . . 23
3.2.2 The non-APN Gold exponents . . . 25
3.3 The Kasami exponents . . . 27
3.3.1 The case $\gcd(r,n)=1$ . . . 28
3.3.2 The case $\frac{n}{\gcd(n,r)}$ odd . . . 32
3.3.3 The case $\frac{n}{\gcd(n,r)}$ even . . . 38
3.3.4 Kasami inverses with special structure . . . 39
3.4 The Bracken-Leander exponent . . . 41
3.5 Conclusion . . . 43

4 Equivalences of monomials and permutations of the form $L_1(x^d)+L_2(x)$ **45**
4.1 A connection between equivalences of vectorial Boolean functions and permutation polynomials of the form $L_1(x^d)+L_2(x)$ . . . 45
4.2 The inverse function . . . 48
4.2.1 Vector spaces of Kloosterman zeros . . . 56
4.3 Other monomials . . . 60
4.4 Conclusion . . . 69

5 Autocorrelation of vectorial Boolean functions **71**
5.1 The DLCT and the autocorrelation . . . 72
5.1.1 Some characterizations and properties of the autocorrelation . . . 73
5.1.2 Bounds on the absolute indicator . . . 75
5.1.3 Invariance of the autocorrelation under Equivalence Relations . . . 77

5.1.4 Divisibility properties of the autocorrelation 78
5.1.5 Autocorrelation of APN functions 80
5.2 Autocorrelation spectra and absolute indicator of special polynomials 82
5.2.1 Monomials 82
5.2.2 Quadratic functions and their inverses 87
5.3 Outlook 89

6 XOR-counts and lightweight multiplication in binary finite fields 91
6.1 Introduction 91
6.2 XOR-Counts 93
6.3 Efficient Multiplication Matrices in Finite Fields 97
6.4 Quantifying the Gap between the Optimal Implementation and the Naive Implementation 105
6.5 Open Problems 106

A Appendix: Proofs of Section 6.4 107
A.1 Proof of Theorem 6.4.1 107
A.2 Proof of Propositions 6.4.2 and 6.4.3 112

Bibliography 115

Chapter 1

Introduction

With the growing importance of communication technologies in everyday life and their widespread use in the economy, secure communication is critical to protect the privacy of billions of people. Accordingly, there is a high demand for cryptographic algorithms and cryptanalysis. Modern cryptography is based on mathematics to describe and analyze these algorithms. Cryptography relies on a *key* which is information only known to a particular set of people which will allow them (and with a secure cryptographic algorithm *only* them) to decrypt the ciphertext. As usual, we will use a (hypothetical) communication between two parties called *Alice* and *Bob* to succinctly describe the process of encryption and decryption. Cryptosystems can be divided into two different types: Symmetric and asymmetric cryptosystems. The distinction is very simple: In symmetric cryptography, both Alice and Bob use the same (private) key for encryption and decryption while in asymmetric cryptography the encryption key is public and only the decryption key is private. The advantage of asymmetric cryptography is apparent. Alice and Bob do not have to agree on a key before the conversation. In fact, Bob can freely distribute his encryption key and use his (secret) decryption key for communication not only with Alice but with many different parties. In contrast, symmetric key cryptography relies on both Alice and Bob having agreed on a secret key that is known only to them, which leads to the important problem of key managing. However, the currently known and used asymmetric algorithms are significantly slower than symmetric algorithms. In practice, a combination of both approaches is common: Alice and Bob use an asymmetric, public-key algorithm to agree on a secret key, which is then subsequently used in a symmetric algorithm.

Symmetric cryptography can be further divided up into two different subfields, block ciphers and stream ciphers. A block cipher processes the data as a series of blocks of a previously determined, fixed size (often 64 or 128 bits), while stream ciphers encrypt and decrypt the plain text one bit at a time. A key idea in the design of block ciphers is to use a simple transformation called *round function* that operates on the block multiple times (see Figure 1.1 for a schematic example). Those round functions must of course depend on the secret key.

For block ciphers, several different construction techniques for round functions

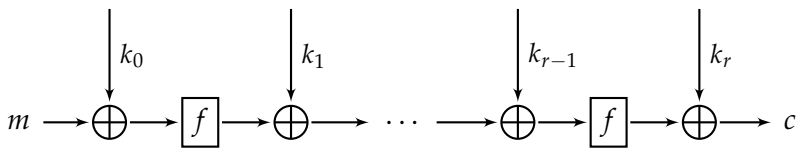

FIGURE 1.1: An iterated block cipher with r rounds and subkeys k_i that encrypts a plain text m into a ciphertext c

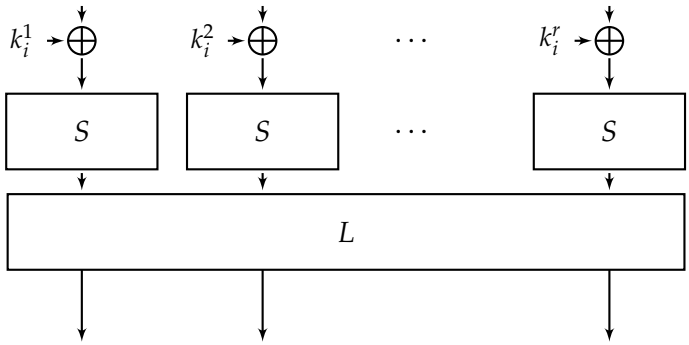

FIGURE 1.2: A high-level view of one round of an SPN with an S-box S, linear layer L and round keys k_i

have been used. The two most widely used constructions are *Substitution Permutation Networks (SPNs)* and *Feistel Networks*. The round function of an SPN is made up of a linear and a non-linear layer. The non-linear layer (the *substitution* part of the SPN) consists of multiple, usually identical permutations called *S-boxes* that work on small parts of the block in parallel. The linear layer then permutes these parts, as depicted in Figure 1.2. The design of an SPN thus comes mainly down to choosing an S-box and a linear layer. Both of these choices have an impact on the speed and the security of encryption and decryption.

The most famous SPN is the Advanced Encryption Standard (AES). Introduced more than 20 years ago, it is probably the block cipher that has been studied the most. It is still considered a secure cipher and remains in use for many different applications. Large parts of this thesis are concerned with contributing to the mathematical background for the design choice of the linear (Chapter 6) and non-linear layer (Chapter 3 to 5) of an SPN.

The thesis is structured as follows: In Chapter 2, we give an overview on the existing theory on vectorial Boolean functions as well as characters of finite fields to the extent that is needed in the subsequent chapters. This chapter does not include any original results.

Chapter 3 deals with the problem of inversion in the ring $\mathbb{Z}_{2^n-1}$. To this end, we introduce a new technique based on the modular add-with-carry approach. The main result of this chapter is the explicit determination of the inverses of all bijective Gold, Kasami and Bracken-Leander functions. With our contribution, all inverses of bijective APN monomials have been explicitly determined. As a corollary, we also obtain the algebraic degree of these inverses. This chapter is based on the paper [83], written by the author of this thesis.

Chapter 4 deals with EA and CCZ-equivalence of monomials in connection to permutation polynomials of the form $L_1(x^d) + L_2(x)$ over $\mathbb{F}_{2^n}$ where L_1, L_2 are $\mathbb{F}_2$-linear functions. The main result in this chapter is a complete characterization of those permutations for the inverse exponent $d = 2^n - 2$ and the Dillon-exponent $d = 2^{n/2} - 1$. As a result, we prove that if $n \geq 5$ all functions that are CCZ-equivalent to the inverse function are already EA-equivalent to it; and that all permutations that are CCZ-equivalent to the inverse function are affine equivalent to it (Theorem 4.2.18). A similar result is achieved for the function $x \mapsto x^{2^{n/2}-1}$ for n even (Theorem 4.3.15). This chapter is based on two papers [57, 81], one of them co-written and the other written exclusively by the author of this thesis, but also contains some significant original results that are as of yet unpublished.

Chapter 5 deals with the autocorrelation of vectorial Boolean functions. We show that the autocorrelation is intimately linked to the Differential-Linear Connectivity Table (DLCT) which is a measure of the resistance of an S-box of a block cipher against differential-linear attacks. We prove several new properties of the autocorrelation for vectorial Boolean functions, including lower bounds, divisibility results and connections to linear and differential properties. We also explicitly derive specific results for special vectorial Boolean functions, including cubic functions, monomials, and inverses of quadratic functions. This chapter is based on [25], co-written by the author of this thesis.

Chapter 6 deals with optimal implementation of linear layers of block ciphers. We discuss and prove some new results about different metrics to count the number of XORs needed in the implementation of certain linear layers. One of the main results of the chapter is a complete classification of all multiplication matrices that can be implemented with exactly 2 XOR-operations (Theorem 6.3.7). We also discuss more generally the efficiency of the "naive" implementation in contrast to more efficient implementations and are able to precisely quantify the gap between naive and efficient implementations (Theorem 6.4.1, Propositions 6.4.2 and 6.4.3). This chapter is based on [84], written by the author of this thesis.

Chapters 3 to 6 are written to be as self-contained as possible and can all be read independently of each other.

Chapter 2

Preliminaries

Large parts of this thesis will be concerned with *vectorial Boolean functions*. A vectorial Boolean function can be defined as a mapping between two finite fields of characteristic 2 (more precise definitions will be given later). If a vectorial Boolean function maps to the field with two elements $\mathbb{F}_2$, it is just called a *Boolean function*. Vectorial Boolean functions play an essential role in modern cryptography. Most importantly, an S-box of a block cipher can always be considered as a vectorial Boolean function. Additionally, Boolean functions are crucial building blocks in stream ciphers. To ensure that the cryptographic principles of *confusion* and *diffusion* developed by Shannon [122] are fulfilled, the vectorial Boolean functions chosen for use in a cryptographic scheme have to satisfy certain properties. Consequently, functions that behave optimally with respect to different cryptographic criteria have been the subject of much research. Generally, it is not possible to find a vectorial Boolean function that is optimal with respect to all cryptographic criteria and, in practice, a compromise between these criteria (and implementation efficiency) often has to be found. In the following, we will describe the most important cryptographic properties of vectorial Boolean functions as well as summarize some fundamental results.

2.1 Notation and Definitions

We briefly introduce the notation and basic definitions that will be used throughout the thesis.

We denote by $\mathbb{F}_q$ the field with q elements and by $\mathbb{Z}_n$ the ring of integers modulo n. It is well known that the field with $q = p^n$ elements for a prime p is unique up to isomorphism. The multiplicative group of $\mathbb{F}_q$ will be written as $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$. We will denote by $\mathbb{F}_q^n$ the set of all n-dimensional vectors with entries in $\mathbb{F}_q$. The field $\mathbb{F}_{q^n}$ can also be considered as an n-dimensional vector space over $\mathbb{F}_q$. Indeed, an isomorphism φ between $\mathbb{F}_q^n$ and $\mathbb{F}_{q^n}$ (as vector spaces) can easily be constructed by fixing a basis $B = \{b_1, \ldots, b_n\}$ of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ and setting

$$\varphi\colon \mathbb{F}_q^n \to \mathbb{F}_{q^n}, \quad \varphi((u_1, \ldots, u_n)) = \sum_{i=1}^{n} u_i b_i. \tag{2.1}$$

We will denote by $\mathrm{Tr}_n\colon \mathbb{F}_{p^n} \to \mathbb{F}_p$ the absolute trace function, i.e.

$$\mathrm{Tr}_n(x) = x + x^p + x^{p^2} + \cdots + x^{p^{n-1}}.$$

If the value of n is clear from the context, we will just write Tr. Using the trace function, we can define the *trace bilinear form* on $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$ in the following way:

$$\langle a, b \rangle = \mathrm{Tr}(ab)$$

for all $a, b \in \mathbb{F}_{p^n}$.

Moreover, we denote by $\mathrm{N}_{q^n/q} \colon \mathbb{F}_{q^n}^* \to \mathbb{F}_q^*$ the *norm function*, defined by

$$\mathrm{N}_{q^n/q}(a) = a^{q^0} \cdot a^{q^1} \cdot \dots \cdot a^{q^{n-1}} = a^{(q^n-1)/(q-1)}$$

for all $a \in \mathbb{F}_{q^n}^*$. In this thesis, we mostly deal with binary finite fields, i.e. extension fields of the field with two elements $\mathbb{F}_2 = \{0, 1\}$, and mappings between those fields.

Definition 2.1.1. *A vectorial Boolean function is a function $F \colon \mathbb{F}_2^n \to \mathbb{F}_2^m$. We will also refer to it as an (n, m)-function. A Boolean function is an $(n, 1)$-function.*

We will denote vectorial Boolean functions with $m \geq 2$ always with uppercase letters $F, G, \dots$ and Boolean functions always with lowercase letters.

Definition 2.1.2 (Algebraic Normal Form)**.** *Let $f \colon \mathbb{F}_2^n \to \mathbb{F}_2$ be a Boolean function. Each Boolean function has a unique representation*

$$f(x_1, \dots, x_n) = \sum_{u \in \mathbb{F}_2^n} a_u \left(\prod_{j=1}^{n} x_j^{u_j} \right),$$

where $a_u \in \mathbb{F}_2$ and $u_j \in \{0, 1\}$. This representation is called the Algebraic Normal Form *(ANF) of f.*

A vectorial Boolean function $F \colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ can be written as m Boolean functions

$$F(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), f_2(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)).$$

The Boolean functions $f_1, \dots, f_m$ are called the coordinate functions *of F.*

For a bit string u we denote by $\mathrm{wt}(u)$ the number of ones of u.

Definition 2.1.3 (Algebraic Degree)**.** *Let f be a Boolean function given by its ANF. The* algebraic degree *of a Boolean function f is defined as the maximal value of $\mathrm{wt}((u_1, \dots, u_n))$ with $a_u \neq 0$.*

The algebraic degree of a vectorial Boolean function is defined as the maximum of the algebraic degrees of its coordinate functions.

A vectorial Boolean function F with algebraic degree 1 is called affine; if additionally $F(0) = 0$ holds, we call it linear. Functions with algebraic degree 2 are called quadratic and functions with algebraic degree 3 cubic.

From now on, we will always use the finite field view of vectorial Boolean functions, i.e. we identify $\mathbb{F}_2^n$ and $\mathbb{F}_2^m$ with $\mathbb{F}_{2^n}$ and $\mathbb{F}_{2^m}$ using the isomorphism in Eq.(2.1) with an arbitrary, fixed basis. The choice of the basis does not impact the mathematical properties of the vectorial Boolean function, in particular it will not impact the algebraic degree we just introduced. We want to note that the choice of basis is relevant when it comes to the implementation of the function, we will consider this problem in Chapter 6. Some of the topics in this thesis (especially in Chapter 5) can also be treated equivalently using the vector space view. For a more in-depth treatment of (vectorial) Boolean functions and their cryptographic properties using the vector space terminology, we refer the reader to [28, 30, 41].

Definition 2.1.4 (Component functions)**.** *Let F be an (n,m)-function. The 2^m-1 component functions of F are the Boolean functions $F_b\colon \mathbb{F}_{2^n} \mapsto \mathbb{F}_2$ defined by $x \mapsto \mathrm{Tr}(bF(x))$ for $b \in \mathbb{F}_{2^m}^*$.*

Definition 2.1.5 (Balanced functions)**.** *An (n,m)-function F is called* balanced *if it attains every value from $\mathbb{F}_{2^m}$ the same number of times, i.e. if for all $v \in \mathbb{F}_{2^m}$*

$$|\{x \in \mathbb{F}_{2^n}\colon F(x) = v\}| = 2^{n-m}.$$

In particular, a Boolean function is balanced if it attains both values 0 and 1 exactly 2^{n-1} times.

(n,n)-functions are of special importance in many applications (e.g. the S-box of an SPN is always an (n,n)-function). Every (n,n)-function can be uniquely written as a univariate polynomial over $\mathbb{F}_{2^n}$ in the form:

$$F(x) = \sum_{i=0}^{2^n-1} a_i x^i \in \mathbb{F}_{2^n}[x].$$

2.2 Cryptographic properties of vectorial Boolean functions

Many different criteria exist to judge the resistance of vectorial Boolean functions to various attacks in cryptography. In this section we will give a brief overview over the most important of these criteria. An exceptionally bad choice for a cryptographic function is a linear function. Indeed, for a linear function, the entire function can be recovered just from the image of a basis. Many cryptographic properties can be seen as a measurement of nonlinearity.

A vectorial Boolean function used as an S-box in an SPN must be bijective to allow decryption. For this reason, bijective vectorial Boolean functions have been the subject of much research.

Definition 2.2.1 (Permutation polynomial)**.** *We call an (n,n)-function a* permutation *or* permutation polynomial *if it induces a bijective map on $\mathbb{F}_{2^n}$.*

Permutations can be seen as a special case of balanced functions for $n = m$ (see Definition 2.1.5). The question whether F is a permutation can be answered completely by considering its component functions.

Proposition 2.2.2 ([102, Theorem 7.7.])**.** *A function $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is a permutation if and only if all of its component functions are balanced.*

Every integer $0 \le k \le 2^n-1$ has a unique *binary expansion* $(k_{n-1},\dots,k_0) \in \{0,1\}^n$ such that $\sum_{i=0}^{n-1} 2^i k_i = k$. We denote by $\mathrm{wt}(k) = \sum_{i=0}^{n-1} k_i$ the *binary weight* of k. When we identify an element k of the set $\{0,1,\dots,2^n-1\}$ with the corresponding sequence in the set of binary sequences of length n of the form $(k_{n-1},\dots,k_0)$, we have $\mathrm{wt}(k) = \mathrm{wt}((k_{n-1},\dots,k_0))$. Similarly, we can identify $\{0,1,\dots,2^n-2\}$ with $\mathbb{Z}_{2^n-1}$, and we define the weight of an element $k \in \mathbb{Z}_{2^n-1}$ as the weight of its corresponding residue in $\{0,1,\dots,2^n-2\}$. With these notations in place, we will often not distinguish between $\{0,1,\dots,2^n-1\}$, $\mathbb{Z}_{2^n-1}$ and the set of binary sequences of length n. Here, we have to pay special attention to the sequence $(1,\dots,1)$, corresponding to the integer 2^n-1, which has weight n, but has no direct correspondent in $\mathbb{Z}_{2^n-1}$. In some cases, it makes sense to identify this element also with $0 \in \mathbb{Z}_{2^n-1}$ and assign it weight 0, mainly because the monomial $x \mapsto x^{2^n-1}$ is identical to $x \mapsto x^0$ on $\mathbb{F}_{2^n}^*$. We will always make it clear when this happens.

Lemma 2.2.3 ([30]). *Let $F(x)\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be an (n,n)-function with univariate representation*

$$F(x) = \sum_{i=0}^{2^n-1} a_i x^i \in \mathbb{F}_{2^n}[x].$$

The algebraic degree of F is the maximum binary weight of i such that $a_i \neq 0$, i.e. the algebraic degree of F is

$$\max_{0\leq i\leq 2^n-1\colon a_i\neq 0} \mathrm{wt}(i).$$

Clearly, the algebraic degree of an (n,n)-function is at most n.

In particular, linear functions are precisely functions where the weights of the exponents are 1, i.e. they have the univariate representation

$$F(x) = \sum_{i=0}^{n-1} a_i x^{2^i} \in \mathbb{F}_{2^n}[x].$$

The term *linear* is motivated by the fact that these functions are precisely the ones that are $\mathbb{F}_2$-linear, i.e. they satisfy the equation $L(x+y) = L(x) + L(y)$ for all $x, y \in \mathbb{F}_{2^n}$. Consequently, we can view linear functions as linear mappings on the vector space $\mathbb{F}_{2^n}$ over $\mathbb{F}_2$, and we may use concepts from linear algebra to investigate these functions.

Definition 2.2.4 (Adjoint polynomial). *Let $L\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be a linear function. We define by $L^*\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ its* adjoint polynomial. *It is the unique polynomial that satisfies*

$$\langle L(x), y\rangle = \langle x, L^*(y)\rangle$$

for all $x, y \in \mathbb{F}_{2^n}$. Here $\langle\cdot,\cdot\rangle$ denotes the trace bilinear form introduced earlier.

The adjoint polynomial of L can be explicitly determined from L. In fact, if $L = \sum_{i=0}^{n-1} a_i x^{2^i}$ then $L^* = \sum_{i=0}^{n-1} a_i^{2^{n-i}} x^{2^{n-i}}$. Note that L^* is also linear and $L^{**} = L$.

The algebraic degree can be seen as a measure of linearity: The lower the algebraic degree of a function, the better it can be approximated by a linear function. If a function with low algebraic degree is chosen as an S-box of a block cipher, it may be more vulnerable to a number of different attacks, for example higher-order differential attacks [82]. The following result follows directly from the well-known Hermite's criterion and establishes that there are no permutations that take on the maximal algebraic degree.

Proposition 2.2.5. *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be a permutation. Then its algebraic degree is at most $n-1$.*

Definition 2.2.6 (Walsh transform, extended Walsh spectrum). *Let F be an (n,m)-function. We define the* Walsh transform *of F as*

$$W_F(a,b) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}_m(bF(x))+\mathrm{Tr}_n(ax)}$$

for all $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^m}$. We define the extended Walsh spectrum *as the multiset*

$$\Lambda_F = \{*|W_F(a,b)|\colon a \in \mathbb{F}_{2^n}, b \in \mathbb{F}_{2^m}^* *\}.$$

For a Boolean function f the only nonzero choice for b is $b = 1$, so we write in this case also $W_f(a) = W_f(a,1)$. Clearly, $W_F(a,0) = 0$ for all $a \in \mathbb{F}_{2^n}^*$ and all possible

(n,m)-functions F. If F is a permutation, then we also have $W_F(0,b) = 0$ for all $b \in \mathbb{F}_{2^n}^*$.

Definition 2.2.7 (Linearity, Nonlinearity)**.** *Let F be an (n,m)-function. The* nonlinearity *of F is defined as*

$$NL(F) = 2^{n-1} - \frac{1}{2} \max_{a \in \mathbb{F}_{2^n}, b \in \mathbb{F}_{2^m}^*} |W_F(a,b)| = 2^{n-1} - \frac{1}{2} \max_{\lambda \in \Lambda_F} \lambda.$$

Similarly, the Linearity *of F is defined as*

$$L(F) = \max_{\lambda \in \Lambda_F} \lambda.$$

As the names indicate, both linearity and nonlinearity are also measures for the linearity of a vectorial Boolean function. It is easy to verify that the nonlinearity of a linear function is 0. Generally, the higher the nonlinearity (equivalently: the lower the linearity) of a function, the better is its resistance to linear attacks [108]. The nonlinearity is bounded from above by what is sometimes called the covering radius bound or universal bound (see e.g. [30]).

Proposition 2.2.8. *Let F be an (n,m)-function. Then*

$$NL(F) \leq 2^{n-1} - 2^{\frac{n}{2}-1}.$$

Functions that satisfy this bound with equality are called *vectorial bent functions*. Vectorial bent functions can of course only exist for even n. Additionally, we get a constraint on m:

Proposition 2.2.9 (Nyberg bound, [114])**.** *Let F be a vectorial bent (n,m)-function. Then $m \leq n/2$.*

Vectorial bent functions can be characterized in different ways:

Proposition 2.2.10 ([30, Section 3.1.1.])**.** *Let F be an (n,m)-function. The following statements are equivalent:*

1. *F is a vectorial bent function.*
2. *The extended Walsh spectrum of F contains only the value $2^{n/2}$.*
3. *For any $c \in \mathbb{F}_{2^m}^*$, the component function $x \mapsto \mathrm{Tr}(cF(x))$ is bent.*
4. *For any $a \in \mathbb{F}_{2^n}^*$, the function $x \mapsto F(x) + F(x+a)$ is balanced.*

The Nyberg bound in particular implies that the bound in Proposition 2.2.8 is not sharp for (n,n)-functions. In this case, we get the following stronger bound:

Theorem 2.2.11 (Sidelnikov-Chabaud-Vaudenay bound, [34])**.** *Let F be an (n,n)-function. Then*

$$NL(F) \leq 2^{n-1} - 2^{\frac{n-1}{2}}.$$

Definition 2.2.12 (Almost bent function)**.** *An (n,n)-function that satisfies the Sidelnikov-Chabaud-Vaudenay bound in Theorem 2.2.11 with equality is called* almost bent *(AB).*

Since the nonlinearity is always an integer, AB functions only exist for odd n. For n even, it is conjectured that $NL(F) \leq 2^{n-1} - 2^{\frac{n}{2}}$ [50]. The extended Walsh spectrum of an almost bent function only contains 2 values, $2^{(n+1)/2}$ and 0. Almost bent functions cannot have a high algebraic degree:

Proposition 2.2.13 ([31]). *The algebraic degree of an almost bent function is bounded from above by* $\frac{n+1}{2}$.

Definition 2.2.14 (Differential spectrum, derivative, differential uniformity). *Let F be an (n,m)-function. We define $\delta_F(a,b)$ by*

$$\delta_F(a,b) = |\{x \in \mathbb{F}_{2^n} \colon F(x+a) + F(x) = b\}|$$

for all $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^m}$. The function $D_aF(x) = F(x+a) + F(x)$ is the derivative *of F in the direction of $a \neq 0$. Further, we define the* differential spectrum *of F as the multiset*

$$\{*\delta_F(a,b) \colon a \in \mathbb{F}_{2^n}^*, b \in \mathbb{F}_{2^m} *\}$$

and the differential uniformity d_F *by*

$$d_F = \max_{a \in \mathbb{F}_{2^n}^*, b \in \mathbb{F}_{2^m}} \delta_F(a,b).$$

The algebraic degree of a derivative is always lower than the algebraic degree of the original function F (as long as F is not constant of course). Note that $\delta_F(a, b)$ is always even. Indeed, if x is a solution of $F(x+a) + F(x) = b$, then so is $x + a$. The differential uniformity of an (n,n)-function plays a key role in differential cryptanalysis [7]. The lower the differential uniformity of a function, the higher is its resistance to a differential attack. As an example, a linear function L over $\mathbb{F}_{2^n}$ has the highest possible differential uniformity 2^n since in this case $L(x+a) + L(x) = b$ is satisfied for all $x \in \mathbb{F}_{2^n}$ if $b = L(a)$. In this sense, differential uniformity can also be seen as a measurement of nonlinearity. The best possible differential uniformity is 2.

Definition 2.2.15 (APN function). *An (n,n)-function with differential uniformity 2 is called* almost perfect nonlinear *(APN).*

The differential spectrum and the extended Walsh spectrum of a vectorial Boolean function (and thus its resistance to linear and differential attacks) are related. In particular, the functions that have optimal nonlinearity also have optimal differential uniformity:

Theorem 2.2.16 ([34]). *Every AB function is APN.*

We will later give examples that show that the converse is not true. We also have the following link between the differential spectrum and the Walsh transform.

Proposition 2.2.17 ([34]). *Let F be an (n,n)-function. We have*

$$W_F(\mu,\lambda)^2 = \sum_{a,b \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(a\lambda + b\mu)} \delta_F(a,b)$$
$$\delta_F(a,b) = 2^{-2n} \sum_{\lambda,\mu \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(a\lambda + b\mu)} W_F(\mu,\lambda)^2$$

for all $a, b, \lambda, \mu \in \mathbb{F}_{2^n}$.

Differential and linear attacks are the two most common attacks against block ciphers, so nonlinearity and differential uniformity can be seen as the two most important cryptographic properties of a vectorial Boolean function. However, other attacks exist and resistance against these attacks can often also be described as a

property of the S-box. For example, the *Boomerang attack* introduced in [133] is another attack that (like the differential attack) exploits differential properties of the S-box. The resistance of a vectorial Boolean function to the Boomerang attack can be measured by the Boomerang Connectivity Table [40]. Another common attack against block ciphers are *differential-linear attacks* introduced in [95]. In Chapter 5 we will give a more detailed analysis of the resistance of vectorial Boolean functions to those attacks.

2.3 Equivalences of vectorial Boolean functions

It can be observed that certain operations on functions will leave some of the cryptographic properties discussed in the previous section invariant. Accordingly, it makes sense to partition the set of all vectorial Boolean functions into equivalence classes, such that all functions in an equivalence class share some cryptographic properties. In this section we will introduce different notions of equivalence, describe which properties are left invariant and compare the equivalence relations with each other.

Definition 2.3.1 (Affine equivalence)**.** *Let F and F' be two (n,m)-functions. We say F and F' are* affine equivalent *if there exist an affine permutation A_1 of $\mathbb{F}_{2^m}$ and another affine permutation A_2 of $\mathbb{F}_{2^n}$ such that*

$$F' = A_1 \circ F \circ A_2.$$

Clearly, the size of the image set is invariant under affine equivalence. In particular, every function that is affine equivalent to a permutation has to be a permutation.

Definition 2.3.2 (EA-equivalence)**.** *Let F and F' be two (n,m)-functions. We say F and F' are* extended affine equivalent *(EA-equivalent) if there exist affine permutations A_1, A_2 of $\mathbb{F}_{2^m}$ and $\mathbb{F}_{2^n}$, respectively, as well as another affine function $A\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ such that*

$$F' = A_1 \circ F \circ A_2 + A.$$

It is obvious that two functions that are affine equivalent are also EA-equivalent. Note that the size of the image set is generally not invariant under EA-equivalence. However, the algebraic degree of a vectorial Boolean function is invariant under EA-equivalence as long as it is at least 2. An even more general concept of equivalence is *CCZ-equivalence* (named after Carlet, Charpin and Zinoviev who introduced it in [31]). We define the *graph* of an (n,m)-function F as the set

$$G_F = \{(x, F(x))\colon x \in \mathbb{F}_{2^n}\} \subset \mathbb{F}_{2^n} \times \mathbb{F}_{2^m}.$$

Definition 2.3.3 (CCZ-equivalence)**.** *Let F and F' be two (n,m)-functions. We say F and F' are* CCZ-equivalent *if there are linear mappings*

$$\alpha\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}, \quad \beta\colon \mathbb{F}_{2^m} \to \mathbb{F}_{2^n},$$
$$\gamma\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}, \quad \delta\colon \mathbb{F}_{2^m} \to \mathbb{F}_{2^m}$$

and $a \in \mathbb{F}_{2^n}$, $b \in \mathbb{F}_{2^m}$, such that $\mathcal{L}(G_F) + (a,b) = G_{F'}$, where $\mathcal{L}\colon \mathbb{F}_{2^n} \times \mathbb{F}_{2^m} \to \mathbb{F}_{2^n} \times \mathbb{F}_{2^m}$ is a bijective mapping defined by

$$\mathcal{L}(x,y) = (\alpha(x) + \beta(y), \gamma(x) + \delta(y))$$

for all $x \in \mathbb{F}_{2^n}$ and $y \in \mathbb{F}_{2^m}$.

Because CCZ-equivalence is essentially just an equivalence of the corresponding graphs, it is sometimes also called *graph equivalence*.

It was shown in [31] that EA-equivalence is a special case of CCZ-equivalence. More precisely, we get the following statement.

Proposition 2.3.4 ([31])**.** *Let F and F' be two (n,m)-functions that are CCZ-equivalent. Using the notation from Definition 2.3.3, if $\mathcal{L}$ can be chosen in a way that $\beta = 0$, then F and F' are EA-equivalent. If $\mathcal{L}$ can be chosen in a way that $\beta = \gamma = 0$, then F and F' are affine equivalent.*

We have another special case of CCZ-equivalence.

Proposition 2.3.5 ([31])**.** *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be a permutation. Then F is CCZ-equivalent to its compositional inverse F^{-1}. Using the notation from Definition 2.3.3, the corresponding function $\mathcal{L}$ is determined by setting $\alpha = \delta = 0$ and $\beta = \gamma = x$.*

Since F and F^{-1} generally do not have the same algebraic degree, this in particular shows that the algebraic degree is not invariant under CCZ-equivalence and that CCZ-equivalence is more general than EA-equivalence.

Both the differential spectrum and the extended Walsh spectrum are preserved under CCZ-equivalence.

Proposition 2.3.6 ([31])**.** *Let F and F' be two CCZ-equivalent (n,m)-functions. The differential spectrum, differential uniformity, extended Walsh spectrum and nonlinearity of F and F' coincide.*

By Proposition 2.3.6, two vectorial Boolean functions that are CCZ-equivalent have the same resistance to linear and differential attacks.

We also want to mention one very specific kind of equivalence that is often used when dealing with monomials, i.e. (n,n)-functions of the form $F(x) = x^d$.

Definition 2.3.7 (Cyclotomic equivalence; cyclotomic cosets)**.** *Let $d, d' \in \{0, 1, \ldots, 2^n - 2\}$. We call d and d'* cyclotomic equivalent *in $\mathbb{Z}_{2^n-1}$ if $d' \equiv 2^i d \pmod{2^n - 1}$ for some $i \in \mathbb{N}$. We call two monomials $F = x^d$, $F' = x^{d'}$ over $\mathbb{F}_{2^n}$ cyclotomic equivalent if d and d' are cyclotomic equivalent in $\mathbb{Z}_{2^n-1}$. The set of all exponents that are cyclotomic equivalent to d is called the* cyclotomic coset *of d.*

Clearly, cyclotomic equivalence of monomials is a special case of affine equivalence. Indeed, if $F(x) = x^d$, then $F(x^{2^i}) = x^{2^i d}$.

2.4 APN functions

By definition, APN functions are the functions that have the best resistance against differential attacks. APN functions are very rare and finding them is generally very difficult. Most APN functions known today are CCZ-equivalent to either a monomial or a quadratic function. For a list of known infinite families of quadratic APN functions up to CCZ-equivalence, we refer to [20]. We want to note that quadratic APN functions are never permutations in even dimension.

Proposition 2.4.1 ([6, Corollary 3])**.** *Let F be a quadratic APN function over $\mathbb{F}_{2^n}$ with n even. Then F is not a permutation.*

We will now deal with the case of APN monomials in more detail.

Definition 2.4.2 (APN exponent). *Let $d \in \{0, 1, \ldots, 2^n - 2\}$. We call d an* APN exponent over $\mathbb{F}_{2^n}$ *if $F = x^d$ is an APN function on $\mathbb{F}_{2^n}$.*

There is the following well-known simple criterion to check whether a monomial is a permutation or not.

Proposition 2.4.3. *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be the monomial $F = x^d$. F is a permutation if and only if $\gcd(d, 2^n - 1) = 1$. In this case, $F^{-1} = x^{1/d}$ where $1/d$ is the inverse of d in $\mathbb{Z}_{2^n-1}$.*

One reason why checking the APN property is easier for monomials is the following well-known proposition.

Proposition 2.4.4. *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be a monomial $F = x^d$. Then $\delta_F(a, b) = \delta_F(1, b/a^d)$ for all $a \in \mathbb{F}_{2^n}^*$.*

The proposition shows that for a monomial it is enough to check that $\delta_F(1, b) \leq 2$ for all $b \in \mathbb{F}_{2^n}$ to prove that F is APN. Note that, by the previous section, all monomials that are cyclotomic equivalent to an APN monomial are also APN. Moreover, if $F = x^d$ is an APN permutation over $\mathbb{F}_{2^n}$, then its compositional inverse $F^{-1} = x^{1/d}$ is also APN. The following result shows that inversion and cyclotomic equivalence alone completely determine CCZ-equivalence for monomials.

Proposition 2.4.5 ([43]). *Let $F, F'\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be two monomials $F = x^d$ and $F' = x^{d'}$. F and F' are CCZ-equivalent if and only if d' is cyclotomic equivalent to d or $1/d$ (if it exists). Here $1/d$ denotes the inverse of d in $\mathbb{Z}_{2^n-1}$.*

Table 2.1 lists all known infinite families of APN monomials up to inversion and cyclotomic equivalence. It is generally believed that the list is complete.

	Exponent	Conditions	AB?	Proof
Gold	$2^r + 1$	$\gcd(r, n) = 1$, $r \leq n/2$	n odd	[56, 113]
Kasami	$2^{2r} - 2^r + 1$	$\gcd(r, n) = 1$ $r \leq n/2$	n odd	[74]
Welch	$2^t + 3$	$n = 2t + 1$	Yes	[46, 68, 23]
Niho	$2^t - 2^{\frac{t}{2}} - 1$	$n = 2t + 1$, t even	Yes	[48, 68]
	$2^t - 2^{\frac{3t+1}{2}} - 1$	$n = 2t + 1$, t odd	Yes	
Inverse	$2^n - 2$	n odd	No	[113]
Dobbertin	$2^{4r} + 2^{3r} + 2^{2r} + 2^r - 1$	$5r = n$	No	[47]

TABLE 2.1: List of known APN exponents over $\mathbb{F}_{2^n}$ up to inversion and cyclotomic equivalence.

There is one big difference between the exponents listed in Table 2.1: For Gold and Kasami exponents, the exponent does not depend on the field size $\mathbb{F}_{2^n}$, while all other exponents do. An exponent that yields an APN function over infinitely many different fields is called an *exceptional* APN exponent. It was proven that the Gold and Kasami exponents are the only exceptional APN exponents [66].

The following proposition determines when an APN monomial is a permutation.

Proposition 2.4.6 ([30, Proposition 9.19.]). *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be an APN monomial, $F = x^d$. Then*

$$\gcd(d, 2^n - 1) = \begin{cases} 1, & \text{if } n \text{ odd}, \\ 3, & \text{if } n \text{ even}. \end{cases}$$

In particular, F is a permutation if and only if n is odd.

The algebraic degree of a monomial $F = x^d$ is just $\mathrm{wt}(d)$. Recall that the algebraic degree is not invariant under taking the inverse. In Chapter 3, we will take a closer look at the inverses of the bijective APN monomials.

Proposition 2.4.6 shows that APN monomials are never permutations in even dimension. In fact, there is (up to equivalence) only one APN permutation known in even dimension - Dillon's permutation over $\mathbb{F}_{2^6}$ [17]. Finding new APN permutations in even dimension (or proving they don't exist) is one of the biggest challenges in the research area and known as the *Big APN problem*. Because of this, it is also interesting to consider permutations in even dimension with differential uniformity 4. Table 2.2 lists the known infinite families of permutation monomials in even dimension with differential uniformity 4.

	Exponent	Conditions	Proof
Gold	2^r+1	t odd, $\gcd(r,n)=2, r \leq n/2$	[56, 113]
Kasami	$2^{2r}-2^r+1$	t odd, $\gcd(r,n)=2, r \leq n/2$	[74]
Inverse	2^n-2		[113]
Bracken-Leander	$2^{2r}+2^r+1$	$4r=n$, r odd	[16]

TABLE 2.2: List of exponents yielding 4 differentially uniform permutations over $\mathbb{F}_{2^n}$ with $n = 2t$ up to inversion and cyclotomic equivalence

2.5 A short introduction to characters

In this section, we give a brief introduction on characters and character sums over finite fields. A detailed treatment can be found in [102, Chapter 5] or [70, Chapter 3]. We introduce characters in the more general setting of finite abelian groups.

Definition 2.5.1 (Character)**.** *Let $(G,\cdot)$ be a finite abelian group with identity element 1_G. A* character *χ of G is a homomorphism from G to the unit circle $U \subset \mathbb{C}$, i.e. a mapping from G to U such that*

$$\chi(g_1 g_2) = \chi(g_1)\chi(g_2)$$

for all $g_1, g_2 \in G$. We denote the set of all characters of G by $\widehat{G}$.

Clearly, $\chi(g)^{|G|} = \chi(g^{|G|}) = \chi(1_G) = 1$, so the values of χ are the $|G|$-th roots of unity. Moreover, since $\chi(g)\chi(g^{-1}) = 1$, we have $\chi(g^{-1}) = (\chi(g))^{-1} = \overline{\chi(g)}$ for all $g \in G$ (here the bar denotes complex conjugation).

Definition 2.5.2 (Trivial and conjugate characters)**.** *The character $\chi \in \widehat{G}$ defined by $\chi(g) = 1$ for all $g \in G$ is called the* trivial character *of G. All other characters are called* nontrivial*. To each character $\chi \in \widehat{G}$ we define its* conjugate character *$\overline{\chi}$ by $\overline{\chi}(g) = \overline{\chi(g)}$.*

We can define a multiplication on $\widehat{G}$ via $(\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g)$ for all $g \in \widehat{G}$. Under this multiplication, $\widehat{G}$ itself becomes a group.

Proposition 2.5.3 ([70, Proposition 3.1.])**.** *Let G be a finite abelian group. Then $G \cong \widehat{G}$. In particular, $|G| = |\widehat{G}|$.*

Proposition 2.5.4 (Orthogonality relations)**.** *Let χ and ψ be characters of G. Then*

$$\sum_{g \in G} \chi(g)\overline{\psi(g)} = \begin{cases} |G|, & \text{if } \chi = \psi \\ 0, & \text{if } \chi \neq \psi. \end{cases}$$

If g and h are elements in G

$$\sum_{\chi \in \widehat{G}} \chi(g)\overline{\psi(h)} = \begin{cases} |G|, & \text{if } g = h \\ 0, & \text{if } g \neq h. \end{cases}$$

When we are working with a finite field $\mathbb{F}_q$, we are naturally interested in two different abelian groups: The additive group $(\mathbb{F}_q, +)$ and the multiplicative group $(\mathbb{F}_q^*, \cdot)$. Accordingly, we will use characters of both these groups. The results on characters covered in this section will hold for finite fields of all characteristics, so we will not make any restrictions on the characteristic.

Definition 2.5.5 (Additive and multiplicative characters)**.** *We call characters of $(\mathbb{F}_q, +)$* additive characters *and characters of $(\mathbb{F}_q^*, \cdot)$* multiplicative characters *of $\mathbb{F}_q$. We write $\widehat{(\mathbb{F}_q, +)} = \widehat{\mathbb{F}_q}$ and $\widehat{(\mathbb{F}_q^*, \cdot)} = \widehat{\mathbb{F}_q^*}$.*

From now on, we will always use χ to refer to an additive character and ψ to refer to a multiplicative character.

Let us consider additive characters of $\mathbb{F}_q$. Let $q = p^n$ with p prime. We define

$$\chi_b(x) = e^{2\pi i \operatorname{Tr}(bx)/p}. \tag{2.2}$$

for all $x \in \mathbb{F}_q$. It can easily be verified that $\chi_{b_1} \neq \chi_{b_2}$ for $b_1 \neq b_2$ and that χ_b is an additive character. Since $|G| = |\widehat{G}|$, all additive characters are obtained through this construction, for instance the trivial additive character corresponds to the case $b = 0$. Additionally, since $\chi_b(x) = \chi_1(bx)$ we can represent all characters using only χ_1. Because of this, we will call χ_1 the *canonical additive character*.

Example 2.5.6. Consider the canonical additive character of the binary finite field $\mathbb{F}_{2^n}$. By Eq. (2.2), we have

$$\chi_1(x) = (-1)^{\operatorname{Tr}(x)}.$$

We can thus write the Walsh transform of a vectorial Boolean (n, n)-function using the canonical additive character:

$$\begin{aligned} W_F(a, b) &= \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\operatorname{Tr}(bF(x)+ax)} = \sum_{x \in \mathbb{F}_{2^n}} \chi_1(bF(x) + ax) \\ &= \sum_{x \in \mathbb{F}_{2^n}} \chi_1(bF(x))\chi_1(ax) = \sum_{x \in \mathbb{F}_{2^n}} \chi_b(F(x))\chi_a(x). \end{aligned}$$

Definition 2.5.7 (Discrete Fourier Transform on finite abelian groups)**.** *Let G be a finite abelian group and $f\colon G \to \mathbb{C}$ be a function. The* Fourier transform *of f is a function $\tilde{f}\colon \widehat{G} \to \mathbb{C}$ defined by*

$$\tilde{f}(\chi) = \sum_{g \in G} f(g)\overline{\chi}(g)$$

for all $\chi \in \widehat{G}$.

Example 2.5.8. The Walsh transform $W_F(a, b) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\operatorname{Tr}(bF(x))}\chi_a(x)$ of an (n, n)-function F can be seen as a discrete Fourier transform on the group $G = (\mathbb{F}_{2^n}, +)$ by setting $f = (-1)^{\operatorname{Tr}(bF(x))}$.

The inversion formula makes it possible to recover a function from its Fourier transform.

Proposition 2.5.9 (Fourier inversion formula, [70, Proposition 3.4.]). *Let G be a finite abelian group and $f\colon G \to \mathbb{C}$ a function. Then*

$$f(g) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \tilde{f}(\chi)\chi(g)$$

for all $g \in G$.

Example 2.5.10. Applying the Fourier inversion formula to the Walsh transform we get the following well-known identity:

$$(-1)^{\mathrm{Tr}(bF(x))} = \frac{1}{2^n} \sum_{a \in \mathbb{F}_{2^n}} \chi_a(x) W_F(a,b).$$

Proposition 2.5.11 (Parseval's identity, [70, Proposition 3.4.]). *Let G be a finite abelian group and $f\colon G \to \mathbb{C}$. Then*

$$\sum_{\chi \in \widehat{G}} |\tilde{f}(\chi)|^2 = |G| \sum_{g \in G} |f(g)|^2.$$

Example 2.5.12. Let us again apply Parseval's identity to the Walsh transform. Recall that in this case $f = (-1)^{\mathrm{Tr}(bF(x))}$, so $|f(g)| = 1$ for all g. We then get

$$\sum_{a \in \mathbb{F}_{2^n}} |W_F(a,b)|^2 = 2^{2n}.$$

Let us now focus on multiplicative characters of $\mathbb{F}_q$. By Proposition 2.5.3, the group of multiplicative characters is a cyclic group of order $q-1$. With this knowledge, the multiplicative characters can be found easily.

Proposition 2.5.13. *Let g be a fixed primitive element of $\mathbb{F}_q$, i.e. a generator of the multiplicative group of $\mathbb{F}_q$. For each $0 \le j \le q-2$ the function defined by*

$$\psi_j(g^k) = e^{2\pi i j k/(q-1)}$$

for each $0 \le k \le q-2$ is a multiplicative character. The ψ_j are distinct and we have $\widehat{\mathbb{F}_q^*} = \{\psi_j\colon 0 \le j \le q-2\}$.

Note that the index j in the previous proposition depends on the choice of the primitive element g. However, independent of the choice of g, the trivial multiplicative character is always ψ_0. Accordingly, we will always use the notation ψ_0 for the trivial multiplicative character.

Example 2.5.14 (The quadratic character)**.** Let q be odd and define $\eta\colon \mathbb{F}_q^* \to \mathbb{C}$ by

$$\eta(x) = \begin{cases} 1, & x \text{ is a square} \\ -1, & x \text{ is a non-square.} \end{cases}$$

As usual, we say x is a square in $\mathbb{F}_q^*$ if there is a $y \in \mathbb{F}_q^*$ such that $y^2 = x$. η is a multiplicative character of $\mathbb{F}_q$. In fact, it is the unique character of order 2 in $\widehat{\mathbb{F}_q^*}$. Among the characters listed in Proposition 2.5.13, we can identify $\eta = \psi_{(q-1)/2}$ independent of the choice of the primitive element g in the Proposition.

We are often interested in bounds on additive character sums of the form $\sum_{x\in\mathbb{F}_q}\chi(F(x))$ for some polynomial $F\in\mathbb{F}_q[x]$. The most famous result on sums of this form is the *Weil bound*.

Theorem 2.5.15 (Weil bound, [102, Theorem 5.38.]). *Let $q=p^n$ with p prime, $F\in\mathbb{F}_q[x]$ be a polynomial with degree $d\geq 1$ and $\chi\in\widehat{\mathbb{F}_q}$. Assume further that F cannot be written as $F=G^p-G+b$ for some other polynomial $G\in\mathbb{F}_q[x]$. Then*

$$\left|\sum_{x\in\mathbb{F}_q}\chi(F(x))\right|\leq(d-1)q^{1/2}.$$

Note that the condition on F is necessary. Indeed, if $F=G^p-G+b$ then, by the properties of the trace-function, we have $\chi(F(x))=\chi(b)$ for all $x\in\mathbb{F}_q$, so $\sum_{x\in\mathbb{F}_q}\chi(F(x))=\pm q$. The Weil bound yields only a nontrivial bound if the degree of F is at most $q^{1/2}$. A similar bound also holds for sums of multiplicative characters, see e.g. [102, Theorem 5.41.].

Definition 2.5.16 (Gauss Sums). *Let χ_1 be the canonical additive character of a finite field $\mathbb{F}_q$. For any multiplicative character ψ of $\mathbb{F}_q$ we define the* Gauss sum *$G(\psi)$ by*

$$G(\psi)=\sum_{a\in\mathbb{F}_q^*}\chi_1(a)\psi(a).$$

The explicit evaluation of Gauss sums is in general very difficult and only known for a few special characters. For the trivial multiplicative character, it is easy to check that $G(\psi_0)=-1$. The most famous and celebrated example is the explicit evaluation of Gauss sums for the quadratic character η, see [102, Theorem 5.15.] for a classical proof or, for a beautiful proof using complex analysis, [70, Theorems 3.9.-3.11.].

Applying the Fourier inversion formula (Proposition 2.5.9) to the definition of the Gauss sum, we immediately get the following connection.

Proposition 2.5.17. *Let $x\in\mathbb{F}_q^*$. Then*

$$\chi_1(x)=\frac{1}{q-1}\sum_{\psi\in\widehat{\mathbb{F}_q^*}}G(\psi)\overline{\psi}(x).$$

For each multiplicative character $\psi\in\widehat{\mathbb{F}_q^*}$ and every positive integer n we can define $\psi^{(n)}=\psi\circ\mathrm{N}_{q^n/q}$. Clearly, $\psi^{(n)}\in\widehat{\mathbb{F}_{q^n}^*}$. We call $\psi^{(n)}$ the *lifted character* of ψ to $\mathbb{F}_{q^n}$. The Davenport-Hasse Theorem explains how the corresponding Gauss sums $G(\psi)$ and $G(\psi^{(n)})$ are related.

Theorem 2.5.18 (Davenport-Hasse Theorem, [102, Theorem 5.14.]). *Let ψ be a multiplicative character of $\mathbb{F}_q$ and $\psi^{(n)}=\psi\circ\mathrm{N}_{q^n/q}\in\widehat{\mathbb{F}_{q^n}^*}$ the lifted character of ψ. We have*

$$G(\psi^{(n)})=(-1)^{n-1}G(\psi)^n.$$

2.5.1 Stickelberger's congruence

We will also need a divisibility result on Gauss sums. The classical result here is a deep theorem in algebraic number theory known as *Stickelberger's congruence* or Stickelberger theorem [126]. This is the only part of the thesis that requires tools

from algebraic number theory, so we will only give a brief sketch of the underlying theory to the extent that is needed to formulate Stickelberger's congruence.

A brief and self-contained introduction to algebraic number theory that culminates in the proof of Stickelberger's congruence can be found in [70, Chapter 4], for a very thorough treatment we refer to [92, Chapter 2 onwards]. A general overview of algebraic number theory can be found for example in [91]. A more advanced book focusing on cyclotomic fields is [134], although it does not contain a proof of Stickelberger's congruence.

Let p be a prime and $q = p^n$ for some positive integer n. Denote by $\zeta_s = e^{2\pi i/s}$ the s-th root of unity. We consider the field extensions $\mathbb{Q} \subseteq \mathbb{Q}(\zeta_{q-1}) \subseteq \mathbb{Q}(\zeta_{q-1}, \zeta_p)$. Recall that $\mathbb{Z}[\zeta_s]$ is the ring of integers of $\mathbb{Q}(\zeta_s)$.

Let $\mathfrak{p}$ be a prime ideal of $\mathbb{Z}[\zeta_{q-1}]$ *lying above* (p), i.e. $\mathfrak{p} \cap \mathbb{Z} = (p)$. $\mathbb{Z}[\zeta_{q-1}]/\mathfrak{p}$ is then a field with characteristic p. Since n is the smallest positive integer such that $p^n \equiv 1 \pmod{q-1}$, the extension degree of $\mathbb{Z}[\zeta_{q-1}]/\mathfrak{p}$ over its prime field is n, so $\mathbb{Z}[\zeta_{q-1}]/\mathfrak{p}$ has $q = p^n$ elements, and we can identify it with the field $\mathbb{F}_q$. The roots of the polynomial $x^{q-1} - 1 \in (\mathbb{Z}[\zeta_{q-1}])[x]$ are ζ_{q-1}^i for all $0 \leq i \leq q-2$. Denote by R the set of these roots, i.e. $R = \langle \zeta_{q-1} \rangle$. Then

$$\mathbb{Z}[\zeta_{q-1}]/\mathfrak{p} = \{0 + \mathfrak{p}, \zeta_{q-1}^0 + \mathfrak{p}, \zeta_{q-1} + \mathfrak{p}, \ldots, \zeta_{q-1}^{q-2} + \mathfrak{p}\}.$$

Definition 2.5.19 (Teichmüller character)**.** *We define a mapping* $\psi_\mathfrak{p} \colon (\mathbb{Z}[\zeta_{q-1}]/\mathfrak{p}) \setminus \{0 + \mathfrak{p}\} \to R$ *via*

$$\psi_\mathfrak{p}(r + \mathfrak{p}) = r \tag{2.3}$$

for all $r \in R$. $\psi_\mathfrak{p}$ *is a multiplicative character of* $\mathbb{Z}[\zeta_{q-1}]/\mathfrak{p}$. *We call* $\psi_\mathfrak{p}$ *the* Teichmüller character.

It is clear that the Teichmüller character has order $q-1$, so we can identify it with a generator of $\widehat{\mathbb{F}_q^*}$.

(p) is totally ramified in $\mathbb{Q}(\zeta_p)$ and $\mathfrak{p}$ is totally ramified in $\mathbb{Q}(\zeta_{q-1}, \zeta_p)$, so there is a unique prime ideal $\mathfrak{P}$ of $\mathbb{Z}[\zeta_{q-1}, \zeta_p]$ lying above $\mathfrak{p}$.

Let j be an integer, $0 \leq j \leq q-2$. We write $j = j_0 + j_1 p + j_2 p^2 + \cdots + j_{n-1} p^{n-1}$ with $0 \leq j_i \leq p-1$ in base-p expansion. With this notation, we are ready to state Stickelberger's congruence.

Theorem 2.5.20 (Stickelberger's congruence, [92, Theorem 2.1.])**.** *Let* $1 \leq j \leq q-2$ *and write* j *in base* p, *i.e.* $j = j_0 + j_1 p + j_2 p^2 + \cdots + j_{n-1} p^{n-1}$. *Then we have*

$$\frac{G(\psi_\mathfrak{p}^{-j})}{(\zeta_p - 1)^{j_0 + j_1 + \cdots + j_{n-1}}} \equiv \frac{-1}{(j_0)!(j_1)! \ldots (j_{n-1})!} \pmod{\mathfrak{P}},$$

where $\psi_\mathfrak{p}$ *and* $\mathfrak{P}$ *are defined as described above.*

Applied to binary finite fields, we have $\zeta_2 = -1$, $j_0 + j_1 + j_2 + \cdots + j_{n-1} = \mathrm{wt}(j)$ and $(j_i)! = 1$ for all i. Recall that $\mathfrak{P}$ lies above (2), so we get the following relation.

Corollary 2.5.21 (Stickelberger's congruence in characteristic 2)**.** *With the same notation as before and* $p = 2$, *we have*

$$G(\psi_\mathfrak{p}^{-j}) \equiv 2^{\mathrm{wt}(j)} \pmod{2^{\mathrm{wt}(j)+1}}.$$

Identifying the Teichmüller character $\psi_\mathfrak{p}$ with a generator of $\widehat{\mathbb{F}_{2^n}^*}$, we can use Corollary 2.5.21 to determine the 2-divisibility of Gauss sums over a binary finite

field. This divisibility result for Gauss sums has been used in a variety of contexts before. For example, it is a key element of the classical theorem by Ax about the number of solutions of a multivariate polynomial over a finite field [2]. Moreover, it has been used to determine the weight divisibility of cyclic codes via McEliece's theorem [110], in the proof of the Welch and Niho conjectures [68, 23], and to study monomial bent functions [93].

When we identify $\mathbb{Z}[\zeta_{q-1}]/\mathfrak{p}$ with $\mathbb{F}_q$ and the Teichmüller character with a generator ψ of $\widehat{\mathbb{F}_q^*}$, Eq. (2.3) from Definition 2.5.19 implies

$$\psi(x) \equiv x \pmod{p} \tag{2.4}$$

for all $x \in \mathbb{F}_q^*$.

Chapter 3

Inversion in $\mathbb{Z}_{2^n-1}$

The work in this chapter is based on [83], written by the author of this thesis, which is accepted for publication in *Designs, Codes and Cryptography*.

As outlined in the previous chapter, a bijective vectorial Boolean function and its inverse are CCZ-equivalent, so the function and its inverse share many cryptographic properties (see Proposition 2.3.6). It is thus desirable to find an explicit formula for inverses of bijective functions that display good cryptographic properties, like APN functions. Moreover, not all cryptographic properties of a function are invariant under inversion, the most notable exception is the algebraic degree. For arbitrary bijective functions, finding the inverse is a very challenging task. In the case of monomials, determining the inverse of a function $x \mapsto x^l$ on $\mathbb{F}_{2^n}$ amounts to finding the inverse of l in the ring $\mathbb{Z}_{2^n-1}$ (see Proposition 2.4.3). Of course, for fixed values of l and n, this can easily be done using (for instance) the Euclidean algorithm. However, for infinite families of monomials (like the APN monomials displayed in Table 2.1), other theoretical tools are needed. The problems comes in two "flavors": The exponent l may depend on n (as for Welch/Niho/Dobbertin exponents in Table 2.1) or not (Gold and Kasami exponents). In many ways, exponents of the second type are more difficult to handle. In [87], a method to find the inverse of a fixed exponent l modulo $2^n - 1$ for arbitrary n was given. This technique was used to determine the algebraic degree of the inverses of the Gold exponents as well as the second Kasami exponent $K_2 = 13$. Unfortunately, it is unclear how to use this approach to determine the inverses of all (infinitely many) Kasami exponents. In fact, just determining the binary weight of the inverses of Kasami exponents is mentioned as an open problem in [87].

In this chapter, we develop a new technique that can be used to determine the inverse of elements l in $\mathbb{Z}_{2^n-1}$ for infinitely many n. Here, l may depend on n or not. We will use this technique to find the inverses of all invertible Gold and Kasami exponents, including the non-APN exponents, as well as the Bracken-Leander exponent. With these results, all inverses of invertible APN exponents and all inverses of exponents that give rise to bijective monomials with differential uniformity 4 have been found. For an overview, see Tables 3.1 and 3.2.

3.1 A new method for inversion in $\mathbb{Z}_{2^n-1}$: The modular add-with-carry approach

Our approach uses as the key tool the *modular add-with-carry approach* that was first formally introduced by Hollmann and Xiang [68].

Theorem 3.1.1 (Modular add-with-carry approach, [68, Theorem 13])**.** *Let* $a, s \in \{1, \ldots, 2^n - 2\}$ *and* $l \in \mathbb{N}$. *We denote by* $a = (a_{n-1}, \ldots, a_0)$ *and* $s = (s_{n-1}, \ldots, s_0)$ *the*

	Exponent	Conditions	d	d_{inv}	Inverse found in
Gold	2^r+1	$\gcd(r,n)=1$ $r<n/2$	2	$\frac{n+1}{2}$	[113]
Kasami	$2^{2r}-2^r+1$	$\gcd(r,n)=1$ $r<n/2$	$r+1$	Varies	Here
Welch	2^t+3		3	t or $t+1$	[87]
Niho	$2^t-2^{\frac{t}{2}}-1$ $2^t-2^{\frac{3t+1}{2}}-1$	t even t odd	$\frac{t+2}{2}$ $t+1$	$\frac{3n+5}{8}$ or $\frac{3n+9}{8}$ $\frac{3n+7}{8}$ or $\frac{3n+11}{8}$	[118],[87]
Inverse	$2^{2t}-1$		$n-1$	$n-1$	Obvious
Dobbertin	$2^{4r}+2^{3r}+2^{2r}+2^r-1$	$5r=n$	$r+3$	$\frac{n+3}{2}$	[87]

TABLE 3.1: List of known APN exponents over $\mathbb{F}_{2^n}$ with $n=2t+1$ up to inversion and cyclotomic equivalence. d and d_{inv} denote the algebraic degree of the monomial and its inverse, respectively.

	Exponent	Conditions	d	d_{inv}	Inverse found in
Gold	2^r+1	t odd, $\gcd(r,n)=2$ $r\leq n/2$	2	$\frac{n}{2}$	[87], here
Kasami	$2^{2r}-2^r+1$	t odd, $\gcd(r,n)=2$ $r\leq n/2$	$r+1$	Varies	Here
Inverse	2^n-2		$n-1$	$n-1$	Obvious
Bracken-Leander	$2^{2r}+2^r+1$	$4r=n$, r odd	3	$\frac{n+2}{2}$	Here

TABLE 3.2: List of exponents yielding 4 differentially uniform permutations over $\mathbb{F}_{2^n}$ with $n=2t$ up to inversion and cyclotomic equivalence. d and d_{inv} denote the algebraic degree of the monomial and its inverse, respectively.

binary expansions of a and s. Let $l=\sum_j t_j 2^j$ with $t_j\in\mathbb{Z}$. Further, let $t_+=\sum_{j,t_j>0} t_j$ and $t_-=\sum_{j,t_j<0} t_j$. The following are equivalent:

(a) $s\equiv l\cdot a \pmod{2^n-1}$

(b) There exists a sequence $c=(c_{n-1},\ldots,c_0)$ with $c_i\in\{t_-,\ldots,t_+-1\}$ (called the carry sequence*) such that*

$$2c_i-c_{i-1}+s_i=\sum_j t_j a_{i-j} \tag{3.1}$$

holds for all i. Here, the indices are seen as elements in $\mathbb{Z}_n$.

The carry sequence in (b) is unique.

Remark 3.1.2. Note that the representation $l=\sum_j t_j 2^j$ with integer coefficients t_j in Theorem 3.1.1 is not unique. In fact, this is one of the major strengths of this theorem since it makes it possible to choose a representation that has more structure than the (usual) binary representation. This makes a big difference especially for the Kasami exponents. Indeed, the r-th Kasami exponent K_r can be written (as it is done usually) as $K_r=2^{2r}-2^r+1$, i.e. with $t_{2r}=t_0=1$ and $t_r=-1$. This is certainly a much simpler representation than the binary representation that has $r+1$ ones. In the general case, it seems to be desirable to choose a representation such that both t_+ and t_- have low absolute value so that the range of the possible values for the carry sequence is small.

The basic idea of finding the inverse of some value l modulo 2^n-1 is now quite simple: We use Theorem 3.1.1 and set $s=1$. Then we try to find sequences a and c that satisfy Eq. (3.1). While we apply the approach in this chapter only to a few

selected exponents, the idea can in principle be used for arbitrary values of l. However, the corresponding sequences a and c are highly dependent on the choice of l, so a general treatment seems to be impossible. Still, this approach gives a good framework to find inverses in $\mathbb{Z}_{2^n-1}$.

In many cases, educated guesses based on experimental results for low values of n are enough to find the inverse. In particular, the carry sequence c often has a strong and visible structure. Since the carry sequence uniquely determines the sequence a, the strategy for the proofs is to find/guess the structure of the carry sequence and then construct the inverse from the carry sequence.

3.2 The Gold exponents

To illustrate our method using Theorem 3.1.1, we use it to derive the inverses of the invertible Gold exponents $G_r = 2^r + 1$ in $\mathbb{Z}_{2^n-1}$. The inverses of the APN Gold exponents (i.e. with the condition $\gcd(r, n) = 1$) are explicitly given in [113]. Moreover, the algebraic degree of the inverses of all Gold exponents is known [87, Theorem 3.7.]. However, as far as we know, the explicit binary expansion of the inverses of the non-APN Gold exponents has not appeared anywhere in the literature yet. In this section, we apply the add-with-carry approach to find the binary expansion of the inverses of all invertible Gold exponents. In particular, this also yields a new proof for the algebraic degree of the Gold functions.

Applied to the Gold exponent, Theorem 3.1.1 yields the following.

Theorem 3.2.1. *Let $a, s \in \{1, \ldots, 2^n - 2\}$ and $G_r = 2^r + 1$ be the r-th Gold exponent. We denote by $a = (a_{n-1}, \ldots, a_0)$ and $s = (s_{n-1}, \ldots, s_0)$ the binary expansions of a and s. The following are equivalent:*

(a) $s \equiv G_r \cdot a \pmod{2^n - 1}$

(b) There exists a carry sequence $c = (c_{n-1}, \ldots, c_0)$ with $c_i \in \{0, 1\}$ such that

$$2c_i - c_{i-1} + s_i = a_{i-r} + a_i \tag{3.2}$$

holds for all i. Here, the indices are seen as elements in $\mathbb{Z}_n$.

The carry sequence in (b) is unique.

The following lemma characterizes all invertible Gold exponents.

Lemma 3.2.2 (e.g. [109, Lemma 11.1.])**.** *Let r and n be positive integers. The Gold exponent $G_r = 2^r + 1$ is invertible in $\mathbb{Z}_{2^n-1}$ if and only if $\frac{n}{\gcd(n,r)}$ is odd.*

3.2.1 The APN Gold exponents

We first deal with the APN Gold exponents $G_r = 2^r + 1$ over $\mathbb{F}_{2^n}$ with $\gcd(r, n) = 1$. We will use some notation from [94], where the modular add-with-carry approach was used to find the Walsh support of the Kasami functions. In particular, we will use the notion of r-ordered sequences.

Since $\gcd(r, n) = 1$ we can reorder the sequences a and c in Theorem 3.2.1 in the following way:

$$a_0, a_{-r}, a_{-2r}, \ldots, a_{-(n-1)r} \text{ and } c_0, c_{-r}, c_{-2r}, \ldots, c_{-(n-1)r}.$$

Here, we view again the indices as elements in $\mathbb{Z}_n$. This ordering is technically a decimation of the sequence by $-r$. Since we will be using this ordering a lot, we will call it the *r-ordering of a sequence* and also denote these sequences by

$$a_0, a_1, a_2, \ldots, a_{n-1} \text{ and } c_0, c_1, c_2, \ldots, c_{n-1},$$

where we will always make sure to specify whether we use the regular ordering or the r-ordering.

By Lemma 3.2.2, G_r is invertible if and only if n is odd. We denote by e the least positive residue of the inverse of r modulo n. Using r-ordered sequences, the key equation in Theorem 3.2.1 takes on the following simpler form:

Theorem 3.2.3. *Let $n \in \mathbb{N}$, $a \in \{1, \ldots, 2^n - 2\}$ and G_r be the r-th Gold exponent with $\gcd(r, n) = 1$. Let e be the least positive residue of the inverse of r modulo n and $(a_0, \ldots, a_{n-1})$ be the r-ordered sequence of the binary representation of a, i.e. $a \equiv \sum_{i=0}^{n-1} a_i 2^{-ir} \pmod{2^n - 1}$. The following are equivalent:*

(a) a is the inverse of G_r modulo $2^n - 1$.

(b) There exists an r-ordered carry sequence $c = (c_0, c_1, \ldots, c_{n-1})$ with $c_i \in \{0, 1\}$ such that

$$2c_0 - c_e + 1 = a_1 + a_0 \tag{3.3}$$
$$2c_i - c_{i+e} = a_{i+1} + a_i \tag{3.4}$$

holds for all $i \in \mathbb{Z}_n$, $i \neq 0$.

The carry sequence in (b) is unique.

Now, we can use Theorem 3.2.3 to give a simple alternative proof for the inverse of the APN Gold exponents. As pointed out earlier, the idea is to guess the structure of the carry sequence from examples for low n and then compute the inverse from the carry sequence. We will make a detailed example to give an intuition for this process. Note that the case of APN Gold functions is easier than other cases (especially the Kasami cases in the next section), but the approach will always remain the same.

Example 3.2.4. Let $n = 7$, $r = 3$ and consider the invertible APN Gold exponent $G_3 = 2^3 + 1 = 9$. We have $3 \cdot 5 \equiv 1 \pmod 7$, so $e = 5$. The inverse of 9 modulo $2^7 - 1$ is $113 = 2^6 + 2^5 + 2^4 + 2^0$, and the binary sequence of 113 is in regular ordering $(a_6, a_5, \ldots, a_0) = (1, 1, 1, 0, 0, 0, 1)$ and in r-ordering $G_3^{-1} = (a_0, \ldots, a_6) = (1, 1, 0, 1, 0, 1, 0)$. We have $a_1 = a_0 = 1$ so by Eq. (3.3) necessarily $c_0 = c_5 = 1$. If $i \neq 0$ we have $a_i + a_{i+1} = 1$ and thus by Eq. (3.4) $c_i = c_{i+5} = 1$. We conclude that the carry sequence consists exclusively of ones.

From this example (and possibly other examples for low n) we guess that the carry sequence always consists only of ones. Let us now consider the Eq.s (3.3) and (3.4) for this choice and arbitrary n. It necessarily yields $a_1 = a_0 = 1$ and $a_{i+1} + a_i = 1$. From this, we immediately conclude that the r-ordered sequence of the inverse G_r^{-1} is $(a_0, \ldots, a_{n-1}) = (1, 1, 0, 1, 0, 1, \ldots, 0, 1, 0)$. Since both Eq.s (3.3) and (3.4) are satisfied, this must be the r-ordered sequence of the inverse of the Gold exponent. We have thus proven the following:

Proposition 3.2.5 (Inverses of APN Gold exponents, [113, Proposition 5]). *Let $G_r = 2^r + 1$ with $\gcd(r, n) = 1$ and n odd. Then G_r is invertible in $\mathbb{Z}_{2^n-1}$ and the least positive residue of its inverse is*

$$G_r^{-1} = \sum_{i=0}^{\frac{n-1}{2}} 2^{2ir}.$$

In particular, $\mathrm{wt}(G_r^{-1}) = \frac{n+1}{2}$, so the algebraic degree of $x \mapsto x^{G_r^{-1}}$ over $\mathbb{F}_{2^n}$ is $\frac{n+1}{2}$.

Proof. By the considerations above, the r-ordered sequence of the inverse G_r^{-1} is $(a_0, \ldots, a_{n-1}) = (1, 1, 0, 1, 0, 1, \ldots, 0, 1, 0)$. We conclude

$$G_r^{-1} \equiv 1 + \sum_{\substack{i \in \{0,\ldots,n-1\} \\ i \text{ odd}}} 2^{-ir} \equiv \sum_{\substack{i \in \{0,\ldots,n-1\} \\ i \text{ even}}} 2^{ir} \equiv \sum_{i=0}^{\frac{n-1}{2}} 2^{2ir} \pmod{2^n - 1}.$$

□

The main takeaway from the example is that the carry sequence has a simpler structure than the sequence $(a_0, \ldots, a_{n-1})$ of the inverse. This observation will also hold for all other exponents considered in this chapter. Indeed, while it is still possible to discern the structure of the APN Gold exponents with relative ease without looking at the carry sequence, this will be close to impossible in the case of the Kasami exponents.

3.2.2 The non-APN Gold exponents

We now deal with the more general case of Gold exponents G_r with $\gcd(n, r) > 1$.

Since $\gcd(n, r) > 1$, we cannot use the r-ordering of sequences that we used in the Proposition 3.2.5. We expand the concept in a natural way.

Definition 3.2.6 (r-matrices). *Let $a = (a_{n-1}, \ldots, a_0)$ be a sequence of integers and r be a positive integer. Set $d = \gcd(n, r)$. We define the associated $(d \times \frac{n}{d})$-matrix $M_{a,r}$ by*

$$M_{a,r} = \begin{pmatrix} a_0 & a_{-r} & a_{-2r} & \cdots & a_{-(\frac{n}{d}-1)r} \\ a_1 & a_{1-r} & a_{1-2r} & \cdots & a_{1-(\frac{n}{d}-1)r} \\ \vdots & & & & \vdots \\ a_{d-1} & a_{d-1-r} & a_{d-1-2r} & \cdots & a_{d-1-(\frac{n}{d}-1)r} \end{pmatrix},$$

*where the indices are seen as elements in $\mathbb{Z}_n$. We call $M_{a,r}$ the r-*matrix of *a. If a is a binary sequence, then we call $M_{a,r}$ also the r-matrix of the corresponding element in $\mathbb{Z}_{2^n-1}$ or $\{0, 1, \ldots, 2^n - 2\}$.*

Since the r-matrices are constructed from sequences, we use the slightly unusual convention of indexing from 0, i.e. the first row/column will be called row/column 0. With this convention, the r-ordered sequences considered in the previous section are just a special case of r-matrices with only one row. Again in accordance to the notation used earlier, we denote by e the least positive residue of the inverse of $\frac{r}{\gcd(n,r)}$ modulo $\frac{n}{\gcd(n,r)}$.

We now use r-matrices to rephrase Theorem 3.2.1.

Theorem 3.2.7. *Let $a \in \{1, \ldots, 2^n - 2\}$, $n \in N$ and G_r be the r-th Gold exponent with $\gcd(r, n) = d$ and e be the least positive residue of the inverse of $\frac{r}{d}$ modulo $\frac{n}{d}$. Moreover, let*

$$M_{a,r} = \begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} & \ldots & a_{0,\frac{n}{d}-1} \\ a_{1,0} & a_{1,1} & a_{1,2} & \ldots & a_{1,\frac{n}{d}-1} \\ \vdots & & & & \vdots \\ a_{d-1,0} & a_{d-1,1} & a_{d-1,2} & \ldots & a_{d-1,\frac{n}{d}-1} \end{pmatrix}$$

be the r-matrix of a, i.e. $a \equiv \sum_{i=0}^{d-1} \sum_{j=0}^{\frac{n}{d}-1} a_{i,j} 2^{i-jr} \pmod{2^n - 1}$. The following are equivalent:

(a) a is the inverse of G_r modulo $2^n - 1$.

(b) There exists an r-matrix for the carry sequence c of the form

$$M_{c,r} = \begin{pmatrix} c_{0,0} & c_{0,1} & c_{0,2} & \ldots & c_{0,\frac{n}{d}-1} \\ c_{1,0} & c_{1,1} & c_{1,2} & \ldots & c_{1,\frac{n}{d}-1} \\ \vdots & & & & \vdots \\ c_{d-1,0} & c_{d-1,1} & c_{d-1,2} & \ldots & c_{d-1,\frac{n}{d}-1} \end{pmatrix}$$

with $c_{i,j} \in \{0, 1\}$ such that the following equations hold:

$$2c_{0,0} - c_{d-1,e} + 1 = a_{0,1} + a_{0,0} \tag{3.5}$$

$$2c_{0,j} - c_{d-1,j+e} = a_{0,j+1} + a_{0,j} \text{ for all } j \in \{1, \ldots, \frac{n}{d} - 1\} \tag{3.6}$$

$$2c_{i,j} - c_{i-1,j} = a_{i,j+1} + a_{i,j} \text{ for all } i \in \{1, \ldots, d-1\}, j \in \{0, \ldots, \frac{n}{d} - 1\}. \tag{3.7}$$

The carry sequence (and thus its associated r-matrix) in (b) is unique.

Proof. The Theorem follows immediately from Theorem 3.2.1 and the definition of the r-matrix. The predecessor of the values $c_{-k_1 r}$ is determined as follows: Observe that $c_{-k_1 r - 1} = c_{\gcd(n,r)-1-k_2 r}$ if and only if $-k_1 r - 1 \equiv \gcd(n, r) - 1 - k_2 r \pmod{n}$, which is equivalent to $-(k_1 - k_2)\frac{r}{\gcd(n,r)} \equiv 1 \pmod{\frac{n}{\gcd(n,r)}}$, so the predecessor of $c_{-k_1 r}$ is $c_{\gcd(n,r)-1-k_2 r}$ with $k_2 = k_1 + e$.

□

With Theorem 3.2.7, we can give the explicit binary representation of all inverses of Gold exponents.

Proposition 3.2.8. *Let $n \in \mathbb{N}$ and G_r be the r-th Gold exponent with $\gcd(r, n) = d > 1$ and $\frac{n}{d}$ odd. Let e be the least positive residue of the inverse of $\frac{r}{d}$. Then $G_r^{-1} \equiv \sum_{i=0}^{d-1} \sum_{j=0}^{\frac{n}{d}-1} a_{i,j} 2^{i-jr} \pmod{2^n - 1}$ where the values $a_{i,j}$ are the entries of the $(d \times \frac{n}{d})$-matrix*

$$M_{a,r} = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & \ldots & 0 & 1 \\ \vdots & & & & & & & & \\ 0 & 0 & 1 & 0 & 1 & 0 & \ldots & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & \ldots & 1 & 0 \end{pmatrix}.$$

In particular, $\mathrm{wt}(G_r^{-1}) = \frac{n-d+2}{2}$.

Proof. The r-matrix of the corresponding carry sequence is the $(d \times \frac{n}{d})$-matrix

$$M_{c,r} = (c_{i,j}) = \begin{pmatrix} 0 & 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & 1 & \dots & 1 \\ \vdots & & & & & \vdots \\ 0 & 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & 1 & \dots & 1 \end{pmatrix}.$$

We now just have to verify Eq.s (3.5) to (3.7). For Eq. (3.5), we get $2c_{0,0} - c_{d-1,e} + 1 = a_{0,1} + a_{0,0} = 0$. For Eq. (3.6), we have $2c_{0,j} - c_{d-1,j+e} = a_{0,j+1} + a_{0,j} = 1$ for all values of $j > 0$. For Eq. (3.7), we have $2c_{i,j} - c_{i-1,j} = a_{i,j+1} + a_{i,j} = 0$ if $j = 0$ and $0 < i < d-1$, $2c_{i,j} - c_{i-1,j} = a_{i,j+1} + a_{i,j} = 2$ if $j = 0$ and $i = d-1$ and $2c_{i,j} - c_{i-1,j} = a_{i,j+1} + a_{i,j} = 1$ in all other possible cases. Thus, all equations are satisfied.

The value of $\mathrm{wt}(G_r^{-1})$ can be determined easily by counting the ones in the matrix $M_{a,r}$. □

Recall that all invertible Gold exponents satisfy $\frac{n}{\gcd(r,n)}$ odd by Lemma 3.2.2, so Propositions 3.2.5 and 3.2.8 cover all invertible Gold exponents.

Note that the $\gcd(r,n) = 1$ case can even be recovered as a special case from Proposition 3.2.8. Indeed, the last row of the r-matrices $M_{a,r}$ and $M_{c,r}$ are precisely the r-sequences we saw in the $\gcd(r,n) = 1$ case in Proposition 3.2.5.

3.3 The Kasami exponents

As shown in Table 3.1, the inverses of all APN exponents except the Kasami exponent have been determined in earlier work. The objective in this section is thus the following: Find the inverse of $K_r = 2^{2r} - 2^r + 1$ modulo $2^n - 1$ for all possible values of r, n. Compared to other APN exponents, determining the inverses of the Kasami exponents is particularly challenging because they are independent from the field size. While this is also true for the Gold exponents we considered in the previous section, finding the inverses of Gold exponents is relatively easy because of their low binary weight. In contrast, the algebraic degree of the Kasami exponents is unbounded which makes the determination of the inverses much harder.

Applied to the Kasami exponent $K_r = 2^{2r} - 2^r + 1$, Theorem 3.1.1 yields the following.

Theorem 3.3.1. *Let $a, s \in \{1, \dots, 2^n - 2\}$ and K_r be the r-th Kasami exponent. We denote by $a = (a_{n-1}, \dots, a_0)$ and $s = (s_{n-1}, \dots, s_0)$ the binary expansions of a and s. The following are equivalent:*

(a) $s \equiv K_r \cdot a \pmod{2^n - 1}$

(b) There exists a carry sequence $c = (c_{n-1}, \dots, c_0)$ with $c_i \in \{-1, 0, 1\}$ such that

$$2c_i - c_{i-1} + s_i = a_{i-2r} - a_{i-r} + a_i \tag{3.8}$$

holds for all i. Here, the indices are seen as elements in $\mathbb{Z}_n$.

The carry sequence in (b) is unique.

We extend the definition of the weight of a sequence $(s_{n-1}, \dots, s_0)$ to the sum of all of its elements. For binary sequences, this corresponds exactly to its binary weight. In particular, this allows us to talk about the weight of the carry sequence

which is in general not a binary sequence. Using this convention, the following Lemma gives an additional condition on the carry sequence for the Kasami exponents.

Lemma 3.3.2 ([68], Lemma 5)**.** *With the notation of Theorem 3.3.1, we have the following:*

(a) $c_i + c_{i-r} \in \{-1, 0, 1\}$. *In particular,* $|\operatorname{wt}(c)| \leq \frac{n}{2}$.

(b) $\operatorname{wt}(c) + \operatorname{wt}(s) = \operatorname{wt}(a)$. *In particular, for* $s = 1$ *we have* $\operatorname{wt}(c) = \operatorname{wt}(a) - 1$.

The following Proposition shows when a Kasami exponent is invertible modulo $2^n - 1$.

Proposition 3.3.3 ([87, Lemma 3.8])**.** *Let n be a positive integer and $K_r = 2^{2r} - 2^r + 1$ be the r-th Kasami exponent. K_r is invertible modulo $2^n - 1$ if and only if one of the following cases occurs:*

- $\frac{n}{\gcd(r,n)}$ *is odd,*
- $\frac{n}{\gcd(r,n)}$ *is even, r is even and* $\gcd(r, n) = \gcd(3r, n)$.

We first deal with the case $\gcd(r, n) = 1$, then with the case $\frac{n}{\gcd(r,n)}$ odd and finally with the case $\frac{n}{\gcd(r,n)}$ even. Technically, the case $\gcd(r, n) = 1$ is included in the case $\frac{n}{\gcd(r,n)}$ odd. However, we single out this case for two reasons: Firstly, it is particularly interesting since those Kasami exponents are precisely the APN exponents. Secondly, the case $\frac{n}{\gcd(r,n)}$ odd is very technical, but can be described much easier by applying the results for the special case $\gcd(r, n) = 1$.

3.3.1 The case $\gcd(r, n) = 1$

We first deal with the APN Kasami exponents $K_r = 2^{2r} - 2^r + 1$ over $\mathbb{F}_{2^n}$ with $\gcd(r, n) = 1$.

By Proposition 3.3.3, K_r is invertible if and only if n is odd. We denote by e the least positive residue of the inverse of r modulo n. Observe that K_r and K_{n-r} are cyclotomic equivalent exponents on $\mathbb{F}_{2^n}$. Indeed, $(2^{2(n-r)} - 2^{n-r} + 1)2^{2r} \equiv 2^{2r} - 2^r + 1 \pmod{2^n - 1}$. Then $K_r^{-1} \equiv 2^{-2r} K_{n-r}^{-1} \pmod{2^n - 1}$, so it suffices to determine the inverse of one of these two values. Since n is odd, we can thus assume without loss of generality that e is odd.

Since $\gcd(r, n) = 1$ we can reorder the sequences a and c in Theorem 3.3.1 using the r-sequences introduced in the previous section. Using r-ordered sequences, the key equation for the Kasami exponents in Theorem 3.3.1 takes on the following form:

Theorem 3.3.4. *Let $n \in \mathbb{N}$, $a \in \{1, \ldots, 2^n - 2\}$ and K_r be the r-th Kasami exponent with* $\gcd(r, n) = 1$. *Let $(a_0, \ldots, a_{n-1})$ be the r-ordered sequence of the binary representation of a, i.e.* $a \equiv \sum_{i=0}^{n-1} a_i 2^{-ir} \pmod{2^n - 1}$. *The following are equivalent:*

(a) a is the inverse of K_r modulo $2^n - 1$.

(b) There exists an r-ordered carry sequence $c = (c_0, c_1, \ldots, c_{n-1})$ with $c_i \in \{-1, 0, 1\}$ such that

$$2c_0 - c_e + 1 = a_2 - a_1 + a_0 \tag{3.9}$$

$$2c_i - c_{i+e} = a_{i+2} - a_{i+1} + a_i \tag{3.10}$$

holds for all $i \in \mathbb{Z}_n$, $i \neq 0$.

The carry sequence in (b) is unique.

Experimental results show that the inverses of the APN Kasami exponents often have binary weight $\frac{n+1}{2}$. In this case, Lemma 3.3.2 immediately shows that the r-ordered carry sequence has weight $\frac{n-1}{2}$ and must be a cyclic shift of the sequence $(0,0,1,0,1,\ldots,0,1)$. Since the carry sequence of the inverse uniquely determines the inverse, these cases can then be solved with comparatively little effort.

In this section, we will always use r-ordered sequences to represent inverses of Kasami exponents because this notation makes the description much easier. Consequently, the inverses will be written in the form $K_r^{-1} \equiv \sum_{i=0}^{n-1} a_i 2^{-ir} \pmod{2^n-1}$ for a sequence $a = (a_0,\ldots,a_{n-1})$. Of course, a translation into the more standard binary representation is easy by reordering the sequence a, i.e. $K_r^{-1} \equiv \sum_{i=0}^{n-1} a_{-ie} 2^i \pmod{2^n-1}$ (recall that e denotes the inverse of r modulo n).

Proposition 3.3.5. *Let n odd, K_r be the r-th Kasami exponent with $\gcd(r,n)=1$. Let e be the least positive residue of the inverse of r modulo n. The inverse of K_r modulo 2^n-1 is*

$$K_r^{-1} \equiv \sum_{i=0}^{n-1} a_i 2^{-ir} \pmod{2^n-1},$$

where $a=(a_0,\ldots,a_{n-1})$ is determined as follows:

- *If $e = 6k+1$, then $a=(1,x,y)$ and $x=(1,0,1,0,\ldots,1,0,1,0)$ is a sequence of length $n-e$ and $y=(1,1,1,0,0,0,1,1,1,0,0,0,\ldots,1,1,1,0,0,0)$ is a sequence of length $6k$.*
- *If $e = 6k+5$, then $a=(0,x,1,1,y)$ and $x=(0,1,0,1,\ldots,0,1,0,1)$ is a sequence of length $n-e+2$ and $y=(0,0,0,1,1,1,0,0,0,1,1,1,\ldots,0,0,0,1,1,1)$ is a sequence of length $6k$.*

In both cases, we have $\mathrm{wt}(K_r^{-1}) = \frac{n+1}{2}$.

Proof. Let $e = 6k+1$. Set $c=(0,1,0,1,\ldots,0,1,0,1,0)$, i.e. $c_i = 0$ if i is even, and $c_i = 1$ otherwise. We show that a and c satisfy the conditions in Theorem 3.3.4. Eq.(3.9) can be easily verified. For Eq. (3.10), we have the following:
Case 1: i odd, $i+e<n$: We have $c_i=1$, $c_{i+e}=0$, $a_i=a_{i+2}=1$ and $a_{i+1}=0$.
Case 2: i even, $i+e<n$: We have $c_i=0$, $c_{i+e}=1$, $a_i=a_{i+2}=0$ and $a_{i+1}=1$.
Case 3: i odd, $i+e\geq n$: We have $c_i=c_{i+e}=1$. Depending on the value of i, the triple (a_i,a_{i+1},a_{i+2}) takes on the values $(1,0,0)$, $(0,0,1)$ or $(1,1,1)$.
Case 4: i even, $i+e\geq n$: We have $c_i=c_{i+e}=0$. Depending on the value of i, the triple (a_i,a_{i+1},a_{i+2}) takes on the values $(0,0,0)$, $(0,1,1)$ or $(1,1,0)$.
So Eq. (3.10) holds for all i.

Now let $e = 6k+5$. Set $c=(0,0,1,0,1,0,1,0,1,\ldots,0,1)$, i.e. $c_i=0$ if $i=0$ or i odd and $c_i=1$ otherwise. We again show that Equations (3.9) and (3.10) are satisfied. Observe that Eq. (3.9) holds. We check the following cases of Eq. (3.10) for $i>0$:
Case 1: i odd, $i+e\leq n$: We have $c_i=0$, $c_{i+e}=1$, $a_i=a_{i+2}=0$ and $a_{i+1}=1$.
Case 2: i even, $i+e\leq n$: We have $c_i=1$, $c_{i+e}=0$, $a_i=a_{i+2}=1$ and $a_{i+1}=0$.
Case 3: i odd, $i+e>n$: We have $c_i=c_{i+e}=0$. Depending on the value of i, the triple (a_i,a_{i+1},a_{i+2}) takes on the values $(0,0,0)$, $(0,1,1)$ or $(1,1,0)$.
Case 4: i even, $i+e>n$: We have $c_i=c_{i+e}=1$. Depending on the value of i, the triple (a_i,a_{i+1},a_{i+2}) takes on the values $(1,0,0)$, $(0,0,1)$ or $(1,1,1)$.
So Eq. (3.10) holds for all i. □

The Kasami APN functions and their inverses are also almost bent functions (see Table 2.1). By Proposition 2.2.13, the algebraic degree of an almost bent functions is bounded from above by $\frac{n+1}{2}$. We have shown that the inverses of the Kasami APN functions defined by the exponents considered in Proposition 3.3.5 attain this bound.

The only case left to check is $e = 6k+3$ (recall that we could assume e odd without loss of generality). This case is a lot more involved and has to be divided into several subcases. The key difference to the cases considered above is that $\mathrm{wt}(K_r^{-1}) < \frac{n+1}{2}$ for $e = 6k+3$, so finding the correct carry sequence is more complicated. However, the strategy of the proof remains the same: Based on experimental results, we guess a carry sequence that then determines the inverse.

Proposition 3.3.6. *Let n odd, K_r be the r-th Kasami exponent with $\gcd(r,n) = 1$. Let $e = 6k+3$ be the least positive residue of the inverse of r modulo n. Define $s,t \in \mathbb{N}$ by $n = se + t$ with $0 \le t < e$. Further, let $x_1 = (0,0,0,1,1,1)$, $x_2 = (0,1,1,1,0,0)$ be sequences of length 6 and*

$$x = (0,1,1,\underbrace{x_1,\ldots,x_1}_{k\text{-times}},0,0,0,\underbrace{x_2,\ldots,x_2}_{k\text{-times}})$$
$$y = (0,0,0,\underbrace{x_2,\ldots,x_2}_{k\text{-times}},0,1,1,\underbrace{x_1,\ldots,x_1}_{k\text{-times}})$$

be sequences of length $2e$. Then

$$K_r^{-1} \equiv \sum_{i=0}^{n-1} a_i 2^{-ir} \pmod{2^n-1},$$

where $a = (a_0,\ldots,a_{n-1})$ is determined as follows:

(a) If $t = 6u+1$ then

$$a = (\underbrace{x_1,\ldots,x_1}_{u\text{-times}},\ \underbrace{y,\ldots,y}_{(s-2)/2\text{-times}},0,0,0,\underbrace{x_2,\ldots,x_2}_{k\text{-times}},0,1,1,\underbrace{x_1,\ldots,x_1}_{u\text{-times}},z,0)+(1,0,\ldots,0),$$

where $z = (0,1,0,1,\ldots,0,1,0,1)$ is a sequence of length $e-3-6u$.

(b) If $t = 6u+2$ then

$$a = (0,1,\underbrace{x_1,\ldots,x_1}_{u\text{-times}},\ \underbrace{y,\ldots,y}_{(s-1)/2\text{-times}},0,0,z,1,z_2),$$

where $z = (1,0,1,0\ldots,1,0,1,0)$ is a sequence of length $6u$ and $z_2 = (x_2,\underbrace{x_1,\ldots,x_1}_{(e-6u-9)/6\text{ -times}})$ is a sequence of length $e-6u-3$.

(c) If $t = 6u+4$ then

$$a = (0,0,0,\underbrace{x_2,\ldots,x_2}_{u\text{-times}},\ \underbrace{x,\ldots,x}_{(s-1)/2\text{-times}},0,1,1,0,z,1,0,1,1,1,\underbrace{x_1,\ldots,x_1}_{(e-6u-9)/6\text{ -times}},0)$$

where $z = (0,1,0,1,\ldots,0,1,0,1)$ is a sequence of length $6u$.

(d) If $t = 6u + 5$ then

$$a = (0,1,1,0,0,\underbrace{x_2,\ldots,x_2}_{u\text{-times}},\underbrace{x,\ldots,x}_{(s-2)/2\text{-times}},0,1,1,\underbrace{x_1,\ldots,x_1}_{k\text{-times}},0,\underbrace{x_1,\ldots,x_1}_{u\text{-times}},z)$$

where $z = (0,1,0,1,\ldots,0,1,0,1)$ is a sequence of length $e-1-6u$.

In the cases (a) and (d) we have $\mathrm{wt}(K_r^{-1}) = \frac{n-s+1}{2}$ *and in the cases (b) and (c)* $\mathrm{wt}(K_r^{-1}) = \frac{n-s}{2}$.

Proof. For all 4 cases, we explicitly give the carry sequence c (in r-ordering) and check that Eq. (3.9) and (3.10) are satisfied. The carry sequences for all cases are quite similar and are composed of the same "building blocks". The verification is simple but tedious, so we will show the correctness of the first case in detail and for the other cases we will just state the carry sequence and omit the verification. We define the auxiliary sequences $s_1 = (0,1,0,1,\ldots,0,1)$ of length $6u$ and $s_2 = (0,0,1,0,1,0,1,0,1,\ldots,0,1)$ of length $e = 6k+3$.
Case (a): Set

$$c = (s_1,\underbrace{s_2,\ldots,s_2}_{s\text{-times}},0).$$

Eq. (3.9) can be easily verified. For Eq. (3.10) we have to distinguish (many) different cases depending on the value of i. We go through each block in the sequence a.
Case a.1: $i > 0$ is in the first block of x_1's. If i is even then we have $c_i = c_{i+e} = 0$ and $(a_i,a_{i+1},a_{i+2}) \in \{(0,1,1),(1,1,0),(0,0,0)\}$. If i is odd then $c_i = c_{i+e} = 1$ and $(a_i,a_{i+1},a_{i+2}) \in \{(0,0,1),(1,1,1),(1,0,0)\}$.
Case a.2: i is in the block of y's. Let $i = 6u+q$. If $q \equiv 1,2,e+1,e+2 \pmod{2e}$ then $c_i = c_{i+e} = 0$. In these first two cases we have $(a_i,a_{i+1},a_{i+2}) = (0,0,0)$ and in the latter two $(a_i,a_{i+1},a_{i+2}) = (0,1,1)$ and $(a_i,a_{i+1},a_{i+2}) = (1,1,0)$, respectively. Let q_1 be the least positive residue of q modulo $2e$. If $3 \le q_1 \le e$ and q_1 odd we have $c_i = c_{i+e} = 0$ and $(a_i,a_{i+1},a_{i+2}) \in \{(0,1,1),(1,1,0),(0,0,0)\}$. If $3 \le q_1 \le e$ and q_1 even, we have $c_i = c_{i+e} = 1$ and $(a_i,a_{i+1},a_{i+2}) \in \{(0,0,1),(1,1,1),(1,0,0)\}$. If $q_1 > e+2$ and q_1 odd we have $c_i = c_{i+e} = 1$ and $(a_i,a_{i+1},a_{i+2}) \in \{(0,0,1),(1,1,1),(1,0,0)\}$ and if $q_1 > e+2$ and q_1 even we have $c_i = c_{i+e} = 0$ and $(a_i,a_{i+1},a_{i+2}) \in \{(0,1,1),(1,1,0),(0,0,0)\}$.
Case a.3: i is in the position of the three zeros after the block of y's. For the first two zeros (i.e. $i = 6u+e(s-2)$ and $i = 6u+e(s-2)+1$) we have $c_i = c_{i+e} = 0$ and $(a_i,a_{i+1},a_{i+2}) = (0,0,0)$. For $i = 6u+e(s-2)+2$ we have $c_i = c_{i+e} = 1$ and $(a_i,a_{i+1},a_{i+2}) = (0,0,1)$.
Case a.4: i is in the block of x_2's, i.e. $i \in \{6u+e(s-2)+3,\ldots,6u+e(s-2)+6k+2\}$. If i is odd, we have $c_i = c_{i+e} = 1$ and $(a_i,a_{i+1},a_{i+2}) \in \{(0,0,1),(1,1,1),(1,0,0)\}$ and if i is even $c_i = c_{i+e} = 0$ and $(a_i,a_{i+1},a_{i+2}) \in \{(0,1,1),(1,1,0),(0,0,0)\}$.
Case a.5: $i \in \{6u+e(s-2)+6k+3,\ldots,6u+e(s-2)+6k+5\}$. If $i = 6u+e(s-2)+6k+3$ then $c_i = 0$, $c_{i+e} = c_{n-1} = 0$ and $(a_i,a_{i+1},a_{i+2}) = (0,1,1)$. For $i = 6u+e(s-2)+6k+4$ we have $c_i = 0$, $c_{i+e} = c_0 = 0$ and $(a_i,a_{i+1},a_{i+2}) = (1,1,0)$ and for $i = 6u+e(s-2)+6k+5$ we have $c_i = 1$, $c_{i+e} = c_1 = 1$ and $(a_i,a_{i+1},a_{i+2}) = (1,0,0)$.
Case a.6: i is in the second block of x_1's, i.e. $i \in \{6u+e(s-2)+6k+6,\ldots,12u+e(s-2)+6k+5\}$. If i is even, we have $c_i = c_{i+e} = 1$ and $(a_i,a_{i+1},a_{i+2}) \in \{(0,1,1),(1,1,0),(0,0,0)\}$. If i is odd and $i \neq 12u+e(s-2)+6k+5$ we have $c_i = c_{i+e} = 0$ and $(a_i,a_{i+1},a_{i+2}) \in \{(0,0,1),(1,1,1),(1,0,0)\}$. If $i = 12u+e(s-2)+6k+5$ then $c_i = 1$, $c_{i+e} = 0$ and $(a_i,a_{i+1},a_{i+2}) = (1,0,1)$.
Case a.7: i is in the subsequence z, i.e. $i \in \{12u+e(s-2)+6k+5, 6u+e(s-1)+$

$6k+1\}$. If i is even then $c_i = 0$, $c_{i+e} = 1$ and $(a_i, a_{i+1}, a_{i+2}) = (0,1,0)$. If i is odd then $c_i = 1$, $c_{i+e} = 0$ and $(a_i, a_{i+1}, a_{i+2}) = (1,0,1)$.
So Eq.(3.10) holds for all $i \neq 0$.
We state the r-ordered carry sequences for the other cases:
Case (b):

$$c = (-1, 1, s_1, \underbrace{s_2, \ldots, s_2}_{s\text{-times}}).$$

Case (c):

$$c = (0, 0, 1, s_1, \underbrace{s_2, \ldots, s_2}_{s\text{-times}}, 0).$$

Case (d):

$$c = (0, 0, 1, 0, 1, s_1, \underbrace{s_2, \ldots, s_2}_{s\text{-times}}).$$

□

Note that Proposition 3.3.6 lists all possible options. Indeed, the cases $t = 6u$ and $t = 6u+3$ do not occur because in these cases $n = se + t$ is divisible by 3, so $e = 6k+3$ is never invertible modulo n.

Corollary 3.3.7. *Let $n \in \mathbb{N}$ and K_r be the r-th Kasami exponent with $\gcd(n,r) = 1$. Let K_r^{-1} be the inverse of K_r modulo $2^n - 1$. Then $\mathrm{wt}(K_r^{-1}) = \frac{n+1}{2}$ for $n \equiv 0 \pmod 3$. Moreover, we have*

$$\mathrm{wt}(K_r^{-1}) \geq \begin{cases} \frac{n+2}{3} & \text{if } n \equiv 1 \pmod 3 \\ \frac{n+1}{3} & \text{if } n \equiv 2 \pmod 3. \end{cases}$$

The lower bound is attained if and only if $e = 3$.

Proof. If $n \equiv 0 \pmod 3$ then e is not divisible by 3 since $\gcd(e,n) = 1$. The result then follows from Proposition 3.3.5.

For the other cases, using the notation of Proposition 3.3.6, the binary weight $\mathrm{wt}(K_r^{-1})$ is minimal when s is maximal. For $n = se + t$ with $0 < t < e$ this clearly implies minimizing e, so $e = 3$ and $t \in \{1,2\}$. For these cases we have

$$\mathrm{wt}(K_r^{-1}) = \begin{cases} \frac{n - \frac{n-1}{3} + 1}{2} = \frac{n+2}{3} & \text{if } t = 1 \\ \frac{n - \frac{n-2}{3}}{2} = \frac{n+1}{3} & \text{if } t = 2 \end{cases}$$

and the result follows.

□

Since EA equivalence preserves the algebraic degree, we get the following simple corollary.

Corollary 3.3.8. *Let $n \in \mathbb{N}$ odd and K_r be the r-th Kasami exponent with $\gcd(n,r) = 1$ and $r < \frac{n}{2}$. Let $F = x^{K_r}$ be the r-th Kasami function on $\mathbb{F}_{2^n}$. If $n \equiv 0 \pmod 3$ and $r \neq \frac{n-1}{2}$ then F is not EA equivalent to F^{-1}. If $n \not\equiv 0 \pmod 3$ and $r < \frac{n-2}{3}$ then F is not EA-equivalent to F^{-1}.*

3.3.2 The case $\frac{n}{\gcd(n,r)}$ odd

We now deal with the Kasami exponents K_r with $\gcd(n,r) > 1$ and $\frac{n}{\gcd(n,r}$ odd. While these Kasami exponents are not APN, they still have some interesting properties.

For example, for $\gcd(r,n) = 2$ and $\frac{n}{2}$ odd, the function $x \mapsto x^{K_r}$ (and thus also its inverse) is a permutation with differential uniformity 4 (see Table 2.2).

Since $\gcd(n,r) > 1$, we cannot use the r-ordering of sequences that we used in the previous section. Just like in the case of Gold functions, we will thus use r-matrices (introduced in Definition 3.2.6).

In accordance to the notation used in the previous subsection, we denote by e the least positive residue of the inverse of $\frac{r}{\gcd(n,r)}$ modulo $\frac{n}{\gcd(n,r)}$. Since $\gcd(n,r) = \gcd(n-r,r)$, $\frac{n}{\gcd(n,r)}$ odd and K_r is cyclotomic equivalent to K_{n-r}, it again suffices to determine the inverses of K_r where e is odd.

Using r-matrices, Theorem 3.3.1 takes on the following form.

Theorem 3.3.9. *Let $a \in \{1,\dots,2^n-2\}$, $n \in N$ and K_r be the r-th Kasami exponent with $\gcd(r,n) = d$ and e be the least positive residue of $\frac{r}{d}$ modulo $\frac{n}{d}$. Moreover, let*

$$M_{a,r} = \begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} & \dots & a_{0,\frac{n}{d}-1} \\ a_{1,0} & a_{1,1} & a_{1,2} & \dots & a_{1,\frac{n}{d}-1} \\ \vdots & & & & \vdots \\ a_{d-1,0} & a_{d-1,1} & a_{d-1,2} & \dots & a_{d-1,\frac{n}{d}-1} \end{pmatrix}$$

be the r-matrix of a, i.e. $a \equiv \sum_{i=0}^{d-1}\sum_{j=0}^{\frac{n}{d}-1} a_{i,j}2^{i-jr} \pmod{2^n-1}$. The following are equivalent:

(a) a is the inverse of K_r modulo 2^n-1.

(b) There exists an r-matrix for the carry sequence c of the form

$$M_{c,r} = \begin{pmatrix} c_{0,0} & c_{0,1} & c_{0,2} & \dots & c_{0,\frac{n}{d}-1} \\ c_{1,0} & c_{1,1} & c_{1,2} & \dots & c_{1,\frac{n}{d}-1} \\ \vdots & & & & \vdots \\ c_{d-1,0} & c_{d-1,1} & c_{d-1,2} & \dots & c_{d-1,\frac{n}{d}-1} \end{pmatrix}$$

with $c_{i,j} \in \{-1,0,1\}$ such that the following equations hold:

$$2c_{0,0} - c_{d-1,e} + 1 = a_{0,2} - a_{0,1} + a_{0,0} \tag{3.11}$$

$$2c_{0,j} - c_{d-1,j+e} = a_{0,j+2} - a_{0,j+1} + a_{0,j} \text{ for all } j \in \{1,\dots,\frac{n}{d}-1\} \tag{3.12}$$

$$2c_{i,j} - c_{i-1,j} = a_{i,j+2} - a_{i,j+1} + a_{i,j} \text{ for all } i \in \{1,\dots,d-1\}, j \in \{0,\dots,\frac{n}{d}-1\}. \tag{3.13}$$

The carry sequence (and thus its associated r-matrix) in (b) is unique.

Proof. The Theorem follows immediately from Theorem 3.3.1 and the definition of the r-matrix. The process is identical to the corresponding case for the Gold function in Theorem 3.2.7. □

Again, we find $M_{a,r}$ and $M_{c,r}$ such that Eq. (3.11)-(3.13) hold. These verifications become quite tedious (especially since we have to distinguish several cases). However, the basic idea does not change: The r-matrices of the carry sequences have a visible structure that can be used to determine the inverse. It turns out that the inverse of K_r on modulo 2^n-1 with $\gcd(r,n) = d$ is closely related to the inverse of $K_{\frac{r}{d}}$

modulo $2^{\frac{n}{d}}-1$ which was already determined in the previous section. To improve readability, we first deal with the case $\frac{n}{\gcd(n,r)} = 6v+3$ for a $v \in \mathbb{N}_0$ separately.

Proposition 3.3.10. *Let $n \in \mathbb{N}$ and K_r be the r-th Kasami exponent with $\gcd(r,n) = d > 1$ and $\frac{n}{d} = 6v+3$. Let e be the least positive residue of the inverse of $\frac{r}{d}$ modulo $\frac{n}{d}$. Then $K_r^{-1} \equiv \sum_{i=0}^{d-1}\sum_{j=0}^{\frac{n}{d}-1} a_{i,j}2^{i-jr} \pmod{2^n-1}$ where the values $a_{i,j}$ are the entries of the matrix $M_{a,r}$*

$$M_{a,r} = \begin{pmatrix} a_1 \\ \vdots \\ a_1 \\ a_2 \end{pmatrix},$$

where the rows a_1 and a_2 are defined as follows:

(a) If $e = 6k+1$:

$$a_1 = (0,0,\underbrace{x_1,\ldots,x_1}_{\frac{n/d-e-2}{6}\text{ -times}},\underbrace{x_2\ldots,x_2}_{k\text{ -times}},0)$$
$$a_2 = (1,x_3,\underbrace{x_4\ldots,x_4}_{k\text{ -times}}).$$

(b) If $e = 6k+5$:

$$a_1 = (0,1,0,0,0,\underbrace{x_4,\ldots,x_4}_{\frac{n/d-e-4}{6}\text{ -times}},1,1,0,0,\underbrace{x_6\ldots,x_6}_{k\text{ -times}},0)$$
$$a_2 = (0,x_5,\underbrace{x_2\ldots,x_2}_{k\text{ -times}}),$$

where $x_1 = (0,0,1,1,1,0)$, $x_2 = (0,0,0,1,1,1)$, $x_4 = (1,1,1,0,0,0)$, $x_6 = (0,1,1,1,0,0)$ are sequences of length 6, $x_3 = (1,0,1,0\ldots,1,0,1,0)$ is a sequence of length $\frac{n}{d}-e$ and $x_5 = (0,1,0,1,\ldots,0,1)$ is a sequence of length $\frac{n}{d}-e$.

In both cases we have $\mathrm{wt}(K_r^{-1}) = \frac{n-3d+4}{2}$.

Proof. Case (a): The r-matrix of the corresponding carry sequence is

$$M_{c,r} = (c_{i,j}) = \begin{pmatrix} c' \\ \vdots \\ c' \\ c'' \end{pmatrix}$$

where $c'' = (c_0,\ldots,c_{\frac{n}{d}-1}) = (0,1,0,1,\ldots,0,1,0,1,0)$ and $c' = (c_e-1, c_{e+1},\ldots,c_{\frac{n}{d}-1}, c_0,c_1,\ldots,c_{e-1})$.

Using Theorem 3.3.9, we just have to verify Eq. (3.11) - (3.13). For our choice of $M_{c,r}$, we have in Eq. (3.11) $2c_{0,0} - c_{d-1,e} + 1 = 2(c_e-1) - c_e + 1 = c_e - 1$. Similarly, in Eq. (3.13) we have for $0 < i < d-1$ and $j = 0$ the relation $2c_{i,j} - c_{i-1,j} = 2(c_e - 1) - (c_e - 1) = c_e - 1$. From these two observations, we conclude

$$c_e - 1 = a_{i,2} - a_{i,1} + a_{i,0} \text{ for all } i \in \{0,\ldots,d-2\}. \tag{3.14}$$

For $j \neq 0$ we have for $i = 0$ (consulting Eq. (3.12)) $2c_{0,j} - c_{d-1,e+j} = 2(c_{e+j}) - c_{e+j} = c_{e+j}$. Looking at Eq. (3.13) for $j \neq 0$ and $i \neq 0$, we have $2c_{i,j} - c_{i-1,j} = 2c_{e+j} - c_{e+j} =$

c_{e+j}. We conclude

$$c_{e+j} = a_{i,j+2} - a_{i,j+1} + a_{i,j} \text{ for all } i \in \{0,\dots,d-1\}, j \in \{1,\dots,\frac{n}{d}-1\}. \tag{3.15}$$

Let us now consider the case $i = d-1$, $j \neq 0$. Then Eq. (3.13) becomes $2c_{d-1,j} - c_{d-2,j} = 2c_j - c_{e+j}$ and we get

$$2c_j - c_{e+j} = a_{d-1,j+2} - a_{d-1,j+1} + a_{d-1,j} \text{ for all } j \in \{1,\dots,\frac{n}{d}-1\}. \tag{3.16}$$

Finally, for the case $i = d-1$ and $j = 0$ we have (again considering Eq. (3.13)) $2c_{d-1,0} - c_{d-2,0} = 2c_0 - (c_e - 1) = 2c_0 - c_e + 1$. We conclude

$$2c_0 - c_e + 1 = a_{d-1,2} - a_{d-1,1} + a_{d-1,0} \tag{3.17}$$

Observe that, by Proposition 3.3.5, a_2 is the r-ordered sequence of the inverse of $K_{\frac{r}{d}}$ modulo $2^{\frac{n}{d}} - 1$ with the corresponding carry sequence c''. Theorem 3.3.4 then shows that Eq. (3.16) and (3.17) are satisfied. We check Eq. (3.14) and (3.15) by hand. In both equations we do not consider the last row of $M_{a,r}$ and since all but the last row in $M_{a,r}$ are identical, it suffices to check the first row.

Eq. (3.14) holds because $c_e = 1$ and $a_{0,2} = a_{0,1} = a_{0,0} = 0$. We check Eq. (3.15): If $e + j < n$ and j odd, then $c_{e+j} = 0$ and $(a_{0,j}, a_{0,j+1}, a_{0,j+2}) \in \{(0,0,0),(1,1,0),(0,1,1)\}$. If $e + j < n$ and j even, then $c_{e+j} = 1$ and $(a_{0,j}, a_{0,j+1}, a_{0,j+2}) \in \{(1,0,0),(1,1,1),(0,0,1)\}$. If $e + j \geq n$ and j is odd then $c_{e+j} = 1$ and $(a_{0,j}, a_{0,j+1}, a_{0,j+2}) \in \{(1,0,0),(1,1,1),(0,0,1)\}$ and if $e + j \geq n$ and j is even then $c_{e+j} = 0$ and $(a_{0,j}, a_{0,j+1}, a_{0,j+2}) \in \{(0,0,0),(1,1,0),(0,1,1)\}$.

Case (b): The proof is similar to the proof of the first case. We define the r-matrix of the corresponding carry sequence

$$M_{c,r} = (c_{i,j}) = \begin{pmatrix} c' \\ \vdots \\ c' \\ c'' \end{pmatrix}$$

where $c'' = (c_0,\dots,c_{\frac{n}{d}-1}) = (0,0,1,0,1,0,1,\dots,0,1,0,1)$ and $c' = (c_e - 1, c_{e+1},\dots, c_{\frac{n}{d}-1}, c_0, c_1,\dots,c_{e-1})$. This leads to precisely the same equations (3.14)-(3.17). Again, by Proposition 3.3.5, a_2 and c'' are the r-ordered sequences of the inverse of the Kasami exponent $K_{\frac{r}{d}}$ modulo $2^{\frac{n}{d}} - 1$ and the corresponding carry sequence, respectively. The validity of Eq. (3.16) and (3.17) follows. Equations (3.14) and (3.15) can be checked just as in the previous case; we omit the calculations.

By adding all entries in $M_{c,r}$, we see that in both cases the weight of the carry sequence is $d\frac{n/d-3}{2} + 1 = \frac{n-3d+2}{2}$. Lemma 3.3.2 then implies $\mathrm{wt}(K_r^{-1}) = \frac{n-3d+4}{2}$. □

Note that the case $e = 6k+3$ does not occur because e is invertible modulo $\frac{n}{d} = 6v+3$. We now deal with the remaining cases $\frac{n}{d} = 6v+1$ and $\frac{n}{d} = 6v+5$.

Proposition 3.3.11. *Let $n \in \mathbb{N}$ and K_r be the r-th Kasami exponent with $\gcd(r,n) = d$ and $\frac{n}{d}$ odd. Let e be the least positive residue of the inverse of $\frac{r}{d}$ modulo $\frac{n}{d}$ and $\frac{n}{d} = se + t$, $0 \leq t < e$. Then $K_r^{-1} \equiv \sum_{i=0}^{d-1}\sum_{j=0}^{\frac{n}{d}-1} a_{i,j}2^{i-jr} \pmod{2^n-1}$ where the values $a_{i,j}$ are the*

entries of the matrix

$$M_{a,r} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_2 \end{pmatrix}.$$

Here, a_1 is the sequence of the inverse of $K_{\frac{r}{d}}$ modulo $2^{\frac{n}{d}} - 1$ in r-ordering as determined in the previous section and a_2 is as follows. We use the auxiliary sequences $x_1 = (0,0,0,1,1,1)$, $x_2 = (1,1,0,0,0,1)$, $x_3 = (0,1,1,1,0,0)$ of length 6 and

$$y = (0,0,0,\underbrace{x_3,\ldots,x_3}_{k\text{-times}},0,1,1,\underbrace{x_1,\ldots,x_1}_{k\text{-times}})$$
$$z = (0,1,1,\underbrace{x_1,\ldots,x_1}_{k\text{-times}},0,0,0,\underbrace{x_3,\ldots,x_3}_{k\text{-times}})$$

of length $12k+6$.

(a) If $e = 6k+1$ and $\frac{n}{d} = 6v+1$

$$a_2 = (\underbrace{x_1,\ldots,x_1}_{v\text{-times}},0).$$

(b) If $e = 6k+1$ and $\frac{n}{d} = 6v+5$

$$a_2 = (\underbrace{x_2,\ldots,x_2}_{v\text{-times}},1,1,0,0,0).$$

(c) If $e = 6k+5$ and $\frac{n}{d} = 6v+1$

$$a_2 = (0,\underbrace{x_1,\ldots,x_1}_{v\text{-times}}).$$

(d) If $e = 6k+5$ and $\frac{n}{d} = 6v+5$

$$a_2 = (0,1,1,\underbrace{x_1,\ldots,x_1}_{v\text{-times}},0,0).$$

(e) If $e = 6k+3$ and $t = 6u+1$

$$a_2 = (\underbrace{x_1,\ldots,x_1}_{u\text{-times}},\underbrace{y,\ldots,y}_{\frac{s}{2}\text{-times}},0).$$

(f) If $e = 6k+3$ and $t = 6u+2$

$$a_2 = (0,1,\underbrace{x_1,\ldots,x_1}_{u\text{-times}},\underbrace{y,\ldots,y}_{\frac{s-1}{2}\text{-times}},0,0,0,\underbrace{x_3,\ldots,x_3}_{k\text{-times}}).$$

(g) If $e = 6k+3$ and $t = 6u+4$

$$a_2 = (0,0,0,\underbrace{x_3,\ldots,x_3}_{u\text{-times}},\underbrace{z,\ldots,z}_{\frac{s-1}{2}\text{-times}},0,1,1,\underbrace{x_1,\ldots,x_1}_{u\text{-times}},0).$$

(h) If $e = 6k + 3$ *and* $t = 6u + 5$

$$a_2 = (0, 1, 1, 0, 0, \underbrace{x_3, \ldots, x_3}_{u\text{-times}}, \underbrace{z, \ldots, z}_{\frac{s}{2}\text{-times}}).$$

In the cases (a)-(d) we have $\mathrm{wt}(K_r^{-1}) = \frac{n-d+2}{2}$*, in the cases (e) and (h)* $\mathrm{wt}(K_r^{-1}) = \frac{n-d(s+1)+2}{2}$ *and in cases (f) and (g)* $\mathrm{wt}(K_r^{-1}) = \frac{n-d(s+2)+2}{2}$.

Proof. In all cases the r-matrix of the carry sequence c has identical rows, i.e.

$$M_{c,r} = \begin{pmatrix} c' \\ \vdots \\ c' \end{pmatrix},$$

where $c' = (c_0, \ldots, c_{\frac{n}{d}-1})$ is the r-ordered carry sequence for the inverse of $K_{\frac{r}{d}}$ modulo $2^{\frac{n}{d}} - 1$ determined in the proofs of Propositions 3.3.5 and 3.3.6. With this carry sequence, the equations (3.11)-(3.13) of Theorem 3.3.9 take on the following form:

$$2c_0 - c_e + 1 = a_{0,2} - a_{0,1} + a_{0,0} \tag{3.18}$$

$$2c_j - c_{j+e} = a_{0,j+2} - a_{0,j+1} + a_{0,j} \text{ for all } j \in \{1, \ldots, \frac{n}{d} - 1\} \tag{3.19}$$

$$c_j = a_{i,j+2} - a_{i,j+1} + a_{i,j} \text{ for all } i \in \{1, \ldots, d-1\}, j \in \{0, \ldots, \frac{n}{d} - 1\}. \tag{3.20}$$

The validity of Eq. (3.18) and (3.19) follows from Theorem 3.3.4 and the choice of a_1 and c'. So we only need to verify Eq. (3.20) for each case. We will show the verification for the first case, the other cases are identical in nature.

In Case (a) we have $c' = (0, 1, 0, 1, \ldots, 0, 1, 0)$ from Proposition 3.3.5, i.e. c_j is 0 if j is even and 1 of j is odd. When j is odd, then $(a_{i,j}, a_{i,j+1}, a_{i,j+2}) \in \{(0, 0, 1), (1, 1, 1), (1, 0, 0)\}$ and if j is even then $(a_{i,j}, a_{i,j+1}, a_{i,j+2}) \in \{(0, 0, 0), (1, 1, 0), (0, 1, 1)\}$ for all $i > 0$, so Eq. (3.20) holds.

Using Lemma 3.3.2, we have

$$\mathrm{wt}(K_r^{-1}) = \mathrm{wt}(c) + 1 = d\,\mathrm{wt}(c') + 1 = d(\mathrm{wt}(K_{\frac{r}{d}}^{-1}) - 1) + 1,$$

where $K_{\frac{r}{d}}^{-1}$ is the least positive residue of the inverse of $K_{\frac{r}{d}}$ modulo $2^{\frac{n}{d}} - 1$. The results on the binary weights then follow from the results in Propositions 3.3.5 and 3.3.6. For example for the cases (a)-(d), Proposition 3.3.5 yields $\mathrm{wt}(K_{\frac{r}{d}}^{-1}) = \frac{\frac{n}{d}+1}{2}$. This leads to $\mathrm{wt}(K_r^{-1}) = d\frac{\frac{n}{d}-1}{2} + 1 = \frac{n-d+2}{2}$.

□

Propositions 3.3.10 and 3.3.11 show that K_r^{-1} has a strong structure because its r-matrix has $d - 1$ identical rows. By the definition of the r-matrix, this means that K_r^{-1} has $\frac{n}{d}$ runs of $(d - 1)$ consecutive ones or zeroes.

The results presented in this section yield the following result for the binary weight of the inverse of Kasami exponents.

Corollary 3.3.12. *Let* $n \in \mathbb{N}$ *and* K_r *be the* r*-th Kasami exponent with* $\gcd(n, r) = d$ *and* $\frac{n}{d}$ *odd. Let* K_r^{-1} *be the inverse of* K_r *modulo* $2^n - 1$*. Then* $\mathrm{wt}(K_r^{-1}) = \frac{n-3d+4}{2}$ *for* $n \equiv 0$

$\pmod 3$ *and* $\mathrm{wt}(K_r^{-1}) \leq \frac{n-d+2}{2}$ *for* $n \not\equiv 0 \pmod 3$. *Moreover, we have*

$$\mathrm{wt}(K_r^{-1}) \geq \begin{cases} \frac{n-d+3}{3} & \text{if } \frac{n}{d} \equiv 1 \pmod 3 \\ \frac{n-2d+3}{3} & \text{if } \frac{n}{d} \equiv 2 \pmod 3. \end{cases}$$

Proof. For $n \equiv 0 \pmod 3$ the result follows from Proposition 3.3.10.

For the other cases, using the notation of Proposition 3.3.11, the binary weight $\mathrm{wt}(K_r^{-1})$ is minimal when e is divisible by 3 and s is maximal. For $n/d = se + t$ with $0 < t < e$ this clearly implies minimizing e, so $e = 3$ and $t \in \{1, 2\}$. With Case (e) and (f) from Proposition 3.3.11, we have

$$\mathrm{wt}(K_r^{-1}) \geq \begin{cases} \frac{1}{2}(n - \frac{n+2d}{3} + 2) = \frac{n-d+3}{3} & \text{if } t = 1 \\ \frac{1}{2}(n - \frac{n+4d}{3} + 2) = \frac{n-2d+3}{3} & \text{if } t = 2 \end{cases}$$

and the result follows.

□

3.3.3 The case $\frac{n}{\gcd(n,r)}$ even

We now deal with the case $\frac{n}{\gcd(n,r)}$ even. Proposition 3.3.3 implies that if K_r is invertible modulo $2^n - 1$ then both n and r are even and $\frac{n}{\gcd(n,r)}$ is not divisible by 3. We will again denote by e the inverse of $\frac{r}{\gcd(n,r)}$ modulo $\frac{n}{\gcd(n,r)}$. Note that since $\frac{n}{\gcd(n,r)}$ is even, e must be odd.

Proposition 3.3.13. *Let* $n \in \mathbb{N}$ *and* K_r *be the* r*-th Kasami exponent with* $\gcd(r, n) = d$, r *even,* $\frac{n}{d}$ *even and not divisible by* 3. *Then* $K_r^{-1} \equiv \sum_{i=0}^{d-1} \sum_{j=0}^{\frac{n}{d}-1} a_{i,j} 2^{i-jr} \pmod{2^n - 1}$ *where the values* $a_{i,j}$ *are the entries of the matrix*

$$M_{a,r} = \begin{pmatrix} a_1 \\ x \\ y \\ \vdots \\ x \\ y \\ x \end{pmatrix}.$$

where a_1, x, y *are as follows. We use the auxiliary sequences* $x_1 = (1, 1, 0, 0, 0, 1)$ *and* $x_2 = (1, 0, 0, 0, 1, 1)$ *of length 6.*

(a) If $\frac{n}{d} = 6k + 2$ *then*

$$a_1 = (1, 1, \underbrace{x_1, \ldots, x_1}_{k\text{-times}}), x = (1, 0, 1, 0, \ldots, 1, 0), y = (0, 1, 0, 1, \ldots, 0, 1).$$

(b) If $\frac{n}{d} = 6k + 4$ *then*

$$a_1 = (1, 0, 1, 1, \underbrace{x_2, \ldots, x_2}_{k\text{-times}}), x = (0, 1, 0, 1, \ldots, 0, 1), y = (1, 0, 1, 0, \ldots, 1, 0).$$

In both cases we have $\mathrm{wt}(K_r^{-1}) = \frac{n+2}{2}$.

Proof. Case (a): The r-matrix of the carry sequence is

$$M_{c,r} = (c_{i,j}) = \begin{pmatrix} 0 & 1 & 0 & 1 & \dots & 0 & 1 \\ 1 & 0 & 1 & 0 & \dots & 1 & 0 \\ \vdots & & & & & & \vdots \\ 0 & 1 & 0 & 1 & \dots & 0 & 1 \\ 1 & 0 & 1 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

We check Eq. (3.11)-(3.13) from Theorem 3.3.9.

Eq. (3.11) holds because $c_{0,0} = 0$, $c_{d-1,e} = 0$ (recall that e is odd) and $a_{0,2} = a_{0,1} = a_{0,0} = 1$.

We verify Eq. (3.12): If j is odd then $c_{0,j} = c_{d-1,j+e} = 1$ and $(a_{0,j}, a_{0,j+1}, a_{0,j+2}) \in \{(1,0,0),(1,1,1),(0,0,1)\}$. If $j > 0$ is even, then $c_{0,j} = c_{d-1,j+e} = 0$ and $(a_{0,j}, a_{0,j+1}, a_{0,j+2}) \in \{(1,1,0),(0,1,1),(0,0,0)\}$.

Lastly, we verify Eq. (3.13): If $i + j$ is even then $c_{i,j} = 0$, $c_{i-1,j} = 1$, $a_{i,j+2} = a_{i,j} = 0$ and $a_{i,j+1} = 1$. If $i + j$ is odd then $c_{i,j} = 1$, $c_{i-1,j} = 0$, $a_{i,j+2} = a_{i,j} = 1$ and $a_{i,j+1} = 0$.

Case (b): In this case, the r-matrix of the carry sequence is

$$M_{c,r} = (c_{i,j}) = \begin{pmatrix} 1 & 0 & 1 & 0 & \dots & 1 & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 & 1 \\ \vdots & & & & & & \vdots \\ 1 & 0 & 1 & 0 & \dots & 1 & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 & 1 \end{pmatrix}.$$

Eq. (3.11) is valid since $c_{0,0} = 1$, $c_{d-1,e} = 1$ and $a_{0,0} = a_{0,2} = 1$ and $a_{0,1} = 0$. The verification process for Eq. (3.12) and (3.13) is identical to Case (a) with odd and even swapped.

□

3.3.4 Kasami inverses with special structure

We now investigate cases where the inverses of Kasami exponents have some special structure. These cases will also illustrate the results in the previous sections and show how to get from the representation using r-matrices to the "usual" binary representation.

In [87, Proposition 3.13], it was shown that the inverse of K_r modulo $2^{5r} - 1$ is cyclotomic equivalent to the Kasami exponent K_{2r}. It was conjectured that K_r^{-1} modulo $2^{\frac{5r}{b}} - 1$ for $b|r$ and $5 \nmid b$ is always cyclotomic equivalent to a Kasami exponent. This conjecture can be proven using Proposition 3.3.11.

Proposition 3.3.14. *Let $d = \frac{r}{b}$ with $b|r$, $n = 5d$ and K_r^{-1} be the least positive residue of the inverse of K_r modulo $2^n - 1$. Then*

$$K_r^{-1} \equiv \begin{cases} 2^{2d} K_{2d} \pmod{2^n - 1} & \text{if } b \equiv 1 \pmod 5 \\ 2^{2d} K_d \pmod{2^n - 1} & \text{if } b \equiv 2 \pmod 5 \\ 2^{2(d-r)} K_d \pmod{2^n - 1} & \text{if } b \equiv 3 \pmod 5 \\ 2^{2(d-r)} K_{2d} \pmod{2^n - 1} & \text{if } b \equiv 4 \pmod 5. \end{cases}$$

Proof. We use the notation of Proposition 3.3.11. We have $d = \gcd(n, r) = \frac{r}{b}$ and $\frac{n}{d} = 5$. Further, we have $\frac{r}{d} = b$. The only two possible odd values for e are $e = 1$ and

$e = 3$ that are attained for $b \equiv 1 \pmod 5$ and $b \equiv 2 \pmod 5$, respectively. These correspond to case (b) and (f) in Proposition 3.3.11. We get $K_r^{-1} \equiv \sum_{i=0}^{d-1} \sum_{j=0}^{4} a_{i,j} 2^{i-jr} \pmod{2^n - 1}$ where the values $a_{i,j}$ are the entries of the matrix M_1 if $e = 1$ and M_2 if $e = 3$:

$$M_1 = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ \vdots & & & & \vdots \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}, M_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

We now write K_r^{-1} in its usual binary representation. To do this, we write from right to left in the following way: We start with the first column, and then proceed in steps of length e to the left (cyclically). So, for the case $e = 1$, we start with column 0 of M_1, then column 4, then 3, then 2 and then 1, resulting in:

$$K_r^{-1} = (\underbrace{1, 1, \ldots, 1, 1}_{d\text{-times}}, \underbrace{0, 0, \ldots, 0, 0}_{2d-1\text{-times}}, 1, \underbrace{0, 0, \ldots, 0, 0}_{d\text{-times}}, \underbrace{1, 1, \ldots, 1, 1}_{d\text{-times}})$$

and for the case $e = 3$ the order of the columns is $0, 2, 4, 1, 3$, resulting in:

$$K_r^{-1} = (\underbrace{0, 0, \ldots, 0, 0}_{d\text{-times}}, \underbrace{1, 1, \ldots, 1, 1}_{d\text{-times}}, \underbrace{0, 0, \ldots, 0, 0}_{d-1\text{-times}}, 1, \underbrace{0, 0, \ldots, 0, 0}_{2d\text{-times}}).$$

In the first case, we have $K_r^{-1} \equiv 2^{2d} K_{2d} \pmod{2^n - 1}$ and in the second case $K_r^{-1} \equiv 2^{2d} K_d \pmod{2^n - 1}$. If $e = 2$ and $e = 4$ (corresponding to the values $b \equiv 3 \pmod 5$ and $b \equiv 4 \pmod 5$) we use the relation $K_r^{-1} \equiv 2^{-2r} K_{n-r}^{-1} \pmod{2^n - 1}$ and apply the procedure above to K_{n-r}.

□

In fact, in [87] several nice formulas for the inverses of K_r modulo $2^{kr} - 1$ for small fixed values of k have been found. Our framework gives an explanation why these inverses have a strong structure: We have $\frac{kr}{\gcd(r,kr)} = k$, so the r-matrices always have k columns. By Proposition 3.3.10 and 3.3.11, all but one row in the r-matrix are identical, so we get long runs of zeroes and ones (as observed in the proof of Proposition 3.3.14). All of the formulas given in [87] can also be obtained using our framework. In particular, it was shown in [87] that if $n = \frac{3r}{b}$ with $b|r$ and $\gcd(3, b) = 1$ then the inverse of K_r modulo $2^n - 1$ has the lowest possible weight 2. Using the results we obtained in the previous sections, we give an alternative proof and show additionally that (apart from sporadic cases for low values of n) these are the only cases where the inverses of Kasami exponents have weight 2.

Proposition 3.3.15. *Let K_r be invertible modulo $2^n - 1$ with $n \geq 6$ and K_r^{-1} be the least positive residue of the inverse of K_r modulo $2^n - 1$. Then $\mathrm{wt}(K_r^{-1}) = 2$ if and only if $n = \frac{3r}{b}$ with $b|r$ and $\gcd(b, 3) = 1$. In these cases we have*

$$K_r^{-1} \equiv \begin{cases} 2^{n-1} + 2^{\frac{n}{3}-1} \pmod{2^n - 1} & \text{if } b \equiv 1 \pmod 3 \\ 2^{n-1} + 2^{\frac{2n}{3}-1} \pmod{2^n - 1} & \text{if } b \equiv 2 \pmod 3. \end{cases}$$

Proof. We go through the results in the earlier sections and check when $\mathrm{wt}(K_r^{-1}) = 2$ is fulfilled. In Proposition 3.3.10, we have $\mathrm{wt}(K_r^{-1}) = \frac{n-3d+4}{2}$ where $d = \gcd(r, n)$. We have $\frac{n-3d+4}{2} = 2$ if and only if $n = 3d$. So, $n = \frac{3r}{b}$ for some b with $\gcd(b, 3) = 1$. We differentiate the two possible cases $e = 1$ and $e = 2$ corresponding to $b \equiv 1 \pmod 3$ and $b \equiv 2 \pmod 3$, respectively. If $e = 1$, we are in Case (a) of

Proposition 3.3.10 and the matrix $M_{a,r}$ looks as follows:

$$M_{a,r} = \begin{pmatrix} 0 & 0 & 0 \\ \vdots & & \vdots \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Consequently, $K_r^{-1} \equiv 2^{\frac{n}{3}-1} + 2^{\frac{n}{3}-1-r} \equiv 2^{n-1} + 2^{\frac{n}{3}-1} \pmod{2^n - 1}$. Here we used that $r \equiv \frac{n}{3} \pmod{n}$ since $b \equiv 1 \pmod 3$. If $e = 2$, we apply the same procedure to K_{n-r}, so $K_{n-r}^{-1} \equiv 2^{\frac{n}{3}-1} + 2^{\frac{n}{3}-1-(n-r)} \equiv 2^{n-1} + 2^{\frac{n}{3}-1} \pmod{2^n - 1}$ since here $r \equiv \frac{2n}{3} \pmod{n}$. Then $K_r^{-1} \equiv 2^{-2r} K_{n-r}^{-1} \equiv 2^{n-1} + 2^{\frac{2n}{3}-1} \pmod{2^n - 1}$.

We now check Proposition 3.3.11. In the Cases (a)-(d) we have $\mathrm{wt}(K_r^{-1}) = \frac{n-d+2}{2}$, so $\mathrm{wt}(K_r^{-1}) = 2$ if and only if $d = n - 2$. Since $d|n$ and $n > 4$, this is not possible.

In the Cases (e) and (h) we have (using the notation from the proposition) $\mathrm{wt}(K_r^{-1}) = \frac{n-d(s+1)+2}{2}$, so $\mathrm{wt}(K_r^{-1}) = 2$ if and only if $n - d(s+1) = 2$. Since $d|n$, this implies $d|2$. Using the bound in Corollary 3.3.12, we infer that $\mathrm{wt}(K_r^{-1}) > 2$ if $n \geq 6$. In the Cases (f) and (g) we have $\mathrm{wt}(K_r^{-1}) = \frac{n-d(s+2)+2}{2}$. Again we get $d|2$ and the same argument as before yields $\mathrm{wt}(K_r^{-1}) > 2$.

In Proposition 3.3.13 the inverses have always binary weight $\frac{n+2}{2}$, so no new cases are found.

□

Note that the condition $n \geq 6$ is necessary. Indeed, for $n = 5$ we get sporadic cases: Consider $K_2 = 13$ over $\mathbb{F}_{2^5}$. We have $\gcd(5, 2) = 1$ and $2 \cdot 3 \equiv 1 \pmod 5$, so $e = 3$ and the inverse of 13 modulo $2^5 - 1$ has weight 2 by Corollary 3.3.7.

3.4 The Bracken-Leander exponent

We now determine the inverse of the Bracken-Leander exponent $BL_r = 2^{2r} + 2^r + 1$ modulo $2^{4r} - 1$ with r odd. In this case, the exponent is not independent from the field size. Because of this, finding the inverse is much easier. We again use the modular add-with-carry approach. Theorem 3.1.1 applied to the Bracken-Leander exponents yields the following condition for the carry sequence.

Theorem 3.4.1. *Let r odd, $n = 4r$, $a \in \{1, \ldots, 2^n - 2\}$ and BL_r be the Bracken-Leander exponent. We denote by $a = (a_{n-1}, \ldots, a_0)$ the binary expansion of a. The following are equivalent:*

(a) a is the inverse of BL_r modulo $2^n - 1$.

(b) There exists a carry sequence $c = (c_{n-1}, \ldots, c_0)$ with $c_i \in \{0, 1, 2\}$ such that

$$2c_0 - c_{-1} + 1 = a_{-2r} + a_{-r} + a_0 \tag{3.21}$$
$$2c_i - c_{i-1} = a_{i-2r} + a_{i-r} + a_i \text{ for all } i > 0. \tag{3.22}$$

Here, the indices are seen as elements in $\mathbb{Z}_n$.

The carry sequence in (b) is unique.

Observe that $\gcd(r, n) = r$ and $\frac{n}{\gcd(r,n)} = 4$. The case here is thus similar to the $\frac{n}{\gcd(r,n)}$ even case of the Kasami functions. We again use r-matrices so that Eq. (3.21) and (3.22) have an easier structure.

Theorem 3.4.2. *Let r odd, $n = 4r$, $a \in \{1, \ldots, 2^n - 2\}$ and $BL_r = 2^{2r} + 2^r + 1$ be the Bracken-Leander exponent. We denote by $a = (a_{n-1}, \ldots, a_0)$ the binary expansion of a. Moreover, let*

$$M_{a,r} = \begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} \\ a_{1,0} & a_{1,1} & a_{1,2} & a_{1,3} \\ \vdots & & & \vdots \\ a_{r-1,0} & a_{r-1,1} & a_{r-1,2} & a_{r-1,3} \end{pmatrix}$$

be the r-matrix of a, i.e. $a \equiv \sum_{i=0}^{r-1} \sum_{j=0}^{3} a_{i,j} 2^{i-jr} \pmod{2^n - 1}$. The following are equivalent:

(a) a is the inverse of K_r modulo $2^n - 1$.

(b) There exists an r-matrix for the carry sequence c of the form

$$M_{c,r} = \begin{pmatrix} c_{0,0} & c_{0,1} & c_{0,2} & c_{0,3} \\ c_{1,0} & c_{1,1} & c_{1,2} & c_{1,3} \\ \vdots & & & \vdots \\ c_{r-1,0} & c_{r-1,1} & c_{r-1,2} & c_{d-1,3} \end{pmatrix}$$

with $c_{i,j} \in \{0, 1, 2\}$ such that the following equations hold:

$$2c_{0,0} - c_{r-1,1} + 1 = a_{0,2} + a_{0,1} + a_{0,0} \tag{3.23}$$

$$2c_{0,j} - c_{r-1,j+1} = a_{0,j+2} + a_{0,j+1} + a_{0,j} \text{ for } j \in \{1, 2, 3\} \tag{3.24}$$

$$2c_{i,j} - c_{i-1,j} = a_{i,j+2} + a_{i,j+1} + a_{i,j} \text{ for all } i \in \{1, \ldots, r-1\}, j \in \{0, 1, 2, 3\}. \tag{3.25}$$

The carry sequence (and thus its associated r-matrix) in (b) is unique.

It is easy to derive some strong necessary conditions from the equations. For example Eq. (3.25) implies that, if $c_{i,j} = 0$ for some $i > 0$, then necessarily $c_{i-1,j} = a_{i,j+2} = a_{i,j+1} = a_{i,j} = 0$, which inductively leads to $c_{i',j} = a_{i',j+2} = a_{i',j+1} = a_{i',j} = 0$ for all $0 < i' < i$. With some examples for small values of n, it is then quite easy to guess the correct r-matrices for the sequence a and its associated carry sequence c.

Proposition 3.4.3. *Let r odd, $n = 4r$ and $BL_r = 2^{2r} + 2^r + 1$ be the Bracken-Leander exponent. Then $BL_r^{-1} \equiv \sum_{i=0}^{r-1} \sum_{j=0}^{3} a_{i,j} 2^{i-jr} \pmod{2^n - 1}$ where the values $a_{i,j}$ are the entries of the matrix*

$$M_{a,r} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ \vdots & & & \vdots \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

We have $\mathrm{wt}(BL_r^{-1}) = \frac{n+2}{2}$.

Proof. The r-matrix of the corresponding carry sequence is

$$(c_{i,j}) = \begin{pmatrix} 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \\ \vdots & & & \vdots \\ 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \end{pmatrix}.$$

We verify Eq. (3.23)-(3.25). Eq. (3.23) holds because $c_{0,0} = c_{r-1,1} = 2$ and $a_{0,0} = a_{0,1} = a_{0,2} = 1$. Eq. (3.24) holds because $c_{0,j} = c_{r-1,j+1} = 2$ and $(a_{0,j}, a_{0,j+1}, a_{0,j+2}) \in \{(1,1,0), (1,0,1), (0,1,1)\}$ if $j \in \{1,2,3\}$.

It only remains to check Eq. (3.25). For i odd, we have $c_{i,j} = 1$, $c_{i-1,j} = 2$ and $a_{i,j} = a_{i,j+1} = a_{i,j+2} = 0$. For $i > 0$ even, we have $c_{i,j} = 2$, $c_{i-1,j} = 1$ and $a_{i,j} = a_{i,j+1} = a_{i,j+2} = 1$, so Eq. (3.25) is satisfied.

To determine $\mathrm{wt}(BL_r^{-1})$, we count the number of ones in $M_{a,r}$, so $\mathrm{wt}(BL_r^{-1}) = 4\frac{r+1}{2} - 1 = \frac{n+2}{2}$.

□

3.5 Conclusion

In this chapter, we introduced a new approach to find inverses of elements in $\mathbb{Z}_{2^n-1}$, using the modular add-with-carry approach. We determined the inverse of all Gold exponents $G_r = 2^r + 1$ and Kasami exponents $K_r = 2^{2r} - 2^r + 1$ modulo $2^n - 1$ (if they exist) as well as the inverse of the Bracken-Leander exponent $BL_r = 2^{2r} + 2^r + 1$ modulo $2^{4r} - 1$ with r odd. With our contribution, the binary representations of the inverses of all known APN exponents as well as the inverses of all exponents that give rise to 4-differentially uniform permutations in even dimension are found. The more general problem of inverting a given element l in $\mathbb{Z}_{2^n-1}$ for all n is still not well understood. It is a natural question if the approach using the modular add-with-carry algorithm can be generalized to other exponents. For every invertible l, we can find a defining set of equations for the binary representation of l^{-1} and the corresponding carry sequence in the style of Eq. (3.1) in Theorem 3.1.1. The difficulty then lies in finding the sequences that satisfy the equations. This has to be done on a case by case basis.

Inversion in $\mathbb{Z}_{2^n-1}$ is not only interesting for questions relating to differential uniformity. For example, if l is a *complete permutation polynomial (CPP) exponent over* $\mathbb{F}_q$ (i.e. there exists an $a \in \mathbb{F}_q$ such that ax^l and $ax^l + x$ are permutation polynomials), then also its inverse r^{-1} modulo $q - 1$ is a CPP exponent [112]. Several CPP exponents in even characteristic have been found (e.g. [39, 131, 136]). For a complete classification of CPP exponents, finding explicit formulas for the corresponding inverses is an interesting research problem.

The modular add-with-carry approach can be easily modified to work also in the ring $\mathbb{Z}_{p^n-1}$ for a prime $p > 2$ [64, Theorem 4.1]. In particular, it can be used to tackle the problem of inversion in $\mathbb{Z}_{p^n-1}$ (corresponding to inversion of monomials in odd characteristic). However, the equations in the style of Eq. (3.1) that have to be checked become more complicated.

Chapter 4

Equivalences of monomials and permutations of the form $L_1(x^d) + L_2(x)$

4.1 A connection between equivalences of vectorial Boolean functions and permutation polynomials of the form $L_1(x^d) + L_2(x)$

The results in this section as well as the next section of this chapter on the inverse function are based on two papers [57, 81], (co)-written by the author of this thesis.

In Chapter 2 we introduced the three main equivalence relations between vectorial Boolean functions: Affine equivalence, EA-equivalence and CCZ-equivalence. In particular, the most general equivalence, CCZ-equivalence, is in many ways still not very well understood.

Following the notation used in [27], let us denote by the *EA-class* of F the set of all functions EA-equivalent to an (n, n)-function F, and similarly by *CCZ-class* of F the set of all functions CCZ-equivalent to F. Since EA-equivalence is a special case of CCZ-equivalence, we can partition a CCZ-class into EA-classes. Experimental results show that for many functions $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ the CCZ-class of F coincides with its EA-class if F is not a permutation. If F is a permutation, then its CCZ-class often consists of precisely 2 EA-classes, with F and F^{-1} being representatives for the two EA-classes. Of course, it is possible to have more than two EA-classes in one CCZ-class, this is for example the case for Gold functions in odd dimension [19, Theorem 1]. Additionally, it is possible that a permutation and its inverse are in the same EA-class, this happens for example naturally for involutions like the inverse function $x \mapsto x^{2^n-2}$ over $\mathbb{F}_{2^n}$.

Since CCZ-equivalence is a much more difficult concept than EA-equivalence, it is desirable to understand when the CCZ-class of F contains only one (if F is not a permutation) or two (if F is a permutation) EA-classes.

This is a problem that is open even for most of the APN monomials in Table 2.1. As noted earlier, the CCZ-class of Gold functions in odd dimension contains more than 2 EA-classes and thus cannot be described with EA-equivalence and the inverse transformation alone. For all other monomials, this question is still open. Recently, Budaghyan, Calderini and Villa conjectured the following based on a computer search for small values of n:

Conjecture 4.1.1 ([18, Conjecture 4.14]). Let $F(x) = x^d$ be a non-Gold APN power function or the inverse function over $\mathbb{F}_{2^n}$. Then, every function that is CCZ-equivalent to F is EA-equivalent to F or to F^{-1} (if it exists).

In this chapter, we are going to confirm the conjecture in the case of the inverse function, i.e. we show that the CCZ-class of the function $x \mapsto x^{2^n-2}$ on $\mathbb{F}_{2^n}$ coincides with its EA-class. This is to our knowledge the first theoretical result of this kind. Note that, in many ways, the inverse function is actually the most interesting case because of its widespread use in cryptography, most famously as the S-box in AES.

Because of the importance of permutations for the design of SPNs, it is also interesting to search for permutations inside the CCZ-class of a function F. Indeed, a way to find permutations with good cryptographic properties is to look for a permutation in the CCZ-class of a non-permutation with good cryptographic properties. This is precisely the technique that was used to find the only known APN permutation in even dimension [17]. Thus it is a very interesting question to classify all permutations that are in the CCZ-class of an APN function. Treating this problem for infinite families is however very difficult. To the authors knowledge, the only result in this direction is found in a recent paper [59], where it was shown that there are no permutations in the CCZ-class of the APN Gold functions over $\mathbb{F}_{2^n}$ with n even and that there are no permutations in the CCZ-class of APN Kasami functions over $\mathbb{F}_{2^n}$ if n is divisible by 4. The proof technique used in [59] relies on a careful analysis of the bent component functions of the Gold and Kasami functions.

In this chapter, we will also classify all permutations in the CCZ-class of the inverse function (both in odd and even dimension). We show that (excluding some sporadic cases in low dimension) the only permutations in the CCZ-class are the "trivial" ones, i.e. the ones that are affine equivalent to the inverse function. Of course, since the inverse function does not have any bent components, our approach is necessarily different from the approach in [59]. Instead, we are going to use the following proposition which connects both of the problems we mentioned to a special type of permutation polynomial. We will later also use the same technique to investigate the functions of the form $x \mapsto x^{2^k-1}$ over $\mathbb{F}_{2^{2k}}$.

Proposition 4.1.2. (a) *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ and assume no permutation of the form $F(x) + L(x)$ exists with non-zero linear $L(x)$. Then every permutation that is EA-equivalent to F is already affine equivalent to it. In particular, if such an F is not bijective, then there are no EA-equivalent permutations to F.*

(b) *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ and assume no permutation of the form $L_1(F(x)) + L_2(x)$ exists with non-zero linear L_1, L_2. Then every function that is CCZ-equivalent to F is EA-equivalent to F or F^{-1} (if it exists). Moreover, all permutations that are CCZ-equivalent to F are affine equivalent to F or F^{-1}.*

Proof. (a) Let F_2 be a permutation EA-equivalent to F. Then, there exist $(a, b) \in \mathbb{F}_{2^n}^2$ and a bijective mapping $\mathcal{L}\colon \mathbb{F}_{2^n}^2 \to \mathbb{F}_{2^n}^2$ defined by $\mathcal{L}(x, y) = (\alpha(x), \gamma(x) + \delta(y))$ with linear functions $\alpha, \gamma, \delta\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ such that

$$\mathcal{L}(x, F(x)) + (a, b) = (\alpha(x) + a, \gamma(x) + \delta(F(x)) + b) = (\pi(x), F_2(\pi(x)))$$

where $\pi\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is the permutation given by $\pi(x) = \alpha(x) + a$. Note that the function δ is bijective on $\mathbb{F}_{2^n}$, since $\mathcal{L}$ is bijective on $\mathbb{F}_{2^n}^2$. Also the composition $F_2(\pi(x))$ is bijective on $\mathbb{F}_{2^n}$, implying that $\gamma(x) + \delta(F(x))$ is bijective, and hence also $\delta^{-1}(\gamma(x)) + F(x)$ is a permutation. Since $\delta^{-1}(\gamma(x))$ is linear, our assumption on F yields that $\gamma = 0$, so F_2 is affine equivalent to F.

(b) Let now F_2 be a function CCZ-equivalent to F. By the definition of CCZ-equivalence, there exist $(a, b) \in \mathbb{F}_{2^n}^2$ and a bijective mapping $\mathcal{L}\colon \mathbb{F}_{2^n}^2 \to \mathbb{F}_{2^n}^2$ given by

$\mathcal{L}(x,y) = (\alpha(x) + \beta(y), \gamma(x) + \delta(y))$ with linear $\alpha, \beta, \gamma, \delta \colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ such that

$$\begin{aligned}\mathcal{L}(x, F(x)) + (a, b) &= (\alpha(x) + \beta(F(x)) + a, \gamma(x) + \delta(F(x)) + b) \\ &= (\pi(x), F_2(\pi(x)))\end{aligned}$$

where $\pi \colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is the permutation on $\mathbb{F}_{2^n}$ given by $\pi(x) = \alpha(x) + \beta(F(x)) + a$. By our assumption on F, either $\alpha = 0$ or $\beta = 0$. Assume first that $\alpha = 0$. Then $\pi(x) = \beta(F(x)) + a$ and in particular both F and β are bijective. Further, γ is bijective since $\mathcal{L}$ is bijective. We then have

$$\gamma(x) + \delta(F(x)) + b = F_2(\pi(x)) = F_2(\beta(F(x)) + a).$$

The composition with the inverse $F^{-1}(x)$ yields

$$\gamma(F^{-1}(x)) + \delta(x) + b = F_2(\beta(x) + a),$$

and hence F_2 is EA-equivalent to F^{-1}. In the case $\beta = 0$ we get similarly $\pi(x) = \alpha(x) + a$ and

$$\gamma(x) + \delta(F(x)) + b = F_2(\pi(x)) = F_2(\alpha(x) + a),$$

where the mappings α and δ are bijective. Hence F_2 is EA-equivalent to F.

Now assume that F_2 is additionally a permutation. If F_2 is EA-equivalent to F then F_2 is affine equivalent to F using the statement in (a). Let us now consider the case that F_2 is EA-equivalent to F^{-1}. Observe that $F^{-1}(x) + L(x)$ is a permutation if and only if $L(F(x)) + x$ is a permutation, so there are no permutations of the form $F^{-1}(x) + L(x)$ by the assumption stated in the proposition. Again using (a), we conclude that F_2 is affine equivalent to F^{-1}.

□

Proposition 4.1.2 shows that very strong conclusions can be drawn when no permutations of the form $L_1(F(x)) + L_2(x)$ with $L_1 \neq 0$ and $L_2 \neq 0$ exist. The rest of this chapter is devoted to polynomials of this form.

Permutations of form $L_1(F(x)) + L_2(x)$ are characterized for some special choices of F and L_1, L_2. It was shown in [37] that no permutation of the form $x^d + L(x)$ exists when there is an $a \in \mathbb{F}_{2^n}$ such that $\mathrm{Tr}(ax^d)$ is bent. Corollary 2.3 from [55] implies that $x^d + L(K(x))$ is not bijective on $\mathbb{F}_q$ for an arbitrary function K whenever $\gcd(d, q-1) \neq 1$ and L is a non-bijective linear function. In [99] a characterization of all permutations of the form $x^{2^i+1} + L(x)$ over $\mathbb{F}_{2^n}$ with $\gcd(i, n) = 1$ was given, as well as some results for the more general case $x^d + L(x)$. Permutations of the form $x^{2^i+1} + L(x)$ over $\mathbb{F}_{2^n}$ with $\gcd(i, n) > 1$ were recently considered in [15].

We start with a proposition that gives a criterion when such a polynomial is a permutation.

Proposition 4.1.3. *Let $F \colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ and L_1, L_2 be linear mappings. The function $L_1(F(x)) + L_2(x)$ is a permutation if and only if*

$$W_F(L_2^*(b), L_1^*(b)) = 0$$

for all $b \in \mathbb{F}_{2^n}^*$.

Proof. By Proposition 2.2.2, a function is a permutation if and only if all of its component functions are balanced. Consequently, $L_1(F(x)) + L_2(x)$ is a permutation if

and only if

$$0 = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(L_1(F(x))+L_2(x)))}$$
$$= \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(L_1^*(b)F(x)+L_2^*(b)x)} = W_F(L_2^*(b), L_1^*(b))$$

for all $b \in \mathbb{F}_{2^n}^*$. □

4.2 The inverse function

A particularly interesting case are functions of the type $L_1(x^{2^n-2}) + L_2(x)$ on $\mathbb{F}_{2^n}$ because of the good cryptographic properties (nonlinearity/differential uniformity) of the inverse function $x \mapsto x^{2^n-2}$. For simplicity, we will refer to the inverse function also as $x \mapsto x^{-1}$, where we use as usual the convention $0^{-1} = 0$. It was shown in [60] that $L_1(x^{-1}) + L_2(x)$ is never a permutation in characteristic ≥ 5 (except for the trivial cases $L_1 = 0$ or $L_2 = 0$). In characteristic 3, no permutations of the type $x^{-1} + L(x)$ with $L \neq 0$ exist, except for sporadic cases in the small fields $\mathbb{F}_3$ and $\mathbb{F}_9$. In this section we are only interested in the case of characteristic 2. The following partial results was already obtained in [100].

Theorem 4.2.1 ([100]). *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be defined by $F(x) = x^{-1} + L(x)$ with some linear mapping $L(x) \neq 0$. If $n \geq 5$ then F is not a permutation.*

The following result is an immediate consequence of Theorem 4.2.1.

Corollary 4.2.2. *Let $n \geq 5$ and $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be defined by $F(x) = L_1(x^{-1}) + L_2(x)$, where L_1, L_2 are non-zero linear functions of $\mathbb{F}_{2^n}$. If L_1 or L_2 is bijective, then F is not a permutation on $\mathbb{F}_{2^n}$.*

Proof. Note that $F(x)$ is bijective if and only if $F(x^{-1}) = L_1(x) + L_2(x^{-1})$ is so. Hence without loss of generality suppose L_1 is bijective. Then the composition $L_1^{-1}(F(x)) = x^{-1} + L_1^{-1}(L_2(x))$ is bijective if and only if F is so, and Theorem 4.2.1 completes the proof. □

We continue the study of functions $L_1(x^{-1}) + L_2(x)$ where L_1, L_2 are linear polynomials over $\mathbb{F}_{2^n}$.

Let us briefly introduce some notation that we will use in this section. For a set $A \subseteq \mathbb{F}_{2^n}$ we denote by $1/A$ the set of all inverses of A, i.e. $1/A = \{1/a\colon a \in A \setminus \{0\}\}$. Further, we denote by $H_a = \{x \in \mathbb{F}_{2^n}\colon \mathrm{Tr}(ax) = 0\}$ with $a \neq 0$ the hyperplanes of $\mathbb{F}_{2^n}$, by $A \cdot A$ the product set $A \cdot A = \{a_1 a_2 \mid a_1, a_2 \in A\}$ and by $\sqrt{A} = \{\sqrt{a} \mid a \in A\}$. Note that since we are working in fields of characteristic 2, the function $x \mapsto x^2$ is bijective, so $|\sqrt{A}| = |A|$. We will also use the following well-known lemma about the adjoint mapping, whose simple proof we include for the sake of completeness.

Lemma 4.2.3. *Let $L\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be a linear mapping and L^* its adjoint mapping. Then* $\dim(\mathrm{im}(L^*)) = \dim(\mathrm{im}(L))$ *and* $\dim(\ker(L^*)) = \dim(\ker(L))$.

Proof. Let $v \in \mathrm{im}(L^*)$ and $w \in \ker(L)$. We can write $v = L^*(x)$ for some $x \in \mathbb{F}_{2^n}$. Then $\langle v, w\rangle = \langle L^*(x), w\rangle = \langle x, L(w)\rangle = \langle x, 0\rangle = 0$, so $\mathrm{im}(L^*) \subseteq \ker(L)^\perp$, in particular $\dim(\mathrm{im}(L^*)) \leq \dim(\mathrm{im}(L))$. The other inequality holds with $L^{**} = L$.

The statement on the kernel follows from $\dim(\mathrm{im}(L)) + \dim(\ker(L)) = n$. □

In the case of the inverse function $x \mapsto x^{-1}$, the Walsh transform is closely connected to Kloosterman sums.

Definition 4.2.4 (Kloosterman sum)**.** *For $a \in \mathbb{F}_{2^n}$, the Kloosterman sum of a over $\mathbb{F}_{2^n}$ is defined as*

$$K_n(a) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(x^{-1}+ax)}.$$

An element $a \in \mathbb{F}_{2^n}$ with $K_n(a) = 0$ is called a Kloosterman zero.

Clearly, $K_n(0) = 0$. We call 0 the trivial Kloosterman zero. Note $K_n(a) = W_F(1, a)$ for $F(x) = x^{-1}$. More precisely, for $a \neq 0$ we have

$$W_F(a, b) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(ax^{-1}+bx)} = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(x^{-1}+abx)} = K_n(ab)$$

using the substitution $x \mapsto ax$. For $a = 0$ and $b \neq 0$, we have $K_n(ab) = W_F(a, b) = 0$.

Proposition 4.1.3 can thus be stated using Kloosterman sums:

Corollary 4.2.5. *Let L_1, L_2 be linear functions of $\mathbb{F}_{2^n}$. Then $L_1(x^{-1}) + L_2(x)$ is a permutation on $\mathbb{F}_{2^n}$ if and only if $\ker(L_1^*) \cap \ker(L_2^*) = \{0\}$ and*

$$K_n(L_1^*(b)L_2^*(b)) = 0$$

for all $b \in \mathbb{F}_{2^n}$.

Proof. By Proposition 4.1.3, $L_1(x^{-1}) + L_2(x)$ is a permutation if and only if $W_F(L_1^*(b), L_2^*(b)) = 0$ for all $b \neq 0$. If $b \in \ker(L_1^*) \cap \ker(L_2^*)$, then $W_F(L_1^*(b), L_2^*(b)) = 2^n \neq 0$. In the other cases $W_F(L_1^*(b), L_2^*(b)) = K(L_1^*(b)L_2^*(b))$ by the considerations above. □

Remark 4.2.6. Corollary 4.2.5 shows that a function $L_1(x^{-1}) + L_2(x)$ is bijective on $\mathbb{F}_{2^n}$ only if the set $\{L_1^*(x)L_2^*(x) | x \in \mathbb{F}_{2^n}\}$ is a subset of the set of Kloosterman zeroes. Conversely, in [69] specific functions of shape $L_1(x^{-1}) + L_2(x)$ are used to obtain identities for Kloosterman sums.

Kloosterman sums provide a powerful tool for studying additive properties of the inversion on finite fields. Moreover, Kloosterman zeros are used for the construction of bent and hyperbent functions (see for example [44, 35]).

Few results about the distribution of Kloosterman zeros are known. However, the precise spectrum of Kloosterman sums in characteristic 2 has been determined.

Theorem 4.2.7 ([88])**.** *The set of Kloosterman sums $\{K_n(a) : a \in \mathbb{F}_{2^n}^*\}$ is precisely the set of all integers divisible by 4 that are contained in the the interval $[1 - 2^{n/2+1}, 1 + 2^{n/2+1}]$.*

This theorem in particular shows that nontrivial Kloosterman zeros exist for all n. We want to note that a similar result exists also for Kloosterman sums in characteristic 3 [75], but not in larger characteristic. Indeed, if the characteristic is ≥ 5, then there exist no nontrivial Kloosterman zeros at all [85].

There is a way to compute the number of Kloosterman zeros [88], which relies on determining the class number of binary quadratic forms. However, it is difficult to use this method to derive a theoretical result on the number and distribution of Kloosterman zeros. It is also known that for $n > 4$, nontrivial Kloosterman zeros are never contained in proper subfields of $\mathbb{F}_{2^n}$ [103]. In [124], it is noted that $|\{a \in \mathbb{F}_{2^n} : K_n(a) = 0\}| = O(2^{3n/4})$.

Because of the chaotic distribution of Kloosterman zeros, we will instead use dyadic approximations of Kloosterman sums. A nice survey on this topic is given in [141]. A main tool for our results in this section is the following characterization of Kloosterman sums divisible by 2^4.

Let $Q\colon \mathbb{F}_{2^n} \to \mathbb{F}_2$ be the mapping defined by

$$Q(x) = \sum_{0 \le i < j < n} x^{2^i + 2^j}$$

for all $x \in \mathbb{F}_{2^n}$.

Theorem 4.2.8 ([61]). *Let $n \geq 4$ and $a \in \mathbb{F}_{2^n}$. Then $K_n(a) \equiv 0 \pmod{16}$ if and only if $\mathrm{Tr}(a) = 0$ and $Q(a) = 0$.*

We want to note that stronger dyadic approximations of Kloosterman sums have been found, for a complete characterization of Kloosterman sums modulo $2^8 = 256$, see [58]. However, the additional precision provided cannot be used to simplify or improve the approach given here.

Applying Theorem 4.2.8 to Corollary 4.2.5 gives a necessary condition for $L_1(x^{-1}) + L_2(x)$ to be a permutation.

Corollary 4.2.9. *If $L_1(x^{-1}) + L_2(x)$ is a permutation of $\mathbb{F}_{2^n}$ with $n \geq 4$, then $\mathrm{Tr}(L_1^*(a)L_2^*(a)) = Q(L_1^*(a)L_2^*(a)) = 0$ for all $a \in \mathbb{F}_{2^n}$ and $\ker(L_1^*) \cap \ker(L_2^*) = \{0\}$.*

Note that Q is a quadratic form over $\mathbb{F}_2$, i.e. the mapping $B_Q(x,y) = Q(x+y) + Q(x) + Q(y)$ is a bilinear form. Indeed, we have

$$\begin{aligned} B_Q(x,y) &= \sum_{0 \le i < j < n} x^{2^i+2^j} + \sum_{0 \le i < j < n} y^{2^i+2^j} + \sum_{0 \le i < j < n} (x+y)^{2^i+2^j} \\ &= \sum_{i \neq j} x^{2^i} y^{2^j} = \sum_i x^{2^i} \sum_{j \neq i} y^{2^j} \\ &= \sum_i x^{2^i} (\mathrm{Tr}(y) + y^{2^i}) = \sum_i (xy)^{2^i} + \mathrm{Tr}(y) \sum_i x^{2^i} = \mathrm{Tr}(xy) + \mathrm{Tr}(x)\,\mathrm{Tr}(y). \end{aligned}$$

We call B_Q the bilinear form associated to Q.

We now prove that no permutations of the form $L_1(x^{-1}) + L_2(x)$ with $L_1, L_2 \neq 0$ exist if $n \geq 5$. Our starting point are the conditions given in Corollary 4.2.9. Our proof consists of three steps:

- We show that if $L_1(x^{-1}) + L_2(x)$ permutes $\mathbb{F}_{2^n}$, then the kernels of L_1 and L_2 must be translates of a subfield of $\mathbb{F}_{2^n}$, i.e. of the form $a\mathbb{F}_{2^k}$ with $a \in \mathbb{F}_{2^n}^*$ and $k|n$.
- Under this condition, we show $k = 1$, i.e. that $\dim(\ker L_1) = \dim(\ker L_2) = 1$.
- We show explicitly that there are no permutations with $\dim(\ker L_1) = \dim(\ker L_2) = 1$.

The key step here is the first one. Generally, the difficulty of the problem lies in the fact that the inverse function does not preserve the additive structure given by the linear mappings. However, if $\ker(L_1) = a\mathbb{F}_{2^k}$, then $1/\ker(L_1) \cup \{0\} = \frac{1}{a}\mathbb{F}_{2^k}$, so the kernel retains its structure after inversion, which is the key for the next steps.

We will use the following result from additive combinatorics which characterizes the subsets in an Abelian group with doubling constant 1.

Theorem 4.2.10 ([130, Proposition 2.7]). *Let $(G,\cdot)$ be an Abelian group and $A \subseteq G$ a finite subset of G. Then $|A \cdot A| = |A|$ if and only if $A = gH$, where $g \in G$ and $H \leq G$ is a subgroup of G.*

We are now ready to prove the first step we outlined above.

Theorem 4.2.11. *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ with $n \geq 5$ be defined by $F(x) = L_1(x^{-1}) + L_2(x)$, where L_1, L_2 are nonzero linear mappings over $\mathbb{F}_{2^n}$. If F is a permutation, then $\ker L_1$ and $\ker L_2$ are translates of subfields of $\mathbb{F}_{2^n}$, i.e. they are of the form $a\mathbb{F}_{2^k}$ for $a \neq 0$ and $k|n$. Moreover, we have $\ker L_1 = L_2^*(\ker(L_1^*))$ and $\ker L_2 = L_1^*(\ker(L_2^*))$.*

Proof. We define $R(x) = L_1^*(x)L_2^*(x)$. Assume that F permutes $\mathbb{F}_{2^n}$. Then Corollary 4.2.9 implies $Q(R(x)) = 0$ and $\mathrm{Tr}(R(x)) = 0$ for all $x \in \mathbb{F}_{2^n}$. Let $x \in \mathbb{F}_{2^n}$ and $y \in \ker(L_1^*)$ (recall that we can assume $\ker(L_1^*) \neq \{0\}$ by Corollary 4.2.2). Then

$$\begin{aligned} 0 =& Q(R(x+y)) = Q(R(x) + L_1^*(x)L_2^*(y) + L_1^*(y)L_2^*(x) + R(y)) \\ =& Q(R(x) + L_1^*(x)L_2^*(y)) \\ =& Q(R(x)) + Q(L_1^*(x)L_2^*(y)) + B_Q(R(x), L_1^*(x)L_2^*(y)) \\ =& Q(L_1^*(x)L_2^*(y)) + \mathrm{Tr}(R(x)L_1^*(x)L_2^*(y)), \end{aligned} \tag{4.1}$$

where we use $R(y) = L_1^*(y) = 0$, $Q(R(x)) = 0$ and $\mathrm{Tr}(R(x)) = 0$ throughout the computation, as well as the bilinear form $B_Q(x,y) = \mathrm{Tr}(xy) + \mathrm{Tr}(x)\,\mathrm{Tr}(y)$. For every $z \in \ker(L_1^*)$, we then get using Eq. (4.1)

$$\begin{aligned} 0 &= Q(L_1^*(x+z)L_2^*(y)) + \mathrm{Tr}(R(x+z)L_1^*(x+z)L_2^*(y)) \\ &= Q(L_1^*(x)L_2^*(y)) + \mathrm{Tr}(R(x+z)L_1^*(x)L_2^*(y)) \\ &= Q(L_1^*(x)L_2^*(y)) + \mathrm{Tr}(L_1^*(x)L_2^*(x+z)L_1^*(x)L_2^*(y)) \\ &= Q(L_1^*(x)L_2^*(y)) + \mathrm{Tr}(R(x)L_1^*(x)L_2^*(y)) + \mathrm{Tr}((L_1^*(x))^2L_2^*(z)L_2^*(y)). \end{aligned}$$

Adding the last equation to Eq. (4.1) yields

$$\mathrm{Tr}((L_1^*(x))^2L_2^*(z)L_2^*(y)) = 0 \tag{4.2}$$

for all $y, z \in \ker(L_1^*)$ and $x \in \mathbb{F}_{2^n}$.

Setting $y = z$ in Eq. (4.2) we have $0 = \mathrm{Tr}(L_1^*(x)L_2^*(y)) = \mathrm{Tr}(xL_1(L_2^*(y)))$ for all $x \in \mathbb{F}_{2^n}$. Consequently, $L_2^*(\ker(L_1^*)) \subseteq \ker L_1$. Since $\ker(L_1^*) \cap \ker(L_2^*) = \{0\}$ by Corollary 4.2.9 and $\dim \ker L_1 = \dim \ker L_1^*$ by Lemma 4.2.3, we get $\ker L_1 = L_2^*(\ker(L_1^*))$.

Again using Eq. (4.2), we get

$$\mathrm{Tr}\left(xL_1\left(\sqrt{L_2^*(z)L_2^*(y)}\right)\right) = 0$$

for all $y, z \in \ker(L_1^*)$ and $x \in \mathbb{F}_{2^n}$, so $\sqrt{L_2^*(\ker(L_1^*)) \cdot L_2^*(\ker(L_1^*))} \subseteq \ker L_1$. Now since $L_2^*(\ker(L_1^*)) = \ker L_1$ we have $\sqrt{\ker L_1 \cdot \ker L_1} = \ker L_1$.

This particularly implies that

$$|(\ker L_1 \setminus \{0\}) \cdot (\ker L_1 \setminus \{0\})| = |(\ker L_1 \setminus \{0\})|,$$

so by Theorem 4.2.10 we get $\ker L_1 = aH \cup \{0\}$ where $a \in \mathbb{F}_{2^n}^*$ and $H \leq \mathbb{F}_{2^n}^*$ is a subgroup of the multiplicative group of $\mathbb{F}_{2^n}$. Since $|\ker L_1| = 2^k$ for some $k \in \mathbb{N}$, we

infer $|H| = 2^k - 1$. In particular, H is the multiplicative group of a subfield of $\mathbb{F}_{2^n}$, so $\ker L_1 = a\mathbb{F}_{2^k}$ for some $a \neq 0$ and $k|n$.

Since $F(x) = L_1(x^{-1}) + L_2(x)$ is a permutation if and only if $F(x^{-1}) = L_1(x) + L_2(x^{-1})$ is a permutation, we get the same results also for the kernel of L_2.

□

Before we start with the second step of the proof, we need some lemmas.

Lemma 4.2.12 ([102]). *The quadratic equation $ax^2 + bx + c = 0$ over $\mathbb{F}_{2^n}$ with $b \neq 0$ has solutions in $\mathbb{F}_{2^n}$ if and only if $\mathrm{Tr}(ac/b^2) = 0$.*

Proposition 4.2.13. *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be defined by $F(x) = L_1(x^{-1}) + L_2(x)$, where L_1, L_2 are nonzero linear mappings over $\mathbb{F}_{2^n}$. F is a permutation polynomial if and only if F has only one zero and $L_2(a) \notin L_1(\frac{1}{H_{1/a}})$ for all $a \in \mathbb{F}_{2^n}^*$.*

Proof. F permutes $\mathbb{F}_{2^n}$ if and only if $F(x) + F(x+a) \neq 0$ for all $x \in \mathbb{F}_{2^n}$ and $a \in \mathbb{F}_{2^n}^*$, i.e.

$$L_1(x^{-1} + (x+a)^{-1}) \neq L_2(a). \tag{4.3}$$

We determine the set $M_a = \{c \in \mathbb{F}_{2^n} \mid \exists x \in \mathbb{F}_{2^n} \colon x^{-1} + (x+a)^{-1} = c\}$. We have

$$x^{-1} + (x+a)^{-1} = c \iff c = a^{-1} \text{ or } \mathrm{Tr}(\frac{1}{ac}) = 0.$$

Indeed, $c = a^{-1}$ if we choose $x = 0$ or $x = a$, and in the other cases we can multiply the equation by $x(x+a)$ which results in the quadratic equation

$$cx^2 + acx + a = 0.$$

By Lemma 4.2.12, this equation has a solution in $\mathbb{F}_{2^n}$ if and only if $c \neq 0$ and $\mathrm{Tr}(1/(ac)) = 0$. We conclude that $M_a = 1/H_{\frac{1}{a}} \cup \{a^{-1}\}$, so Eq. (4.3) gives $L_2(a) \notin L_1(M_a)$. Observe that $L_2(a) = L_1(a^{-1})$ if and only if a is a zero of F and the result follows. □

Lemma 4.2.14. *Let a, b, c be three distinct elements in $\mathbb{F}_{2^n}^*$. Then $H_a \cup H_b \cup H_c = \mathbb{F}_{2^n}$ if and only if $a + b = c$. In particular, if $M = r\mathbb{F}_{2^k}$ with $k|n$, $k > 1$ and $r \in \mathbb{F}_{2^n}^*$, we can always find three elements $a, b, c \in M \setminus \{0\}$, such that $H_{1/a} \cup H_{1/b} \cup H_{1/c} = \mathbb{F}_{2^n}$.*

Proof. Clearly, all hyperplanes have 2^{n-1} elements and all intersections of two distinct hyperplanes have 2^{n-2} elements, so

$$\begin{aligned}|H_a \cup H_b \cup H_c| =& |H_a| + |H_b| + |H_c| - |H_a \cap H_b| - |H_a \cap H_c| - |H_b \cap H_c| \\ &+ |H_a \cap H_b \cap H_c| \\ =& 3 \cdot 2^{n-1} - 3 \cdot 2^{n-2} + |H_a \cap H_b \cap H_c| \\ =& 2^{n-1} + 2^{n-2} + |H_a \cap H_b \cap H_c|.\end{aligned}$$

Consequently, $H_a \cup H_b \cup H_c = \mathbb{F}_{2^n}$ if and only if $|H_a \cap H_b \cap H_c| = 2^{n-2}$, which means $H_a \cap H_b \subseteq H_c$. This is equivalent to $a + b = c$.

Now let $M = r\mathbb{F}_{2^k}$ be a translate of a subfield of $\mathbb{F}_{2^n}$ with $k > 1$. Choose two distinct elements $a = rs_1$, $b = rs_2$ with $s_1, s_2 \in \mathbb{F}_{2^k}^*$, then

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{r}\left(\frac{1}{s_1} + \frac{1}{s_2}\right).$$

Clearly, $1/s_1 + 1/s_2 = 1/s$ is an element in $\mathbb{F}_{2^k}^*$. For the three elements $a, b, rs \in M$ we have $\frac{1}{a} + \frac{1}{b} = \frac{1}{rs}$, so we have $H_{1/a} \cup H_{1/b} \cup H_{1/rs} = \mathbb{F}_{2^n}$. □

Theorem 4.2.15. *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be defined by $F(x) = L_1(x^{-1}) + L_2(x)$ with $n \geq 5$, where L_1, L_2 are nonzero linear mappings over $\mathbb{F}_{2^n}$. If F is a permutation, then $|\ker(L_1)| = |\ker(L_2)| = 2$.*

Proof. By Theorem 4.2.11, we know that $\ker L_2 = a\mathbb{F}_{2^k}$ for some $a \in \mathbb{F}_{2^n}^*$ and $k|n$. We want to show that $k = 1$. Assume to the contrary that $k \geq 2$. Then there exist three distinct elements $a, b, c \in \ker L_2$ such that $H_{1/a} \cup H_{1/b} \cup H_{1/c} = \mathbb{F}_{2^n}$ by Lemma 4.2.14. Equivalently, $\frac{1}{H_{1/a}} \cup \frac{1}{H_{1/b}} \cup \frac{1}{H_{1/c}} = \mathbb{F}_{2^n}^*$.

By Proposition 4.2.13, we have $L_2(x) \notin L_1(\frac{1}{H_{1/x}})$ for all $x \in \mathbb{F}_{2^n}^*$. For the elements a, b, c, this equation becomes

$$0 \notin L_1(\frac{1}{H_{1/a}}) \cup L_1(\frac{1}{H_{1/b}}) \cup L_1(\frac{1}{H_{1/c}}) = L_1(\mathbb{F}_{2^n}^*).$$

But since $|\ker L_1| > 1$ by Corollary 4.2.2, this is not possible. We conclude $k = 1$ and $|\ker(L_2)| = 2$.

Since $L_1(x^{-1}) + L_2(x)$ is a permutation if and only if $L_1(x) + L_2(x^{-1})$ is a permutation, the argument works again symmetrically for the kernel of L_1. □

Using Theorem 4.2.15 and a suitable composition with a bijective linear mapping, we can assume without loss of generality that $L_1(x) = x^2 + ax$ for some $a \neq 0$. In fact, we can even assume $a = 1$ as the following argument shows:

$$\begin{aligned} & x^{-2} + ax^{-1} + L_2(x) = c \\ \iff\ & a^{-2}x^{-2} + a^{-1}x^{-1} + a^{-2}L_2(x) = a^{-2}c \end{aligned} \tag{4.4}$$

by multiplying the equation with a^{-2}. After a substitution $x \mapsto x/a$, we get

$$x^{-2} + x^{-1} + a^{-2}L_2(x/a) = a^{-2}c. \tag{4.5}$$

Eq. (4.4) has one solution for every $c \in \mathbb{F}_{2^n}$ if and only if Eq. (4.5) has one solution for each c. Observe that $a^{-2}L_2(x/a)$ is still a linear mapping, so we can consider without loss of generality $L_1(x) = x^2 + x$. In fact, we will instead use $L_1(x) = (x^2 + x)^{2^{n-1}} = x^{2^{n-1}} + x$, which is equivalent to the case $L_1(x) = x^2 + x$. Indeed, if $F(x) = x^{-2} + x^{-1} + L_2(x)$ is a permutation, then so is $F(x^{2^{n-1}}) = x^{-1} + x^{-2^{n-1}} + L_2(x^{2^{n-1}})$. The reason for this transformation is that in this case $L_1^*(x) = x^2 + x$, which makes the following technical calculations slightly easier. In this case we also have $\ker(L_1) = \ker(L_1^*) = \{0, 1\}$. By Theorem 4.2.11, we know $\ker(L_1) = L_2^*(\ker(L_1^*))$, which implies $L_2^*(1) = 1$.

Theorem 4.2.16. *There are no permutations of the form $F(x) = L_1(x^{-1}) + L_2(x)$ on $\mathbb{F}_{2^n}$ with nonzero linear mappings L_1, L_2 if $n \geq 5$.*

Proof. Assume that F is a permutation. By the considerations above we can assume without loss of generality that $L_1^*(x) = x^2 + x$ and $L_2^*(1) = 1$. We set $L_2^*(x) = \sum c_i x^{2^i}$ and derive necessary conditions on the coefficients c_i and show that those conditions contradict each other. As the basis we use the conditions given in Corollary 4.2.9 as

well as Eq. (4.1) for $y = 1$. We get

$$\mathrm{Tr}((x^2+x)L_2^*(x)) = 0 \tag{4.6}$$
$$Q((x^2+x)L_2^*(x)) = 0 \tag{4.7}$$
$$Q(x^2+x) + \mathrm{Tr}((x^4+x^2)L_2^*(x)) = 0 \tag{4.8}$$

for all $x \in \mathbb{F}_{2^n}$.

We start by expanding condition (4.6). We have

$$\begin{aligned}\mathrm{Tr}((x^2+x)\sum_{i=0}^{n-1} c_i x^{2^i}) &= \sum_{s=0}^{n-1}\sum_{i=0}^{n-1} c_i^{2^s} x^{2^{i+s}+2^{s+1}} + \sum_{s=0}^{n-1}\sum_{i=0}^{n-1} c_i^{2^s} x^{2^{i+s}+2^s} \\ &= \sum_{s=0}^{n-1}\sum_{i=0}^{n-1} c_{i-s}^{2^s} x^{2^i+2^{s+1}} + \sum_{s=0}^{n-1}\sum_{i=0}^{n-1} c_{i-s}^{2^s} x^{2^i+2^s} \\ &= \sum_{s=0}^{n-1}\sum_{i=0}^{n-1} (c_{i-s+1}^{2^{s-1}} + c_{i-s}^{2^s}) x^{2^i+2^s},\end{aligned}$$

where we use a transformation $i \mapsto i-s$ in the second step and then a transformation $s \mapsto s-1$ in the left double sum in the last step. Here we view the indices of the coefficients c_i modulo n. By condition (4.6) this polynomial is equal to the zero polynomial. When we check the coefficient of x^4 (achieved by setting $i = s = 1$), we get

$$c_1 = c_0^2. \tag{4.9}$$

Similarly, checking the coefficient of x^{2^r+1} for $1 \le r \le n-1$ (achieved for $i = 0, s = r$ and $i = r, s = 0$) we get

$$c_{-r+1}^{2^{r-1}} + c_{-r}^{2^r} + c_{r+1}^{2^{-1}} + c_r = 0 \tag{4.10}$$

for all $1 \le r \le n-1$.

We now do the same procedure for condition (4.8). We have

$$\begin{aligned}&Q(x^2+x) + \mathrm{Tr}((x^4+x^2)L_2^*(x)) \\ &= Q(x^2) + Q(x) + \mathrm{Tr}(x^3) + \mathrm{Tr}(x) + \mathrm{Tr}((x^4+x^2)\sum_{i=0}^{n-1} c_i x^{2^i}) \\ &= \mathrm{Tr}(x^3) + \mathrm{Tr}(x) + \sum_{s=0}^{n-1}\sum_{i=0}^{n-1} c_i^{2^s} x^{2^{i+s}+2^{s+2}} + \sum_{s=0}^{n-1}\sum_{i=0}^{n-1} c_i^{2^s} x^{2^{i+s}+2^{s+1}} \\ &= \sum_{i=0}^{n-1} x^{2^{i+1}+2^i} + \sum_{i=0}^{n-1} x^{2^i} + \sum_{s=0}^{n-1}\sum_{i=0}^{n-1} (c_{i-s+2}^{2^{s-2}} + c_{i-s+1}^{2^{s-1}}) x^{2^i+2^s},\end{aligned}$$

where we use $Q(x^2) = Q(x)$ and the last two steps again the transformations $i \mapsto i-s$ and $s \mapsto s-2$ (in the left double sum) and $s \mapsto s-1$ (in the right double sum). Checking the coefficient of x^8 (achieved by $i = s = 2$ in the double sum) we get

$$c_2 = c_1^2 + 1. \tag{4.11}$$

For the coefficients of x^{2^r+1} for $1 \leq r \leq n-1$ (achieved by $i = r$, $s = 0$ and $i = 0$, $s = r$), we similarly get

$$c_{r+2}^{2^{-2}} + c_{r+1}^{2^{-1}} + c_{-r+2}^{2^{r-2}} + c_{-r+1}^{2^{r-1}} = \begin{cases} 1, & r \in \{1, n-1\} \\ 0, & r \in \{2, \ldots, n-2\}, \end{cases} \tag{4.12}$$

where the additional 1 in the cases $r \in \{1, n-1\}$ is due to the $\mathrm{Tr}(x^3)$ term. Substituting $r \mapsto r-1$ and squaring the equation yields

$$c_{r+1}^{2^{-1}} + c_r + c_{-r+3}^{2^{r-2}} + c_{-r+2}^{2^{r-1}} = \begin{cases} 1, & r \in \{0, 2\} \\ 0, & r \in \{3, \ldots, n-1\}. \end{cases}$$

Adding this equation to Eq. (4.10), we get

$$c_{-r+1}^{2^{r-1}} + c_{-r}^{2^r} + c_{-r+3}^{2^{r-2}} + c_{-r+2}^{2^{r-1}} = \begin{cases} 1, & r = 2 \\ 0, & r \in \{3, \ldots, n-1\}. \end{cases}$$

We simplify the equation by substituting $r \mapsto -r$ and taking the resulting equation to the power 2^{r+2}:

$$c_{r+3} + c_{r+2}^2 + c_{r+1}^2 + c_r^4 = \begin{cases} 1, & r = n-2 \\ 0, & r \in \{1, \ldots, n-3\}. \end{cases} \tag{4.13}$$

We show by induction that the constraints we have obtained so far imply

$$c_i = \begin{cases} c_0^{2^i}, & i \text{ odd} \\ c_0^{2^i} + 1, & i \text{ even}. \end{cases} \tag{4.14}$$

for all $i \in \{1, \ldots, n-1\}$. The cases $i = 1$ and $i = 2$ are shown in Eq. (4.9) and (4.11). We verify the case $i = 3$ by using Eq. (4.12) with $r = 1$:

$$0 = 1 + c_3^{2^{-2}} + c_2^{2^{-1}} + c_1^{2^{-1}} + c_0 = 1 + c_3^{2^{-2}} + c_0^2 + 1 + c_0 + c_0 = c_3^{2^{-2}} + c_0^2,$$

which immediately yields $c_3 = c_0^8$ as claimed. We now proceed by induction: Assume that all coefficients up to $k \geq 3$ satisfy Eq. (4.14). Then by Eq. (4.13)

$$c_{k+1} = c_k^2 + c_{k-1}^2 + c_{k-2}^4 = (c_0^{2^{k+1}} + 1) + c_0^{2^k} + (c_0^{2^k} + 1) = c_0^{2^{k+1}}$$

if k is odd and

$$c_{k+1} = c_k^2 + c_{k-1}^2 + c_{k-2}^4 = c_0^{2^{k+1}} + (c_0^{2^k} + 1) + c_0^{2^k} = c_0^{2^{k+1}} + 1$$

if k is even, proving Eq. (4.14).

We compute c_{n-1} in another way using Eq. (4.10) for $r = 1$:

$$0 = c_0 + c_{n-1}^2 + c_2^{2^{-1}} + c_1 = c_{n-1}^2 + c_0 + c_0^2 + 1 + c_0^2, \tag{4.15}$$

so $c_{n-1} = 1 + c_0^{2^{n-1}}$. This immediately implies in connection with Eq. (4.14) that n must be odd.

The last step is to find a contradiction to the coefficients described in Eq. (4.14). For that, we use the condition from Eq. (4.7). First, we expand the condition

$$\begin{aligned} Q((x^2+x)\sum_{i=0}^{n-1} c_i x^{2^i}) &= Q(\sum_{i=0}^{n-1} c_i x^{2^i+2} + \sum_{i=0}^{n-1} c_i x^{2^i+1}) \\ &= \sum_{0\le r<s\le n-1} (\sum_{i=0}^{n-1} (c_{i-r+1}^{2^{r-1}} + c_{i-r}^{2^r}) x^{2^i+2^r}) (\sum_{j=0}^{n-1} (c_{j-s+1}^{2^{s-1}} + c_{j-s}^{2^s}) x^{2^j+2^s}) \\ &= \sum_{0\le r<s\le n-1} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (c_{i-r+1}^{2^{r-1}} + c_{i-r}^{2^r})(c_{j-s+1}^{2^{s-1}} + c_{j-s}^{2^s}) x^{2^i+2^j+2^r+2^s} \\ &= \sum_{0\le r<s\le n-1} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} d_{i,j,r,s} x^{2^i+2^j+2^r+2^s} \end{aligned}$$

with $d_{i,j,r,s} = (c_{i-r+1}^{2^{r-1}} + c_{i-r}^{2^r})(c_{j-s+1}^{2^{s-1}} + c_{j-s}^{2^s})$. To satisfy condition (4.7), this polynomial must be equal to the zero polynomial. Using Eq. (4.14) and Eq. (4.15), we see that

$$d_{i,j,r,s} = \begin{cases} 0, & i = r \text{ or } j = s \\ 1, & \text{else.} \end{cases}$$

We check the coefficient of x^8 of the polynomial $Q((x^2+x)L_2(x))$: The possible choices for i, j, r, s are:

1. $i = j = 0, r = 1, s = 2$ with $d_{i,j,r,s} = 1$
2. $i = r = 0, s = 1, j = 2$ with $d_{i,j,r,s} = 0$
3. $i = r = 0, j = 1, s = 2$ with $d_{i,j,r,s} = 0$
4. $j = r = 0, i = 1, s = 2$ with $d_{i,j,r,s} = 1$
5. $j = r = 0, s = 1, i = 2$ with $d_{i,j,r,s} = 1$.

In particular, the coefficient of x^8 is the sum of the listed values of $d_{i,j,r,s}$, which is 1, so $Q((x^2+x)L_2(x))$ is not the zero polynomial. This contradicts condition (4.7) and proves the theorem.

□

Remark 4.2.17. The condition $n \ge 5$ in Theorem 4.2.16 is necessary. Indeed, it is possible to find permutation polynomials of the form $L_1(x^{-1}) + L_2(x)$ over $\mathbb{F}_{2^4}$ and $\mathbb{F}_{2^3}$ using a simple computer search, examples with $L_1(x) = x$ can be found in [100].

Our main result of this section is a direct consequence from Theorem 4.2.16 and Proposition 4.1.2 (recall that the inverse function is an involution).

Theorem 4.2.18. *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be the inverse function with $n \ge 5$. The CCZ-class of F coincides with the EA-class of F. Moreover, all permutations in the CCZ-class of F are affine equivalent to F.*

4.2.1 Vector spaces of Kloosterman zeros

While working on the proof of Theorem 4.2.16, we were also interested in vector spaces of Kloosterman zeros, i.e. vector spaces that consist of Kloosterman zeros. This was motivated by the following observation.

Proposition 4.2.19. *Let $F(x) = L_1(x^{-1}) + L_2(x)$ where L_1 and L_2 are non-bijective non-zero linear functions of $\mathbb{F}_{2^n}$. If F is a permutation, then there is a vector space of Kloosterman zeros of dimension* $\max\{\dim\ker(L_1), \ker(L_2)\}$.

Proof. Observe that $F(x)$ is a permutation if and only if $F(x^{-1}) = L_1(x) + L_2(x^{-1})$ is so. Hence we may assume without loss of generality that $\max\{\dim\ker(L_1), \ker(L_2)\} = \dim(\ker(L_1))$. Suppose F is a permutation. Then by Corollary 4.2.5 we have $\ker(L_1^*) \cap \ker(L_2^*) = \{0\}$ and $K_n(L_1^*(b)L_2^*(b)) = 0$ for all $b \in \mathbb{F}_{2^n}$. Choose $0 \neq c \in \ker(L_2^*)$. The set

$$V = L_1^*(c + \ker(L_1^*)) \cdot L_2^*(c + \ker(L_1^*)) = L_1^*(c) \cdot L_2^*(\ker(L_1^*))$$

is a vector space that is contained in the image set of $L_1^*(b)L_2^*(b)$. In particular $K_n(v) = 0$ for all $v \in V$. Since $\ker(L_1^*) \cap \ker(L_2^*) = \{0\}$ we have $\dim(V) = \dim\ker L_1^* = \dim\ker L_1$. □

Of course, in the light of Theorem 4.2.16, this proposition does not give any insight. However, vector spaces of Kloosterman zeros are interesting objects on their own and were for example considered in [96] to construct vectorial bent functions by generalizing a classical construction of bent functions by Dillon.

Our objective is this section is to find an upper bound on the size of a vector space of Kloosterman zeros. To do this, we will again use the dyadic approximation given by Theorem 4.2.8. We will also use the theory of quadratic forms.

Let B be a bilinear form from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$. We denote by $\mathrm{rad}(B) = \{y \in \mathbb{F}_{2^n} : B(x, y) = 0 \text{ for all } x \in \mathbb{F}_{2^n}\}$ the radical of B. Given a quadratic form $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$, let $B_f(x, y) = f(x) + f(y) + f(x + y)$ be the bilinear form associated to it. The radical of the quadratic form f is defined as $\mathrm{rad}(B_f) \cap f^{-1}(\{0\})$. A quadratic form is called non-degenerate if $\mathrm{rad}(f) = \{0\}$.

Theorem 4.2.8 implies that the Kloosterman zeroes are contained in the intersection of the set $\{x \in \mathbb{F}_{2^n} : Q(x) = 0\}$ and the hyperplane

$$H = H_1 = \{x \in \mathbb{F}_{2^n} : \mathrm{Tr}(x) = 0\}.$$

Therefore we consider the quadratic form $Q|_H$ which is induced by Q on H. We first determine its radical.

Lemma 4.2.20. *We have*

$$\mathrm{rad}(Q|_H) = \begin{cases} \{0,1\}, & n \equiv 0 \pmod 4 \\ \{0\}, & \textit{else.} \end{cases}$$

Proof. The bilinear form associated to Q was computed in the last section as $B_Q(x, y) = \mathrm{Tr}(xy) + \mathrm{Tr}(x)\,\mathrm{Tr}(y)$. Since $\mathrm{Tr}(x) = 0$ for all $x \in H$, we have

$$B_{Q|_H}(x, y) = \mathrm{Tr}(xy).$$

Then $y \in \mathrm{rad}(B_{Q|_H})$, if $B_{Q|_H}(x, y) = \mathrm{Tr}(xy) = 0$ for all $x \in H$. Hence $\mathrm{rad}(B_{Q|_H}) = \mathbb{F}_2 \cap H$. Observe that $1 \in H$ if and only if n is even, so $\mathrm{rad}(B_{Q|_H}) = \{0\}$ if n is odd and $\mathrm{rad}(B_{Q|_H}) = \mathbb{F}_2$ if n is even. One can easily verify that

$$Q(1) = \frac{n(n-1)}{2} = \begin{cases} 0 & n \equiv 0, 1 \pmod 4 \\ 1 & n \equiv 2, 3 \pmod 4 \end{cases}$$

and the result follows. □

Let $N(Q|_H(x) = u)$ denote the number of solutions of $Q|_H(x) = u$ for $u \in \mathbb{F}_2$. By expanding the characteristic polynomial of an element $x \in \mathbb{F}_{2^n}$, it is easy to see that $N(Q|_H(x) = 0)$ is precisely the number of elements $x \in \mathbb{F}_{2^n}$ whose second and third coefficients of the characteristic polynomial are zero. The value $N(Q|_H(x) = u)$ was investigated in [53, 137, 33], where irreducible polynomials with prescribed coefficients were studied. In particular, the value $N(Q|_H(x) = 0)$ was determined. We summarize some of their results in the following theorem.

Theorem 4.2.21. *Let $N(Q|_H(x) = 0)$ be the number of $x \in H$ with $Q|_H(x) = 0$. Then $N(Q|_H(x) = 0) = 2^{n-2} + e$ where*

$$e = \begin{cases} -2^{\frac{n-2}{2}}, & n \equiv 0 \pmod 8 \\ 2^{\frac{n-3}{2}}, & n \equiv 1,7 \pmod 8 \\ 0, & n \equiv 2,6 \pmod 8 \\ -2^{\frac{n-3}{2}}, & n \equiv 3,5 \pmod 8 \\ 2^{\frac{n-2}{2}}, & n \equiv 4 \pmod 8. \end{cases}$$

Two quadratic forms f and g on a vector space V are called equivalent if f can be transformed into g with a non-singular linear transformation of V. The following result is well known (see e.g. [102, 70]).

Theorem 4.2.22 (Classification of quadratic forms)**.** *Let $f\colon V \to \mathbb{F}_2$ with $\dim(V) = n$ be a quadratic form with $\dim(\mathrm{rad}(f)) = w$. Then f is equivalent to one of three forms:*

$$f \simeq \sum_{i=1}^{v} x_i y_i \qquad \textit{(hyperbolic case)}$$

$$f \simeq z + \sum_{i=1}^{v} x_i y_i \qquad \textit{(parabolic case)}$$

$$f \simeq x_1^2 + x_1 y_1 + y_1^2 + \sum_{i=2}^{v} x_i y_i \qquad \textit{(elliptic case)},$$

where $v = \lfloor (n-w)/2 \rfloor$.

The value of $N(f(x) = 0)$ depends only on n, w and the type of the quadratic form. More precisely,

$$N(f(x) = 0) = 2^{n-1} + \Lambda(f) 2^{\frac{n+w-2}{2}},$$

with

$$\Lambda(f) = \begin{cases} 1, & \textit{if } f \textit{ is hyperbolic} \\ 0, & \textit{if } f \textit{ is parabolic} \\ -1, & \textit{if } f \textit{ is elliptic.} \end{cases}$$

The *Witt index* of a quadratic form is the number of pairs $x_i y_i$ that appear in the decomposition described above. In particular, the Witt index of f is v in the hyperbolic and parabolic case, and $v - 1$ in the elliptic case.

Remark 4.2.23. Just using the classification of quadratic forms in Theorem 4.2.22 and the determination of the radical in Lemma 4.2.20 we can give a simple alternative proof of the cases $n \equiv 2,6 \pmod 8$ in Theorem 4.2.21. Indeed, in these cases $Q|_H$ is necessarily parabolic which immediately gives the value for $N(Q|_H(x) = 0)$.

Let f be a quadratic form on V. A subspace W of V is called totally isotropic if $f(w) = 0$ for all $w \in W$. A subspace W is called maximal totally isotropic if there is no subspace W_2 with $f(w) = 0$ for all $w \in W_2$ and $W \subsetneq W_2 \subseteq V$. Two maximal totally isotropic subspaces have the same dimension, which is the sum of the Witt index and the dimension of the radical of the quadratic form, as the following result implies.

Proposition 4.2.24 ([89, Corollary 4.4.]). *Let $f\colon V \to \mathbb{F}_2$ be a non-degenerate quadratic form on a vector space V over $\mathbb{F}_2$ with $\dim(V) = n$. Let W be a maximal totally isotropic subspace of V. Then, the dimension of W is equal to the Witt index of f. In particular, we have*

$$\dim(W) = \begin{cases} \frac{n}{2}, & \text{if } f \text{ is hyperbolic} \\ \frac{n-1}{2}, & \text{if } f \text{ is parabolic} \\ \frac{n-2}{2}, & \text{if } f \text{ is elliptic.} \end{cases}$$

We collect the above observations to give a sharp upper bound on the size of vector spaces that consist of elements with Kloosterman sum divisible by 16.

Proposition 4.2.25. *Let W be a subspace of $\mathbb{F}_{2^n}$ with $K_n(w) \equiv 0 \pmod{16}$ for all $w \in W$ and $n \geq 5$. Then $\dim W \leq d$ where*

$$d = \begin{cases} \frac{n-2}{2}, & n \equiv 0, 2, 6 \pmod 8 \\ \frac{n-1}{2}, & n \equiv 1, 7 \pmod 8 \\ \frac{n-3}{2}, & n \equiv 3, 5 \pmod 8 \\ \frac{n}{2}, & n \equiv 4 \pmod 8. \end{cases}$$

The bounds are sharp.

Proof. From the Theorems 4.2.21 and 4.2.22 we deduce that $Q|_H$ is elliptic if $n \equiv 0, 3, 5 \pmod 8$, hyperbolic if $n \equiv 1, 4, 7 \pmod 8$ and parabolic if $n \equiv 2, 6 \pmod 8$. In the cases $n \not\equiv 0, 4 \pmod 8$ the quadratic form $Q|_H$ is non-degenerate by Lemma 4.2.20 and we immediately get bounds on $\dim(W)$ from Proposition 4.2.24 (recall that $Q|_H$ is a quadratic form on an $(n-1)$ dimensional space). If $n \equiv 0, 4 \pmod 8$ then $\dim(\operatorname{rad}(Q|_H)) = 1$, so $\dim V \leq 1 + \frac{n-4}{2} = \frac{n-2}{2}$ if $n \equiv 0 \pmod 8$ and $\dim V \leq 1 + \frac{n-2}{2} = \frac{n}{2}$ if $n \equiv 4 \pmod 8$. □

Remark 4.2.26. Every vector space W that contains exclusively Kloosterman zeros is of course also a vector space that contains only Kloosterman sums divisible by 16. In particular, by Propositions 4.2.24 and 4.2.25, all vector spaces of Kloosterman zeros are necessarily contained in a maximal totally isotropic vector space of $Q|_H$. However, these vector spaces are generally not unique.

Using Proposition 4.2.25, we get the following result.

Theorem 4.2.27. *Let W be a subspace of $\mathbb{F}_{2^n}$ such that $K_n(v) = 0$ for all $v \in W$ and $n \geq 5$. Then $\dim W \leq d$ where*

$$d = \begin{cases} \frac{n-2}{2}, & n \equiv 0, 2, 4, 6 \pmod 8 \\ \frac{n-1}{2}, & n \equiv 1, 7 \pmod 8 \\ \frac{n-3}{2}, & n \equiv 3, 5 \pmod 8. \end{cases}$$

Proof. The bound follows from Proposition 4.2.25 for all cases except $n \equiv 4 \pmod 8$. In the latter case the bound of Proposition 4.2.25 can be improved by one using the

following observation for even n. [1] Let $n = 2k$ be even. As noted in [103], there are no nontrivial Kloosterman zeros in the subfield $\mathbb{F}_{2^k}$. We have $\mathbb{F}_{2^k} \subset H$, $W \subset H$ and $W \cap \mathbb{F}_{2^k} = \{0\}$, implying $\dim(V) \leq \frac{n-2}{2}$. □

We would like to mention that the following approach yields a slightly weaker bound than the one given in Theorem 4.2.27. The following identity for sums of Kloosterman sums over a vector space was given in [38, Proposition 3]: For any subspace V of $\mathbb{F}_{2^n}$ with $\dim(V) = k$ we have

$$\sum_{a \in V} (K_n^2(a) - K_n(a)) = 2^{n+k} - 2^{n+1} + 2^k \sum_{u \in V^\perp} K_n(u^{-1}).$$

If V contains exclusively Kloosterman zeros, we get

$$0 = 2^{n+k} - 2^{n+1} + 2^k \sum_{u \in V^\perp} K_n(u^{-1}),$$

recall we set $0^{-1} = 0$. Bounding the Kloosterman sum in the right hand side of the equation using Theorem 4.2.7, $|K_n(a)| \leq 2^{\frac{n}{2}+1}$, we get

$$0 \geq 2^{n+k} - 2^{n+1} - 2^k 2^{n-k} 2^{\frac{n}{2}+1} = 2^{n+k} - 2^{n+1} - 2^{\frac{3n}{2}+1}.$$

This shows that $k = \dim(V) \leq \frac{n}{2} + 1$ for $n \geq 3$.

Remark 4.2.28. Theorem 4.2.27 provides to our knowledge the first general upper bound on the maximal size of subspaces of Kloosterman zeros. However, experimental results indicate that our bound is weak, see Table 4.1. Our bound is sharp for very small n (see right table in Table 4.1), which is not surprising since the approximation modulo 16 is strong for small n. Numerics in Table 4.1 were computed using [88, 80] for the left table and [14] for the right table. The left table shows that the total number of Kloosterman zeros in the field $\mathbb{F}_{2^n}$ is close to $2^{n/2}$ for $n \leq 60$. It is of course not to expect that the set of Kloosterman zeros has a strong additive structure, so we believe that the bound of Theorem 4.2.27 can be significantly improved.

Problem 4.2.29. Find a better bound on the maximal size of a subspace containing exclusively Kloosterman zeros.

4.3 Other monomials

We now consider more generally permutation polynomials of the type $L_1(x^d) + L_2(x)$. The content of this section is original work and has not been published in any form as of writing this thesis. The basic idea is similar to the one used in the last section: We use a dyadic approach, i.e. we use Proposition 4.1.3 and weaken the condition $W_F(L_2^*(b), L_1^*(b)) = 0$ to $W_F(L_2^*(b), L_1^*(b)) \equiv 0 \pmod{2^k}$ for some positive integer k. This allows us to use divisibility results, and leads in several cases already to non-existence results. To obtain statements on the divisibility of the Walsh transform, we use Stickelberger's congruences in characteristic 2 (Corollary 2.5.21) on Gauss sums.

Let us first state a straightforward generalization from [37, Lemma 2].

[1] This argument is due to an anonymous referee of the paper [57].

n	$2^{\frac{-n}{2}}\mathcal{Z}(n)$
5	0.88
10	1.87
15	1.57
20	0.86
25	0.67
30	1.29
35	1.15
40	1.15
45	1.14
50	0.91
55	1.32
60	1.25

n	$\dim(V)$
5	1
6	2
7	3
8	1
9	1
10	2
11	2
12	2
13	1
14	3
15	4
16	2

TABLE 4.1: Left Table: Comparison of the number of Kloosterman zeros over $\mathbb{F}_{2^n}$ to the value $2^{n/2}$. Here, $\mathcal{Z}(n)$ denotes the number of Kloosterman zeros over $\mathbb{F}_{2^n}$.
Right table: the maximal dimension of a subspace W of $\mathbb{F}_{2^n}$ that contains exclusively Kloosterman zeros.

Proposition 4.3.1. *Let d, n be positive integers with $\gcd(d, 2^n - 1) = s$ with $s > 1$. Let $G\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be the monomial $G = x^d$ and $F = L_1(x^d) + L_2(x)$ for linear polynomials L_1, L_2. If F is a permutation, then L_2 is a permutation as well.*

Proof. By Proposition 4.1.3, F is a permutation if and only if

$$0 = W_G(L_2^*(b), L_1^*(b)) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(L_1^*(b)x^d + L_2^*(b)x)} \tag{4.16}$$

for all $b \in \mathbb{F}_{2^n}^*$. Assume that L_2 is not a permutation. Then by Lemma 4.2.3, L_2^* is also not a permutation, so choose an element $b \in \mathbb{F}_{2^n}^*$ with $L_2^*(b) = 0$. If $L_1^*(b) = 0$ then clearly $W_G(L_2^*(b), L_1^*(b)) = 2^n$, contradicting Eq. (4.16). If $L_1^*(b) \neq 0$ then $x \mapsto L_1^*(b)x^d$ is an s-to-1 function, i.e. every element in $\mathbb{F}_{2^n}^*$ has either s or 0 preimages. Hence, the size of the set $\{x\colon \mathrm{Tr}(L_1^*(b)x^d) = 1\}$ is divisible by s and in particular not 2^{n-1} (observe that s cannot be even). We conclude $W_G(L_2^*(b), L_1^*(b)) \neq 0$, so F is not a permutation. □

We start by rewriting the Walsh transform using Gauss sums. The steps are fairly standard and have appeared similarly for example in [93, Eq. (3) onwards], [1, Section 4].

Proposition 4.3.2. *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be a monomial $F(x) = x^d$. Let ψ be a generator of $\widehat{\mathbb{F}_{2^n}^*}$. Then*

$$W_F(a, b) = \frac{2^n}{2^n - 1} + \frac{1}{2^n - 1} \sum_{j=1}^{2^n-2} G(\overline{\psi}^j) G(\psi^{jd}) \psi((b/a^d)^j)$$

for all $a, b \in \mathbb{F}_{2^n}^*$.

Proof. We have $W_F(a, b) = \sum_{x\in\mathbb{F}_{2^n}} \chi_1(bx^d)\chi_1(ax)$ (see Example 2.5.6). By Proposition 2.5.17, we know

$$\chi_1(x) = \frac{1}{2^n - 1} \sum_{\psi\in\widehat{\mathbb{F}_{2^n}^*}} G(\psi)\overline{\psi}(x).$$

for all $x \in \mathbb{F}_{2^n}^*$. Combining these two statements, we get

$$\begin{aligned} W_F(a,b) &= 1 + \sum_{x \in \mathbb{F}_{2^n}^*} \chi_1(bx^d)\chi_1(ax) \\ &= 1 + \frac{1}{(2^n-1)^2} \sum_{x \in \mathbb{F}_{2^n}^*} \sum_{\psi_1 \in \widehat{\mathbb{F}_{2^n}^*}} \sum_{\psi_2 \in \widehat{\mathbb{F}_{2^n}^*}} G(\psi_1)\overline{\psi_1}(bx^d)G(\psi_2)\overline{\psi_2}(ax) \\ &= 1 + \frac{1}{(2^n-1)^2} \sum_{\psi_1 \in \widehat{\mathbb{F}_{2^n}^*}} \sum_{\psi_2 \in \widehat{\mathbb{F}_{2^n}^*}} G(\psi_1)G(\psi_2)\overline{\psi_1}(b)\overline{\psi_2}(a) \sum_{x \in \mathbb{F}_{2^n}^*} \overline{\psi_1}^d(x)\overline{\psi_2}(x) \end{aligned}$$

Using the orthogonality relations (Proposition 2.5.4), we have

$$\sum_{x \in \mathbb{F}_{2^n}^*} \overline{\psi_1}^d(x)\overline{\psi_2}(x) = \begin{cases} 2^n - 1, & \text{if } \psi_2 = \overline{\psi_1}^d \\ 0, & \text{else.} \end{cases}$$

Consequently,

$$\begin{aligned} W_F(a,b) &= 1 + \frac{1}{2^n-1} \sum_{\psi \in \widehat{\mathbb{F}_{2^n}^*}} G(\psi)G(\overline{\psi}^d)\psi(a^d/b) \\ &= \frac{2^n}{2^n-1} + \frac{1}{2^n-1} \sum_{\substack{\psi \in \widehat{\mathbb{F}_{2^n}^*} \\ \psi \neq \psi_0}} G(\psi)G(\overline{\psi}^d)\psi(a^d/b) \\ &= \frac{2^n}{2^n-1} + \frac{1}{2^n-1} \sum_{\substack{\psi \in \widehat{\mathbb{F}_{2^n}^*} \\ \psi \neq \psi_0}} G(\overline{\psi})G(\psi^d)\psi(b/a^d) \end{aligned}$$

where we use $G(\psi_0) = -1$ in the first step and the substitution $\psi \mapsto \overline{\psi}$ in the second step. If ψ is a generator of $\widehat{\mathbb{F}_{2^n}^*}$, we can write equivalently

$$W_F(a,b) = \frac{2^n}{2^n-1} + \frac{1}{2^n-1} \sum_{j=1}^{2^n-2} G(\overline{\psi}^j)G(\psi^{jd})\psi((b/a^d)^j). \qquad \square$$

Example 4.3.3. Let us consider Proposition 4.3.2 for the inverse function on $\mathbb{F}_{2^n}$, i.e. $d = 2^n - 2$. As discussed in the previous section, the Walsh transform is in this case closely connected to Kloosterman sums. Let ψ be a generator of $\widehat{\mathbb{F}_{2^n}^*}$. Then

$$\begin{aligned} W_F(a,b) = K_n(ab) &= \frac{2^n}{2^n-1} + \frac{1}{2^n-1} \sum_{j=1}^{2^n-2} G(\overline{\psi}^j)G(\psi^{-j})\psi((ab)^j) \\ &= \frac{2^n}{2^n-1} + \frac{1}{2^n-1} \sum_{j=1}^{2^n-2} (G(\overline{\psi}^j))^2\psi((ab)^j) \end{aligned}$$

for all $a, b \in \mathbb{F}_{2^n}^*$. In particular for $b = 1$

$$K_n(a) = \frac{2^n}{2^n-1} + \frac{1}{2^n-1} \sum_{j=1}^{2^n-2} (G(\overline{\psi}^j))^2\psi(a^j). \tag{4.17}$$

By Stickelberger's congruence (Corollary 2.5.21), we have $(G(\overline{\psi}^j))^2 \equiv 0 \pmod 4$ for all $1 \leq j \leq 2^n - 2$, so we can immediately see that $K_n(a)$ is always divisible by 4 for

$n \geq 2$. More precisely, as noted by van der Geer and van der Vlugt [132, Remark 3.10.], we have $(G(\overline{\psi}^j))^2 \equiv 4 \pmod 8$ if and only if $\mathrm{wt}(j) = 1$. Since $\psi(a^j) \equiv a^j \pmod 2$ (see Eq. (2.4)), we get for $n \geq 3$

$$K_n(a) \equiv 4 \sum_{i=0}^{n-1} a^{2^i} \equiv 4 \operatorname{Tr}(a) \pmod 8,$$

where we (with slight abuse of notation) view $\operatorname{Tr}(a)$ as an integer. This result by van der Geer and van der Vlugt received little attention for several years until it was rediscovered Helleseth and Zinoviev who found a proof with more elementary techniques [65, Theorem 3].

When we apply Stickelberger's congruence to Proposition 4.3.2 it is clear that the divisibility of the Walsh transform depends on the value $\mathrm{wt}(j) + \mathrm{wt}(-jd)$ (here we mean by $\mathrm{wt}(-jd)$ the weight of $-jd \in \mathbb{Z}_{2^n-1}$, i.e. reduced modulo $2^n - 1$).

Definition 4.3.4 (J-set)**.** *Let d be an integer, $0 \leq d \leq 2^n - 2$. We define M_d^n by*

$$M_d^n = \min_{j \in \{1,2,\ldots,2^n-2\}} (\mathrm{wt}(j) + \mathrm{wt}(-jd)).$$

The J-set *of d over $\mathbb{F}_{2^n}$ is the set*

$$J_d^n = \{j \in \{1, 2, \ldots, 2^n - 2\} \colon \mathrm{wt}(j) + \mathrm{wt}(-jd) = M_d^n\}.$$

Here we view $-jd$ as an element in $\mathbb{Z}_{2^n-1}$, i.e. its smallest non-negative residue modulo $2^n - 1$.

Clearly, J_d^n is never empty. Further, since the binary expansion of $2j$ modulo $2^n - 1$ is just the binary expansion of j shifted by one, we have $\mathrm{wt}(2j) = \mathrm{wt}(j)$ for all $j \in \mathbb{Z}_{2^n-1}$. Hence, the J-set is always a union of cyclotomic cosets. For instance, in Example 4.3.3 we had $M_{2^n-2}^n = 2$ and $J_{2^n-2}^n = \{2^i \colon 0 \leq i \leq n-1\}$, i.e. the cyclotomic coset of 1.

Using these definitions, we can rewrite Proposition 4.3.2 modulo $2^{M_d^n+1}$.

Proposition 4.3.5. *Let $F \colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be a monomial $F(x) = x^d$ with $0 \leq d \leq 2^n - 2$ and $\mathrm{wt}(d) > 1$. Then*

$$W_F(a, b) \equiv 2^{M_d^n} \sum_{j \in J_d^n} (b/a^d)^j \pmod{2^{M_d^n+1}} \tag{4.18}$$

for all $a, b \in \mathbb{F}_{2^n}^*$.

Proof. By Proposition 4.3.2, we have

$$W_F(a, b) = \frac{2^n}{2^n - 1} + \frac{1}{2^n - 1} \sum_{j=1}^{2^n-2} G(\overline{\psi}^j) G(\psi^{jd}) \psi((b/a^d)^j).$$

Note that $M_d^n < n$ (consider for example $\mathrm{wt}(j) + \mathrm{wt}(-jd)$ for $j = 1$). Accordingly, $2^n \equiv 0 \pmod{2^{M_d^n+1}}$. By the definition of the J-set and Stickelberger's congruence, we have

$$G(\overline{\psi}^j) G(\psi^{jd}) \equiv \begin{cases} 2^{M_d^n} \pmod{2^{M_d^n+1}}, & \text{if } j \in J_d^n \\ 0 \pmod{2^{M_d^n+1}}, & \text{if } j \notin J_d^n \end{cases}$$

and by Eq. (2.4) $\psi((b/a^d)^j) \equiv (b/a^d)^j \pmod 2$, and the result follows. □

Since the J-set is a union of cyclotomic cosets, the sum $\sum_{j \in J_d^n}((b/a^d)^j)$ is a sum of absolute traces. In particular, it is always zero or one, so the right hand side of the congruence (4.18) is always an integer as desired. Note that if x^d is almost bent over $\mathbb{F}_{2^n}$, then $M_d^n = \frac{n+1}{2}$. In fact, a major step in the proofs of the Welch and Niho conjectures was precisely to find M_d^n [68, 23].

In conjunction with Proposition 4.1.3, we get the following necessary condition for a function $L_1(x^d) + L_2(x)$ to be a permutation.

Proposition 4.3.6. *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be defined by $F = L_1(x^d) + L_2(x)$ for linear polynomials L_1 and L_2. If F is a permutation, then*

$$\sum_{j \in J_d^n} \left(\frac{L_1^*(b)}{L_2^*(b)^d} \right)^j = 0 \tag{4.19}$$

for all $b \in \mathbb{F}_{2^n}$. If $x \mapsto x^d$ is almost bent over $\mathbb{F}_{2^n}$, then the condition in Eq. (4.19) *is also sufficient for F to be a permutation.*

Proof. Set $G\colon x \mapsto x^d$ on $\mathbb{F}_{2^n}$. By Proposition 4.1.3, $W_G(L_2^*(b), L_1^*(b)) = 0$ for all $b \in \mathbb{F}_{2^n}^*$. If $b \notin \ker L_1^* \cup \ker L_2^*$ then $\sum_{j \in J_d^n} \left(\frac{L_1^*(b)}{L_2^*(b)^d} \right)^j = 0$ using Proposition 4.3.5. If $b \in \ker L_1^* \cup \ker L_2^*$, then Eq. (4.19) of course holds as well using the convention $0^{-1} = 0$.

If G is almost bent on $\mathbb{F}_{2^n}$ then its extended Walsh spectrum contains only the values 0 and $2^{(n+1)/2}$, in particular $M_d^n = (n+1)/2$. Hence, we have $W_G(a,b) = 0$ if and only if $W_G(a,b) \equiv 0 \pmod{2^{M_d^n+1}}$ for all $a, b \in \mathbb{F}_{2^n}^*$, so the condition in Eq. (4.19) is sufficient to show that F is a permutation. □

For most values of d, determining M_d^n and the set J_d^n is very difficult, however in some cases results are possible.

A particularly simple case are the almost bent Gold exponents. Because of the low weight of the Gold exponent, the J-set can be determined fairly easily.

Proposition 4.3.7 ([94, Section 9.1.])**.** *Let n odd and $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be the Gold function $F(x) = x^{G_r}$ with $G_r = 2^r + 1$ and $\gcd(r,n) = 1$. Then $M_{G_r}^n = \frac{n+1}{2}$ and*

$$J_{G_r}^n = \left\{ -2^i (2^r+1)^{-1} \colon i \in \{0, 1, \ldots, n-1\} \right\},$$

where we take the inverse in $\mathbb{Z}_{2^n-1}$.

With Proposition 4.3.6 we can now find all permutation polynomials of the form $F = x^{2^r+1} + L(x)$ over $\mathbb{F}_{2^n}$ with $\gcd(r,n) = 1$. This was already done in [99] with different means. The proof in [99] is very technical and only shown in detail for the case $r = 1$. In contrast, we will give a complete unified proof for all AB Gold exponents.

Proposition 4.3.8 ([99, Theorem 4])**.** *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be defined by $F = x^d + L(x)$ with $d = 2^r + 1$, $\gcd(r,n) = 1$ and some linear polynomial L. F is a permutation if and only if n is odd and $L(x) = a^{2^r} x + a x^{2^r}$ for some $a \in \mathbb{F}_{2^n}$.*

Proof. Observe that F is EA-equivalent to the Gold function, in particular F is a quadratic APN function and thus cannot be a permutation if n is even by Proposition 2.4.1.

So let n be odd. Recall that the Gold function is almost bent. By Propositions 4.3.6 and 4.3.7, F is a permutation if and only if

$$0 = \sum_{i=0}^{n-1} \left(\frac{b}{L^*(b)^{2^r+1}} \right)^{-\frac{2^i}{2^r+1}} = \sum_{i=0}^{n-1} \left(b^{\frac{-1}{2^r+1}} L^*(b) \right)^{2^i} = \mathrm{Tr}\left(b^{\frac{-1}{2^r+1}} L^*(b) \right)$$

for all $b \in \mathbb{F}_{2^n}$.

Using Proposition 3.2.5, we can determine the binary expansion of $-(2^r+1)^{-1}$.

$$\frac{-1}{2^r+1} = -\sum_{i=0}^{\frac{n-1}{2}} 2^{2ir} = \sum_{i=0}^{\frac{n-3}{2}} 2^{r+2ir}.$$

We write $L^*(x) = \sum_{i=0}^{n-1} c_i x^{2^{r+2ri}}$. Then

$$x^{\frac{-1}{2^r+1}} L^*(x) = \sum_{i=0}^{n-1} c_i x^{2^{r+2ri} + \sum_{i=0}^{\frac{n-3}{2}} 2^{r+2ir}}.$$

We can visualize the exponents by writing them as a binary sequence in the order $(s_r, s_{3r}, s_{5r}, \ldots, s_{r+2(n-1)r})$ similar to the r-ordering we used in Chapter 3. The exponent $-(2^r+1)^{-1} = \sum_{i=0}^{\frac{n-3}{2}} 2^{r+2ir}$ will then be represented by the sequence

$$(\underbrace{1,1,1,\ldots,1}_{\frac{n-1}{2}\text{-times}},0,0,\ldots,0).$$

The exponent of the term $c_j x^{2^{r+2rj} + \sum_{i=0}^{\frac{n-3}{2}} 2^{r+2ir}}$ can then be written as a sequence

$$s_j = (\underbrace{1,1,1,\ldots,1}_{\frac{n-1}{2}\text{-times}},0,0,\ldots,0) + (0,0,\ldots,0,1,0,\ldots,0), \tag{4.20}$$

where the 1 on the right is in position j. If $\mathrm{Tr}\left(b^{\frac{-1}{2^r+1}} L^*(b) \right) = 0$ for all $b \in \mathbb{F}_{2^n}$ then all cyclotomic cosets of exponents inside of the trace-function have to cancel out. The exponents of two sequences written in the style of Eq. (4.20) belong to the same cyclotomic coset if and only if the sequences are cyclic shifts of one another. It is clear that all sequences in Eq. (4.20) belong to exponents in different cyclotomic cosets, except when $j = (n-1)/2$ or $j = n-1$ (in this case the two sequences can be transformed into one another with a cyclic shift by one). Accordingly, the only possible nonzero coefficients are $c_{(n-1)/2}$ and c_{n-1}, i.e. $L^*(x) = c_{(n-1)/2}x + c_{n-1}x^{2^{n-r}}$. Recall that because of the ordering of the sequence in Eq. (4.20), a shift by one represents a multiplication with 2^{2r}. We thus get $c_{(n-1)/2} = c_{n-1}^{2^{2r}}$ and $L^*(x) = c_{(n-1)/2}^{2^{n-2r}} x + c_{(n-1)/2} x^{2^{n-r}}$ for some $c_{(n-1)/2} \in \mathbb{F}_{2^n}$. Taking the adjoint and setting $a = c_{(n-1)/2}$, we conclude $L(x) = ax + a^{2^{n-r}} x^{2^r}$, and after a simple transformation $a \mapsto a^{2^r}$ we infer $L(x) = a^{2^r} x + a x^{2^r}$.

□

Remark 4.3.9. It is easy to verify that $F = x^{2^r+1} + a^{2^r} x + a x^{2^r}$ is a permutation polynomial by elementary means. Indeed, $F = (x+a)^{2^r+1} + a^{2^r+1}$ so F is a permutation since the Gold function $x \mapsto x^{2^r+1}$ is a permutation for n odd. The main statement

of Proposition 4.3.8 is that these are the *only* permutation polynomials of the form $x^{2^r+1} + L(x)$.

We now deal with exponents of the form $d = 2^r - 1$, i.e. we consider permutation polynomials of the form $L_1(x^{2^r-1}) + L_2(x)$. The cryptographic properties of the monomials $x \mapsto x^{2^r-1}$ have been investigated in several works, e.g. [10, 12], motivated initially by the fact that this family contains two different interesting monomials: for $r = 2$ we get the first Gold function $x \mapsto x^3$ and for $r = n - 1$ we get the function $x \mapsto x^{2^{n-1}-1}$ which is cyclotomic equivalent to the inverse function. However, other cases also show good properties [12]. For example, when working over a field with even dimension $n = 2k$, another interesting exponent of this form is the Dillon-exponent $d = 2^{n/2} - 1$, which can be used to construct bent functions [44].

The following elementary lemma shows when the function $x \mapsto x^{2^r-1}$ is a permutation of $\mathbb{F}_{2^n}$.

Lemma 4.3.10. *Let r, n be positive integers. Then* $\gcd(2^r - 1, 2^n - 1) = 2^{\gcd(r,n)} - 1$

In order to apply the divisibility results on Gauss sums we developed in this section, we will need to consider the values of $\mathrm{wt}(j) + \mathrm{wt}(-j(2^r - 1))$. To do this, we apply the modular add-with-carry approach we introduced in Chapter 3.

Lemma 4.3.11. *Let r, n be integers, $d = \gcd(r, n)$ and e be the least positive residue of the inverse of $\frac{r}{d}$ modulo $\frac{n}{d}$. Further, set $l = 2^r - 1$ and let $a, s \in \{1, 2, \ldots, 2^n - 2\}$ defined by $a \equiv \sum_{i=0}^{d-1} \sum_{j=0}^{\frac{n}{d}-1} a_{i,j} 2^{i-jr} \pmod{2^n - 1}$ and $s \equiv \sum_{i=0}^{d-1} \sum_{j=0}^{\frac{n}{d}-1} s_{i,j} 2^{i-jr} \pmod{2^n - 1}$ with binary coefficients $a_{i,j}, s_{i,j}$. The following are equivalent:*

(a) $s \equiv a \cdot l \pmod{2^n - 1}$

(b) There exists a matrix $c_{i,j}$ with $0 \le i < d$, $0 \le j < \frac{n}{d}$ and $c_{i,j} \in \{-1, 0\}$ (the r-matrix of the carry sequence*) such that*

$$2c_{0,j} - c_{d-1,j+e} + s_{0,j} = a_{0,j+1} - a_{0,j} \text{ for all } j \in \{0, \ldots, \frac{n}{d} - 1\} \tag{4.21}$$

$$2c_{i,j} - c_{i-1,j} + s_{i,j} = a_{i,j+1} - a_{i,j} \text{ for all } i \in \{1, \ldots, d-1\}, j \in \{0, \ldots, \frac{n}{d} - 1\} \tag{4.22}$$

holds for all i. Here, the indices are seen as elements in $\mathbb{Z}_n$.

If $s \equiv a \cdot l \pmod{2^n - 1}$, then we have

$$\mathrm{wt}(s) = -\sum_{i=0}^{d-1} \sum_{j=0}^{\frac{n}{d}-1} c_{i,j}$$

Proof. We apply Theorem 3.1.1 and use its notation. Let $a = (a_{n-1}, \ldots, a_0)$ and $s = (s_{n-1}, \ldots, s_0)$ be the binary expansions of a and s. We have $t_r = 1$, $t_0 = -1$ and $t_i = 0$ for all other values. Thus $t_+ = 1$ and $t_- = -1$, so $c_i \in \{-1, 0\}$ for the entries in the carry sequence, so $s \equiv al \pmod{2^n - 1}$ if and only if there exists a carry sequence $c = (c_{n-1}, \ldots, c_0)$ such that

$$2c_i - c_{i-1} + s_i = a_{i-r} - a_i \tag{4.23}$$

for all $i \in \mathbb{Z}_n$. To simplify working with Eq. (4.23), we again use the notation of r-matrices introduced in Definition 3.2.6. We denote by $M_{a,r} = (a_{i,j})$ the r-matrix of a

and similarly by $M_{s,r} = (s_{i,j})$ and $M_{c,r} = (c_{i,j})$ the r-matrices of s and c, respectively. Eq. (4.23) can then be rewritten (completely identically to Theorem 3.2.7) as

$$2c_{0,j} - c_{d-1,j+e} + s_{0,j} = a_{0,j+1} - a_{0,j} \text{ for all } j \in \{0, \dots, \frac{n}{d} - 1\}$$
$$2c_{i,j} - c_{i-1,j} + s_{i,j} = a_{i,j+1} + a_{i,j} \text{ for all } i \in \{1, \dots, d-1\}, j \in \{0, \dots, \frac{n}{d} - 1\}.$$

where e is the least positive residue of the inverse of $\frac{r}{d}$ modulo $\frac{n}{d}$.

The value of $\mathrm{wt}(s)$ can be determined by summing Eq. (4.21) and Eq. (4.22) for all values of i and j. □

Lemma 4.3.12. *Let r, n be positive integers such that $d = \gcd(r, n)$. For each $a \in \mathbb{Z}_{n-1}$ precisely one of the following cases occurs:*

- $a(2^r - 1) \equiv 0 \pmod{2^n - 1}$,
- $\mathrm{wt}(-a(2^r - 1)) \geq d$.

Here we take the product $-a(2^r - 1)$ in $\mathbb{Z}_{2^n-1}$.

Proof. We use the notation of Lemma 4.3.11 and set $l = 2^r - 1$.

Assume that the r-matrix of the carry sequence $M_{c,r}$ has a row i that is comprised exclusively of -1's. Let us now look at the $(i+1)$-st row of $M_{c,r}$. Then (see Eq.s (4.21) and (4.22)) either $c_{i+1,j} = -1$, $s_{i+1,j} = 1, a_{i,j} = a_{i+1,j+1}$ or $c_{i+1,j} = 0$, $s_{i+1,j} = 0$, $a_{i+1,j} = 0, a_{i+1,j+1} = 1$. Assume that there is a j such that $a_{i+1,j} = 0$, $a_{i+1,j+1} = 1$. Then there is a j' such that $a_{i+1,j'} = 1$, $a_{i+1,j'+1} = 0$, which cannot occur. So we have necessarily $c_{i+1,j} = -1$, $s_{i+1,j} = 1$, $a_{i+1,j} = a_{i+1,j+1}$ for all j. In particular, if the i-th row of $M_{c,r}$ contains only -1's, then the $(i+1)$-st row also contains only -1's and $s_{i+1,j} = 1$ for all j. We conclude inductively that if $M_{c,r}$ has one row with only -1's, then the entire matrix is comprised only of -1's and the entire matrix $M_{s,r}$ contains only 1's. Accordingly, $a(2^r + 1) \equiv s \equiv 0 \pmod{2^n - 1}$.

In all other cases, every row of $M_{c,r}$ contains at least one 0, so by Lemma 4.3.11, $\mathrm{wt}(a(2^r - 1)) = \mathrm{wt}(s) = -\sum_{i=0}^{d-1} \sum_{j=0}^{\frac{n}{d}-1} c_{i,j} \leq n - d$, so $\mathrm{wt}(-a(2^r - 1)) = n - \mathrm{wt}(a(2^r - 1)) \geq d$.

□

Theorem 4.3.13. *Let r, n be positive integers with $d = \gcd(r, n)$ and $n \geq 4$. Define a function $G\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ by $G = x^{2^r-1}$ and $F = G + L$ for some linear polynomial L. If $n \leq d^2$, then F is not a permutation.*

Moreover, if n is even, there are no permutations of the form $L_1(x^{2^{n/2}-1}) + L_2(x)$ for non-zero linear polynomials L_1, L_2.

Proof. We consider the polynomial $F' = L_1(x^{2^r-1}) + L_2(x)$. Note that $d > 1$, so $\gcd(2^r - 1, 2^n - 1) > 1$ and G is not a permutation. By Proposition 4.3.1, L_2^* must then be a permutation if F' is a permutation. By Proposition 4.3.2

$$W_G(L_2^*(b), L_1^*(b)) = \frac{2^n}{2^n - 1} + \frac{1}{2^n - 1} \sum_{j=1}^{2^n-2} G(\overline{\psi}^j) G(\psi^{j(2^r-1)}) \psi((L_1^*(b)/L_2^*(b)^{2^r-1})^j)$$

for all $b \in \mathbb{F}_{2^n}^*$, where ψ is a generator of $\widehat{\mathbb{F}_{2^n}^*}$. Taking the equation modulo 2^{d+1}, applying Stickelberger's congruence, Lemma 4.3.12 and using the fact that L_2^* is a

permutation, we get

$$W_G(L_2^*(b), L_1^*(b)) \equiv \sum_{\substack{j:\, j(2^r-1)\equiv 0 \pmod{2^n-1} \\ j\neq 0}} G(\overline{\psi}^j)\psi(L_1^*(b)^j) \pmod{2^{d+1}}.$$

The characters ψ^j such that $j(2^r-1) \equiv 0 \pmod{2^n-1}$ are precisely the characters whose order divides 2^r-1. Since the order of course divides 2^n-1, Lemma 4.3.10 implies that the order of ψ^j also divides 2^d-1, so these characters are precisely the lifted characters from the subfield $\mathbb{F}_{2^d}$. Thus, with the Davenport-Hasse theorem (Theorem 2.5.18), we can write

$$W_G(L_2^*(b), L_1^*(b)) \equiv \sum_{j=1}^{2^d-2} (-1)^{n/d}(G(\overline{\psi'^j}))^{n/d}\psi'(\mathrm{N}_{2^n/2^d}(L_1^*(b)^j)) \pmod{2^{d+1}}, \tag{4.24}$$

where ψ' is a generator of $\widehat{\mathbb{F}_{2^d}^*}$. If $n \leq d^2$, then $n/d \leq d$, so we conclude using Stickelberger's congruence on $\mathbb{F}_{2^d}$ and $\psi'(x) \equiv x \pmod 2$ from Eq. (2.4)

$$\begin{aligned} W_G(L_2^*(b), L_1^*(b)) &\equiv \sum_{j\in\mathbb{Z}_{2^d-1}:\ \mathrm{wt}(j)=1} (-1)^{n/d}2^{n/d}(\mathrm{N}_{2^n/2^d}(L_1^*(b)^j)) \pmod{2^{n/d+1}} \\ &\equiv 2^{n/d}\,\mathrm{Tr}_d(\mathrm{N}_{2^n/2^d}(L_1^*(b))) \pmod{2^{n/d+1}}. \end{aligned}$$

It is well known that the norm function $\mathrm{N}_{2^n/2^d}$ is surjective on $\mathbb{F}_{2^d}^*$, in particular there exists $b \in \mathbb{F}_{2^n}^*$ such that $\mathrm{Tr}_d(\mathrm{N}_{2^n/2^d}(b)) = 1$. We conclude that if L_1 is a permutation, then there exists a $b \in \mathbb{F}_{2^n}$ such that $W_G(L_2^*(b), L_1^*(b)) \not\equiv 0 \pmod{2^{n/r+1}}$ and F' is not a permutation.

Now assume that n is even and $r = d = n/2$. Using Eq. (4.17) we rewrite Eq. (4.24)

$$\begin{aligned} W_G(L_2^*(b), L_1^*(b)) &\equiv \sum_{j=1}^{2^{n/2}-2} (G(\overline{\psi'^j}))^2\psi'(\mathrm{N}_{2^n/2^{n/2}}(L_1^*(b))^j) \pmod{2^{n/2+1}} \\ &\equiv (2^{\frac{n}{2}}-1)K_{n/2}(\mathrm{N}_{2^n/2^{n/2}}(L_1^*(b))) - 2^{n/2} \pmod{2^{n/2+1}} \\ &\equiv -(K_{n/2}(\mathrm{N}_{2^n/2^{n/2}}(L_1^*(b))) + 2^{n/2}) \pmod{2^{n/2+1}}, \end{aligned}$$

where we use that $K_{n/2}(x)$ is always even (indeed, as Theorem 4.2.7 shows, Kloosterman sums are always divisible by 4). The right hand side of the congruence is never zero since $|K_{n/2}(x)| \leq 2^{n/4+1}$ by Theorem 4.2.7 for all $x \in \mathbb{F}_{2^{n/2}}$. We conclude that no permutations of the form $L_1(x^{2^{n/2}-1}) + L_2(x)$ exist. □

Remark 4.3.14. The connection between the exponent $2^{n/2}-1$ and Kloosterman sums is well-known. A classical construction of bent functions by Dillon [44] is that $\mathrm{Tr}_n(ax^{2^{n/2}-1})$ for $a \in \mathbb{F}_{2^{n/2}}$ is bent if and only if $K_{n/2}(a) = 0$. In fact, a proof of this fact using similar techniques used to the ones we used in Theorem 4.3.13 can be found in [93].

The following is a direct consequence of Theorem 4.3.13 and Proposition 4.1.2.

Theorem 4.3.15. *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ defined by $F(x) = x^{2^r-1}$ with $\gcd(n, r) = d$, $n \geq 4$ and $n \leq d^2$. There are no permutations EA-equivalent to F.*
If n is even and $r = n/2$, then the CCZ-class of F coincides with its EA-class and there are no permutations CCZ-equivalent to F.

4.4 Conclusion

We have investigated permutation polynomials of the form $L_1(x^d) + L_2(x)$ over $\mathbb{F}_{2^n}$. We achieved a complete characterization for the inverse function $d = 2^n - 2$ as well as for the Dillon exponent $d = 2^{n/2} - 1$. We also achieved partial results in the more general case $d = 2^r - 1$ and gave a new proof of a known result when d is an APN Gold exponent.

Using these results, we were able to proof that the CCZ-classes of the inverse functions and the function $x \mapsto x^{2^{n/2}-1}$ coincide with their EA-classes. To our knowledge, these are the first proofs of this kind. We also classified all permutations that are CCZ-equivalent to those functions.

An interesting avenue of further research is to consider the same questions for other functions with good cryptographic properties (nonlinearity/differential uniformity). In particular, Theorem 4.2.18 proves Conjecture 4.1.1 for the case of the inverse function. However, all other cases have not been answered yet (to our knowledge). Using the approach in this chapter, a possible way to prove the conjecture would be to prove the non-existence of permutation polynomials of the form $L_1(x^d) + L_2(x)$ with $L_1, L_2 \neq 0$ where x^d is a non-Gold APN monomial. Note however, that the non-existence of such a polynomial is a stronger statement than the statement in Conjecture 4.1.1, i.e. finding a permutation polynomial of the form $L_1(x^d) + L_2(x)$ with nonzero L_1, L_2 does not disprove Conjecture 4.1.1.

More generally, an interesting way to expand on the results in this chapter would be to work towards a classification of permutation polynomials of the form $L_1(x^d) + L_2(x)$ (or even just $x^d + L(x)$) over $\mathbb{F}_{2^n}$ where L, L_1, L_2 are linear mappings. Combined with Proposition 4.1.2, this might give insight into the CCZ-classes of monomials. Note that this family of polynomials is similar to other families that have been investigated in the past, for instance the polynomials $x^s + \gamma \operatorname{Tr}(x^t)$ considered in [36].

It would also be of interest to look at the same problem in odd characteristic. As shown in [60], there are no permutation polynomials of the form $L_1(x^{-1}) + L_2(x)$ in characteristic ≥ 5 which is a straightforward observation based on the fact that there are no non-trivial Kloosterman zeros in finite fields of characteristic ≥ 5. The characteristic 3 case is however still open.

Chapter 5

Autocorrelation of vectorial Boolean functions

The differential-linear attack is one of the standard attacks against block ciphers in modern cryptography. The basic idea behind the attack as introduced in [95] is to split the encryption function E into two parts $E = E_1 \circ E_0$. Very roughly speaking, the attack uses the differential properties of E_0 and the linear properties of E_1 to mount an attack on the entire cipher E. Differential-linear attacks have proven to be very effective and yield some of the best known attacks against many ciphers, for example SERPENT [105], CHASKEY [98] and ICEPOLE [71]. For a very good general overview on the details as well as the history of the differential-linear attack, we refer to [11]. In the traditional differential-linear cryptanalysis, the subciphers E_0 and E_1 are usually assumed to be independent of one another. To account for possible dependencies between the subciphers, in [4] a new decomposition is introduced using a middle layer E_m, i.e. $E = E_1' \circ E_m \circ E_0'$, where E_m accounts for the dependency between the subciphers E_1 and E_0 in the original decomposition. Note that this idea was inspired by a similar approach for boomerang attacks [40, 51]. Using this approach, a new tool for differential-linear cryptanalysis was proposed in [4]: The *differential-linear connectivity table*.

Definition 5.0.1 (Differential-linear connectivity table (DLCT))**.** *Let $F\colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ be a vectorial Boolean function. We define its* differential-linear connectivity table *(*DLCT*) as*

$$\mathsf{DLCT}_F(a,b) = |\{x \in \mathbb{F}_{2^n}\colon \operatorname{Tr}(b(F(x) + F(x+a))) = 0\}| - 2^{n-1}$$

for all $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^m}$.

As described in [4], the DLCT of an S-box as defined in Definition 5.0.1 captures the transition between the two subciphers E_1 and E_0.

A theoretical analysis of the DLCT will lead to more insight into differential-linear attacks on various ciphers. In particular the maximal value of $|\mathsf{DLCT}_F(a,b)|$ for non-zero a, b is of interest. In this chapter, we show connections between the DLCT and established properties of vectorial Boolean functions, in particular the autocorrelation of Boolean functions. We also prove upper bounds for the value of $|\mathsf{DLCT}_F(a,b)|$ both in the general case and for specific interesting choices of F. The work in this chapter is based on the preprint [26] written by Anne Canteaut, Friedrich Wiemer and the author of this thesis. In the same timeframe, another research group has worked on the same problem; a merged paper is available at [25].

5.1 The DLCT and the autocorrelation

Let us first note some trivial things about the DLCT. For any $(a,b) \in \mathbb{F}_{2^n} \times \mathbb{F}_{2^m}$, $|\mathsf{DLCT}_F(a,b)| \leq 2^{n-1}$, and $\mathsf{DLCT}_F(a,b) = 2^{n-1}$ when either $a = 0$ or $b = 0$. Therefore, we only need to focus on the cases for $a \in \mathbb{F}_{2^n}^*$ and $b \in \mathbb{F}_{2^m}^*$.

Our first observation on the DLCT is that it essentially coincides with the *autocorrelation* (AC) of F. Recall the definition of the autocorrelation of Boolean functions, see e.g. [28, p. 277].

Definition 5.1.1 (Autocorrelation of a Boolean function)**.** *Given a Boolean function $f\colon \mathbb{F}_{2^n} \to \mathbb{F}_2$, the* autocorrelation *of the function f at $a \in \mathbb{F}_{2^n}$ is defined as*

$$\mathsf{AC}_f(a) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)+f(x+a)}.$$

Furthermore, the absolute indicator *of f is defined as $\Delta_f = \max_{a \in \mathbb{F}_{2^n}^*} |\mathsf{AC}_f(u)|$.*

This notion can naturally be generalized to vectorial Boolean functions as follows via the component functions.

Definition 5.1.2 (Autocorrelation of a vectorial Boolean function)**.** *Let F be an (n,m)-function. For any $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^m}$, the* autocorrelation *of F at (a,b) is defined as*

$$\mathsf{AC}_F(a,b) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(F(x)+F(x+a)))},$$

and the autocorrelation spectrum *of F is defined as the multiset*

$$\left\{ * \, \mathsf{AC}_F(a,b) \, * \;\middle|\; a \in \mathbb{F}_{2^n}^*, b \in \mathbb{F}_{2^m}^* \right\}.$$

Moreover, the absolute indicator *of F is defined as*

$$\Delta_F = \max_{a \in \mathbb{F}_{2^n}^*, b \in \mathbb{F}_{2^m}^*} |\mathsf{AC}_F(a,b)|.$$

It is worth noticing that we can write the autocorrelation using the Walsh transform and the derivative:

$$\mathsf{AC}_F(a,b) = W_{D_aF}(0,b) = W_{D_aF_b}(0). \tag{5.1}$$

From Definitions 5.0.1 and 5.1.2, we immediately observe the following connection between the DLCT and the autocorrelation of vectorial Boolean functions.

Proposition 5.1.3. *Let F be an (n,m)-function. Then for any $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^m}$, the autocorrelation of F at (a,b) is twice the value of the DLCT of F at the same position (a,b), i.e.,*

$$\mathsf{DLCT}_F(a,b) = \frac{1}{2}\mathsf{AC}_F(a,b) \ .$$

Proof. Define $M_i = \{x \in \mathbb{F}_{2^n} | \operatorname{Tr}(bF(x) + F(x+a)) = i\}$. From the definitions of the DLCT and the autocorrelation it follows that

$$\begin{aligned} 2 \cdot \mathsf{DLCT}_F(a,b) &= 2 \cdot \#\{x \in \mathbb{F}_{2^n} | \operatorname{Tr}(bF(x)) = \operatorname{Tr}(bF(x+a))\} - 2^n \\ &= \#M_0 - (2^n - \#M_0) \\ &= \#M_0 - \#M_1 \\ &= \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\operatorname{Tr}(b(F(x)+F(x+a)))} \\ &= \mathsf{AC}_F(a,b). \end{aligned}$$

This gives the desired conclusion. □

For the remainder of this chapter we thus stick to the more established notion of the autocorrelation instead of the DLCT.

5.1.1 Some characterizations and properties of the autocorrelation

In this subsection, we give some characterizations and properties of the autocorrelation of a vectorial Boolean function relating to its Walsh transform and differential properties.

The following proposition shows that the restriction of the autocorrelation function $a \mapsto \mathsf{AC}_F(a,b)$ can be seen as the discrete Fourier transform of the squared Walsh transform of F_b: $\omega \mapsto W_F(\omega,b)^2$. It should be noted that the relations Eq. (5.2) and Eq. (5.4) were already obtained in [62] and [138] for Boolean functions. Here we generalize the results to vectorial Boolean functions.

Proposition 5.1.4. *Let F be an (n,m)-function. Then for any $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^m}$,*

$$W_F(a,b)^2 = \sum_{\omega \in \mathbb{F}_{2^n}} (-1)^{\operatorname{Tr}(a\omega)} \mathsf{AC}_F(\omega,b).$$

Conversely, the inverse Fourier transform leads to

$$\mathsf{AC}_F(\omega,b) = \frac{1}{2^n} \sum_{a \in \mathbb{F}_{2^n}} (-1)^{\operatorname{Tr}(a\omega)} W_F(a,b)^2 \tag{5.2}$$

Moreover, we have

$$\sum_{a \in \mathbb{F}_{2^n}} \mathsf{AC}_F(a,b) = W_F(0,b)^2 \tag{5.3}$$

and

$$\sum_{a \in \mathbb{F}_{2^n}} \mathsf{AC}_F(a,b)^2 = \frac{1}{2^n} \sum_{\omega \in \mathbb{F}_{2^n}} W_F(\omega,b)^4. \tag{5.4}$$

Proof. According to the definition, for any $a \in \mathbb{F}_{2^n}$,

$$\begin{aligned} W_F(a,b)^2 &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(ax)+\mathrm{Tr}(bF(x))} \sum_{y\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(ay)+\mathrm{Tr}(bF(y))} \\ &= \sum_{x,y\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(a(x+y))+\mathrm{Tr}(b(F(x)+F(y)))} \\ &= \sum_{x,\omega\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(a\omega)+\mathrm{Tr}(b(F(x)+F(x+\omega)))} \\ &= \sum_{\omega\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(a\omega)} \sum_{x\in\mathbb{F}_2^n} (-1)^{\mathrm{Tr}(b(F(x)+F(x+\omega)))} \\ &= \sum_{\omega\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(a\omega)} \mathsf{AC}_F(\omega,b). \end{aligned}$$

The inverse Fourier Transform (Proposition 2.5.9) then leads to Eq. (5.2). Then Eq. (5.3) is obtained from Eq. (5.2) by summing over u. Furthermore, Parseval's identity (Proposition 2.5.11) leads to Eq. (5.4). □

It was shown in [139, Section 3] that, for an (n,n)-function, the mapping $b \mapsto \mathsf{AC}_F(a,b)$ corresponds to the Fourier transform of the mapping $b \mapsto \delta_F(a,b)$. This relation coincides with the one provided in [4, Proposition 1]. We slightly extend it here to the case of (n,m)-functions. It is worth noticing that this correspondence together with Proposition 5.1.4 points out the relation between the Walsh transform of F and the values of δ_F as shown in Proposition 2.2.17.

Proposition 5.1.5. *Let F be an (n,m)-function. Then, for any $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^m}$, we have*

$$\mathsf{AC}_F(a,b) = \sum_{\omega\in\mathbb{F}_{2^m}} (-1)^{\mathrm{Tr}(b\omega)} \delta_F(a,\omega) \tag{5.5}$$

$$\delta_F(a,b) = 2^{-m} \sum_{\omega\in\mathbb{F}_{2^m}} (-1)^{\mathrm{Tr}(b\omega)} \mathsf{AC}_F(a,\omega). \tag{5.6}$$

Most notably,

$$\sum_{b\in\mathbb{F}_{2^m}} \mathsf{AC}_F(a,b) = 2^m \delta_F(a,0) \tag{5.7}$$

implying

$$\sum_{a\in\mathbb{F}_{2^n}, b\in\mathbb{F}_{2^m}} \mathsf{AC}_F(a,b) = 2^{m+n}, \tag{5.8}$$

and

$$\sum_{a\in\mathbb{F}_{2^m}} \mathsf{AC}_F(a,b)^2 = 2^m \sum_{\omega\in\mathbb{F}_{2^m}} \delta_F(a,\omega)^2. \tag{5.9}$$

Proof. The first equation holds since

$$\begin{aligned} \mathsf{AC}_F(a,b) &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(F(x)+F(x+a)))} \\ &= \sum_{\omega\in\mathbb{F}_{2^m}} (-1)^{\mathrm{Tr}(b\omega)} \delta_F(a,\omega)\ . \end{aligned}$$

The inverse Fourier transform then leads to Eq. (5.6). Setting $b = 0$ in Eq. (5.6), we then get Eq. (5.7). Summing over all a in Eq. (5.7) yields Eq. (5.8). Finally, Parseval's identity implies Eq. (5.9). □

Remark 5.1.6. Note that in [115, 117, 111] the boomerang connectivity table (BCT) (which measures the resistance of a vectorial Boolean function to a boomerang attack) is linked to the differential properties $\delta(a, b)$. Proposition 5.1.5 then also establishes a link between the BCT and the autocorrelation. For further details about this connection, we refer to [117], in particular Proposition 1.

5.1.2 Bounds on the absolute indicator

Similar to other cryptographic criteria, it is interesting and important to know how "good" the absolute indicator of a vectorial Boolean function could be. It is clear from the definition that the absolute indicator Δ_F of any (n, m)-function is upper bounded by 2^n. Indeed, it is possible to characterize the functions for which this value is attained.

Recall that a linear structure for an (n, m)-function F is a tuple $(a, b) \in \mathbb{F}_{2^n} \times \mathbb{F}_{2^m}$ such that $x \mapsto \mathrm{Tr}(b(F(x) + F(x + a)))$ is constant. The connection between linear structures and the autocorrelation is then easy to see and follows directly from the definition of the autocorrelation.

Lemma 5.1.7. *Let F be an (n, m)-function. Then $\mathsf{AC}_F(a, b) = \pm 2^n$ if and only if (a, b) is a linear structure for F. In particular, $\Delta_F = 2^n$ if and only if F has a linear structure $(a, b) \in \mathbb{F}_{2^n}^* \times \mathbb{F}_{2^m}^*$.*

Finding lower bounds of the absolute indicator is an interesting open question. Of course, the trivial lowest possible value for the absolute indicator is 0. It can be easily determined for which functions this value is attained.

Proposition 5.1.8. *Let F be an (n, m)-function. Then $\Delta_F = 0$ if and only if F is a vectorial bent function.*

Proof. From Eq. (5.1) we have $\mathsf{AC}_F(a, b) = W_{D_aF}(0, b)$. In particular $\mathsf{AC}_F(a, b) = 0$ for all $a \in \mathbb{F}_{2^n}^*$ and $b \in \mathbb{F}_{2^m}^*$ if and only if all derivatives are balanced. By Proposition 2.2.10 this is equivalent to F being a vectorial bent function. □

By the Nyberg bound we conclude that the absolute indicator cannot attain the lowest possible value 0 if n is odd or $m > n/2$. In particular, the most interesting case of (n, n)-functions is not covered. Finding lower bounds outside of the one given in Proposition 5.1.8 is more challenging. For instance, Zhang and Zheng conjectured [138, Conjecture 1] that the absolute indicator of a *balanced* Boolean function in n variables was at least $2^{\frac{n+1}{2}}$. This was later disproved first for odd values of $n \geq 9$ by modifying the Patterson-Wiedemann construction, namely for $n \in \{9, 11\}$ in [78], for $n = 15$ in [106, 76] and for $n = 21$ in [54]. For the case n even, [129] gave a construction for balanced Boolean functions with absolute indicator strictly less than $2^{n/2}$ when $n \equiv 2 \pmod 4$. Very recently, similar examples for $n \equiv 0 \pmod 4$ were exhibited by [77]. However, we now show that such small values for the absolute indicator cannot be achieved for (n, n)-*vectorial functions*.

Proposition 5.1.5 leads to the following upper bound on the sum of all squared autocorrelation coefficients in each row. This result is an extension of a result in [116] (see also [6, Theorem 2]) where it was proven in the case of (n, n)-functions.

Proposition 5.1.9. *Let F be an (n, m)-function. Then, for all $a \in \mathbb{F}_{2^n}$, we have*

$$\sum_{b \in \mathbb{F}_{2^m}} \mathsf{AC}_F(a, b)^2 \geq 2^{n+m+1} .$$

Moreover, equality holds for all nonzero $a \in \mathbb{F}_{2^n}$ if and only if F is APN.

Proof. From Eq. (5.9), we have that, for all $a \in \mathbb{F}_{2^n}$,

$$\sum_{b \in \mathbb{F}_{2^m}} \mathsf{AC}_F(a,b)^2 = 2^m \sum_{\omega \in \mathbb{F}_{2^m}} \delta_F(a,\omega)^2 \tag{5.10}$$

Recall that $\sum_{\omega \in \mathbb{F}_{2^m}} \delta(a,\omega) = 2^n$. The Cauchy-Schwarz inequality implies that

$$2^{2n} = \left(\sum_{\omega \in \mathbb{F}_{2^m}} \delta_F(a,\omega) \right)^2 \leq \left(\sum_{\omega \in \mathbb{F}_{2^m}} (\delta_F(a,\omega))^2 \right) \times \#\{\omega \in \mathbb{F}_{2^m} | \delta_F(a,\omega) \neq 0\} ,$$

with equality if and only if all nonzero elements in $\{\delta_F(a,\omega) | \omega \in \mathbb{F}_{2^m}\}$ are equal. Since $\delta_F(a,\omega)$ is even, we infer

$$\#\{\omega \in \mathbb{F}_{2^m} | \delta_F(a,\omega) \neq 0\} \leq 2^{n-1}$$

with equality for all nonzero a if and only if F is APN. We deduce that

$$\sum_{\omega \in \mathbb{F}_{2^m}} (\delta_F(a,\omega))^2 \geq 2^{n+1}$$

with equality for all nonzero a if and only if F is APN. With Eq. (5.10) we deduce that

$$\sum_{b \in \mathbb{F}_{2^m}} \mathsf{AC}_F^2(a,b) \geq 2^{n+m+1}$$

with equality for all nonzero a if and only if F is APN. □

We want to note that Proposition 5.1.9 is trivial if $m < n$ since $\mathsf{AC}_F(a,0)^2 = 2^{2n}$ for any $a \in \mathbb{F}_{2^n}$.

If $m \geq n$, we can use Proposition 5.1.9 to give a lower bound for the absolute indicator.

Theorem 5.1.10. *Let F be an (n,m)-function, where $m \geq n$. Then*

$$\Delta_F \geq \sqrt{\frac{2^{m+n+1} - 2^{2n}}{2^m - 1}}.$$

Most notably, if $m = n$,

$$\Delta_F > 2^{n/2} .$$

Proof. From the previous proposition, we deduce that

$$\sum_{b \in \mathbb{F}_{2^m}^*} \mathsf{AC}_F(a,b)^2 \geq 2^{n+m+1} - 2^{2n} .$$

Since clearly

$$\sum_{b \in \mathbb{F}_{2^m}^*} \mathsf{AC}_F(a,b)^2 \leq (2^m - 1)\Delta_F^2 ,$$

we get

$$\Delta_F \geq \sqrt{\frac{2^{m+n+1} - 2^{2n}}{2^m - 1}}.$$

□

A general lower bound if $n/2 < m < n$ or $m < n$ with n odd has not been found yet.

Problem 5.1.11. Find a lower bound for the absolute indicator of an (n, m)-function in the cases that $n/2 < m < n$ or $m < n$ with n odd.

5.1.3 Invariance of the autocorrelation under Equivalence Relations

In this subsection, we study the autocorrelation and the absolute indicator with respect to the equivalence relations introduced in Chapter 2. Our first result is that the autocorrelation behaves well under EA-equivalence.

Theorem 5.1.12. *Let F be an (n, m)-function. Further, let $A_1 \colon \mathbb{F}_{2^m} \to \mathbb{F}_{2^m}$, $A_2 \colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be two affine permutations defined by $A_1 = L_1 + a_1$, $A_2 = L_2 + a_2$ with linear parts L_1, L_2 and $A_3 \colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ be an affine function defined by $A_3 = L_3 + a_3$ with linear part L_3.*

Set $F' = A_1 \circ F \circ A_2 + A_3$. Then

$$\mathsf{AC}_{F'}(a, b) = (-1)^{\mathrm{Tr}(bL_3(a))} \mathsf{AC}_F(L_2(a), L_1^*(b)).$$

In particular, the absolute indicator is invariant under EA-equivalence and the autocorrelation spectrum is invariant under affine equivalence.

Proof. For any $a \in \mathbb{F}_{2^n}^*$ and $b \in \mathbb{F}_{2^m}^*$ we have

$$\begin{aligned}
\mathsf{AC}_{F'}(a, b) &= \sum_{x \in \mathbb{F}_2^n} (-1)^{\mathrm{Tr}(b(F'(x)+F'(x+a)))} \\
&= \sum_{x \in \mathbb{F}_2^n} (-1)^{\mathrm{Tr}(b(A_1 \circ F \circ A_2(x)+A_3(x)+A_1 \circ F \circ A_2(x+a)+A_3(x+a)))} \\
&= (-1)^{\mathrm{Tr}(bL_3(a))} \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(A_1 \circ F \circ A_2(x)+A_1 \circ F \circ A_2(x+a)))} \\
&= (-1)^{\mathrm{Tr}(bL_3(a))} \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(L_1(F \circ A_2(x)+F \circ A_2(x+a))))} \\
&= (-1)^{\mathrm{Tr}(bL_3(a))} \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(L_1^*(b)(F \circ A_2(x)+F \circ A_2(x+a)))} \\
&= (-1)^{\mathrm{Tr}(bL_3(a))} \sum_{y \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(L_1^*(b)(F(y)+F(y+L_2(a))))} \\
&= (-1)^{\mathrm{Tr}(bL_3(a))} \mathsf{AC}_F(L_2(a), L_1^*(b)),
\end{aligned}$$

where we set $y = A_2(x)$ and used that A_1, A_2 are permutations. □

To examine the behavior under CCZ-equivalence, we first focus on the autocorrelation of a permutation and its compositional inverse. When $n = m$ and F permutes $\mathbb{F}_{2^n}$, it was shown in [139, Corollary 1] that (using our notation)

$$\mathsf{AC}_{F^{-1}}(a, b) = \frac{1}{2^n} \sum_{u, v \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(bu+av)} \mathsf{AC}_F(u, v). \tag{5.11}$$

The relation in Eq. (5.11) indicates that the autocorrelation spectrum of an (n, n)-permutation F should in general not be equal to that of F^{-1}.

This observation is indeed confirmed by many examples. In particular, by Lemma 5.1.7 the absolute indicator of F attains the maximal value if and only if F has

a nontrivial linear structure. However, there are many examples of permutations F of $\mathbb{F}_{2^n}$ where F has nontrivial linear structures but its inverse has not.

This even happens for functions that are used in practice as S-boxes in block ciphers. For instance, the S-boxes from SAFER [107], SC2000 [123], and FIDES [8] have linear structures in one direction but not in the other direction. Another example where Δ_F does not coincide with $\Delta_{F^{-1}}$ is the infinite family formed by the Gold permutations as analyzed in Section 5.2.2. Summarizing these results, we get:

Proposition 5.1.13. *The autocorrelation spectrum and the absolute indicator are generally not invariant under inversion. In particular, the autocorrelation spectrum and the absolute indicator are generally not invariant under CCZ-equivalence.*

In [97] all optimal permutations over $\mathbb{F}_2^4$ having the best differential uniformity and nonlinearity (both 4) were classified up to affine equivalence. There are only 16 different optimal S-boxes, see Table 5.1.

Based on the classification of optimal S-boxes, we compute the autocorrelation spectra of the optimal S-boxes in Table 5.2, where the superscript of each autocorrelation value indicates the number of its occurrences in the spectrum. Note in particular that the absolute indicators of the functions given in Table 5.2 is not always the same.

TABLE 5.1: Representatives for all 16 classes of optimal 4 bit S-boxes

F_0	$0,1,2,13,4,7,15,6,8,11,12,9,3,14,10,5$
F_1	$0,1,2,13,4,7,15,6,8,11,14,3,5,9,10,12$
F_2	$0,1,2,13,4,7,15,6,8,11,14,3,10,12,5,9$
F_3	$0,1,2,13,4,7,15,6,8,12,5,3,10,14,11,9$
F_4	$0,1,2,13,4,7,15,6,8,12,9,11,10,14,5,3$
F_5	$0,1,2,13,4,7,15,6,8,12,11,9,10,14,3,5$
F_6	$0,1,2,13,4,7,15,6,8,12,11,9,10,14,5,3$
F_7	$0,1,2,13,4,7,15,6,8,12,14,11,10,9,3,5$
F_8	$0,1,2,13,4,7,15,6,8,14,9,5,10,11,3,12$
F_9	$0,1,2,13,4,7,15,6,8,14,11,3,5,9,10,12$
F_{10}	$0,1,2,13,4,7,15,6,8,14,11,5,10,9,3,12$
F_{11}	$0,1,2,13,4,7,15,6,8,14,11,10,5,9,12,3$
F_{12}	$0,1,2,13,4,7,15,6,8,14,11,10,9,3,12,5$
F_{13}	$0,1,2,13,4,7,15,6,8,14,12,9,5,11,10,3$
F_{14}	$0,1,2,13,4,7,15,6,8,14,12,11,3,9,5,10$
F_{15}	$0,1,2,13,4,7,15,6,8,14,12,11,9,3,10,5$

TABLE 5.2: Autocorrelation spectrum of F_i for $0 \leq i \leq 15$

F_i	Autocorrelation spectrum
$i \in \{3,4,5,6,7,11,12,13\}$	$\{-8^{60}, 0^{135}, 8^{30}\}$
$i \in \{0,1,2,8\}$	$\{-16^6, -8^{48}, 0^{144}, 8^{24}, 16^3\}$
$i \in \{9,10,14,15\}$	$\{-16^2, -8^{56}, 0^{138}, 8^{28}, 16^1\}$

5.1.4 Divisibility properties of the autocorrelation

We now want to investigate the divisibility property of the autocorrelation coefficients of vectorial Boolean functions.

Proposition 5.1.14. *Let $n > 2$ and $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ be a vectorial Boolean function with algebraic degree $d > 1$. Then, for any $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^m}$, $\mathsf{AC}_F(a,b)$ is divisible by $2^{\lceil \frac{n-1}{d-1} \rceil + 1}$. In particular, when $m = n$ and F is a permutation, $\mathsf{AC}_F(a,b)$ is divisible by 8.*

Proof. By Eq. (5.1), for any $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^m}$, we have

$$\mathsf{AC}_F(a,b) = W_{D_aF_b}(0).$$

Note that for given $a \in \mathbb{F}_{2^n}$ and $b \in \mathbb{F}_{2^m}$, the Boolean function

$$D_aF_b = \mathrm{Tr}(b(F(x) + F(x+a)))$$

satisfies two properties: its algebraic degree is at most $d-1$ since it is the derivative of F_b which has algebraic degree at most d, and $D_aF_b(x) = D_aF_b(x+a)$.

We now focus on the divisibility of $W_{D_aF_b}(0)$. We can decompose $\mathbb{F}_{2^n} = \{0, a\} \oplus V$ for some subspace V of dimension $n-1$. Since $D_aF_b(x+a) = D_aF_b(x)$, the value of $D_aF_b(x)$ on the entire field $\mathbb{F}_{2^n}$ is already completely determined by its value on V. Hence $D_aF_b(x)$ can be expressed as $D_aF_b(x) = h(x') \colon V \to \mathbb{F}_2$ and the Walsh transform of D_aF_b at the point 0 satisfies

$$W_{D_aF_b}(0) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{D_aF_b(x)} = \sum_{x' \in V, e \in \mathbb{F}_2} (-1)^{D_aF_b(x'+ea)} = 2 \cdot \sum_{x' \in V} (-1)^{h(x')}.$$

It is well-known that the values taken by the Walsh transform of a Boolean function f from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ with algebraic degree d are divisible by $2^{\lceil \frac{n}{d-1} \rceil}$ (see [110] or [28, Section 3.1]). Recall that the autocorrelation spectrum is invariant under affine equivalence by Theorem 5.1.12 and V is isomorphic to $\mathbb{F}_2^{n-1}$ as a vector space over $\mathbb{F}_2$. In particular, $\sum_{x' \in V} (-1)^{h(x')}$ is divisible by $2^{\lceil \frac{n-1}{d-1} \rceil}$, implying that $W_{D_aF_b}(0)$ is divisible by $2^{\lceil \frac{n-1}{d-1} \rceil + 1}$.

If $m = n$ and F is bijective, then $d < n$ by Proposition 2.2.5. We then have that

$$\left\lceil \frac{n-1}{d-1} \right\rceil \geq 2,$$

implying that $\mathsf{AC}_F(a,b)$ is divisible by 8. □

For cubic (n,m)-functions, we have the following stronger result.

Proposition 5.1.15. *Suppose an (n,m)-function F has algebraic degree 3. Then for nonzero a and b, we have*

$$|\mathsf{AC}_F(a,b)| \in \left\{0, 2^{\frac{n+t(a,b)}{2}}\right\},$$

where $t(a,b) = \dim \{w \in \mathbb{F}_{2^n} \mid D_wD_aF_b = c\}$ and $c \in \mathbb{F}_2$ is constant.

Proof. Since F has algebraic degree 3, the derivatives of order two of the component functions F_b, namely $D_wD_aF_b(x)$, are affine over $\mathbb{F}_{2^n}$, so we can write them as $D_wD_aF_b(x) = \mathrm{Tr}(A_{a,b}(w)x + C_{a,b}(w))$, where $w \mapsto A_{a,b}(w)$ and $w \mapsto C_{a,b}(w)$ are functions mapping $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^n}$. Moreover, the function $w \mapsto C_{a,b}(w)$ is linear over the linear subspace $L(a,b) = \{w \in \mathbb{F}_{2^n} : A_{a,b}(w) = 0\} = \{w \in \mathbb{F}_{2^n} : D_wD_aF_b(x) =$

$\mathrm{Tr}(C_{a,b}(w))\}$. From the definition of the autocorrelation, we have

$$\begin{aligned}
\mathsf{AC}_F(a,b)^2 &= \left(\sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(F(x+a)+F(x)))}\right)^2 \\
&= \sum_{x,y\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(F(x+a)+F(x)+F(y+a)+F(y)))} \\
&= \sum_{x,w\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(F(x+a)+F(x)+F(x+w+a)+F(x+w)))} \\
&= \sum_{x,w\in\mathbb{F}_{2^n}} (-1)^{D_w D_a F_b(x)} \\
&= \sum_{w\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(C_{a,b}(w))} \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(A_{a,b}(w)x)}.
\end{aligned}$$

Hence,

$$\mathsf{AC}_F(a,b)^2 = \begin{cases} 0, & \text{if } A_{a,b}(w) \neq 0, \\ 2^{n+t(a,b)}, & \text{if } A_{a,b}(w) = 0. \end{cases}$$

The desired conclusion directly follows. □

Proposition 5.1.15 implies that any entry in the autocorrelation of a cubic function is divisible by $2^{\frac{n+t}{2}}$, where t is the smallest integer among the $t(a,b)$ when a,b run through $\mathbb{F}_{2^n}^*$ and $\mathbb{F}_{2^m}^*$, respectively. If $t \geq 2$, Proposition 5.1.15 improves the result in Proposition 5.1.14.

5.1.5 Autocorrelation of APN functions

Proposition 5.1.16. *Let F be an APN function from $\mathbb{F}_{2^n}$ to itself. For any nonzero $a \in \mathbb{F}_{2^n}$, we define the Boolean function*

$$\gamma_a(x) = \begin{cases} 1, & \text{if } x \in \mathrm{im}(D_aF), \\ 0, & \text{if } x \in \mathbb{F}_{2^n} \setminus \mathrm{im}(D_aF). \end{cases} \tag{5.12}$$

Then the autocorrelation of F can be expressed by the Walsh transform of γ_a as

$$\mathsf{AC}_F(a,b) = -W_{\gamma_a}(b).$$

Proof. Since F is APN, we know that $\mathrm{im}(D_aF)$ has cardinality 2^{n-1} for each $a \in \mathbb{F}_{2^n}^*$. Then,

$$\begin{aligned}
\mathsf{AC}_F(a,b) &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(F(x+a)+F(x)))} \\
&= 2\sum_{y\in\mathrm{im}(D_aF)} (-1)^{\mathrm{Tr}(by)} \\
&= \sum_{y\in\mathrm{im}(D_aF)} (-1)^{\mathrm{Tr}(by)} - \sum_{y\in\mathbb{F}_{2^n}\setminus\mathrm{im}(D_aF)} (-1)^{\mathrm{Tr}(by)} \\
&= -\sum_{y\in\mathbb{F}_{2^n}} (-1)^{\gamma_a(y)+\mathrm{Tr}(by)} \\
&= -W_{\gamma_a}(b).
\end{aligned}$$

□

From Proposition 5.1.16, we see that the autocorrelation of any APN function corresponds to the Walsh transform of the Boolean function γ_a in Eq. (5.12), which is balanced. We then immediately deduce the following Corollary.

Corollary 5.1.17 (Lowest possible absolute indicator for APN functions)**.** *Let n be a positive integer. If there exists an APN function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^n}$ with absolute indicator Δ, then there exists a balanced Boolean function of n variables with linearity Δ.*

To our best knowledge, the smallest known linearity for a balanced function is obtained by Dobbertin's recursive construction [49]. For instance, for $n = 9$, the smallest possible linearity for a balanced Boolean function is known to belong to the set $\{24, 28, 32\}$, which implies that exhibiting an APN function over $\mathbb{F}_{2^9}$ with absolute indicator 24 would determine the smallest linearity for such a function.

One of the functions whose absolute indicator is known is the inverse mapping $F(x) = x^{2^n-2}$ over $\mathbb{F}_{2^n}$.

Proposition 5.1.18 ([38, Corollary 2])**.** *The autocorrelation spectrum of the inverse function $F(x) = x^{2^n-2}$ over $\mathbb{F}_{2^n}$ is given by*

$$\left\{K_n(a) + 2 \times ((-1)^{\mathrm{Tr}(a)} - 1) \middle| a \in \mathbb{F}_{2^n}^*\right\},$$

where $K_n(a)$ denotes the Kloosterman sum over $\mathbb{F}_{2^n}$. Furthermore, the absolute indicator of the inverse function is given by:

1. *when n is even, $\Delta_F = 2^{\frac{n}{2}+1}$;*
2. *when n is odd, $\Delta_F = L(F)$ if $L(F) \equiv 0 \pmod 8$, and $\Delta_F = L(F) \pm 4$ otherwise.*

When n is odd, the inverse mapping is APN. Then, from Proposition 5.1.16, its autocorrelation is directly determined by the corresponding γ. This explains why the absolute indicator of the inverse mapping when n is odd, is derived from its linearity as detailed in the following example.

Example 5.1.19 (Autocorrelation of the inverse mapping, n odd)**.** Let $F = x^{2^n-2}$ be the inverse function with n odd. Let us first determine the Boolean function γ_a of F as it is defined in Proposition 5.1.16.

We consider for $a \in \mathbb{F}_{2^n}^*$ the equation

$$D_a F(x) = (x + a)^{-1} + x^{-1} = b$$

which, for $x \neq a$ and $x \neq 0$, can be rewritten as

$$x + (x + a) = b(x + a)x$$

and by setting $y = a^{-1}x$ when $b \neq 0$,

$$y^2 + y = a^{-1}b^{-1}\,.$$

It follows that this equation has two solutions if and only if $\mathrm{Tr}(a^{-1}b^{-1}) = 0$.

We conclude that γ_a coincides with $(1 + F_{a^{-1}})$ except on two points:

$$\gamma_a(x) = \begin{cases} 1 + \mathrm{Tr}(a^{-1}x^{-1}), & \text{if } x \notin \{0, a^{-1}\} \\ 0, & \text{if } x = 0 \\ 1, & \text{if } x = a^{-1}. \end{cases}$$

From Proposition 5.1.16, we deduce

$$\begin{aligned}\mathsf{AC}_F(a,b) &= -W_{\gamma_a}(b)\\ &= W_{F_{a^{-1}}}(b) + 2\left(1-(-1)^{\mathrm{Tr}(a^{-1}b)}\right),\end{aligned}$$

where the additional term corresponds to the value of the sum defining the Walsh transform $W_{F_{a^{-1}}}(b)$ at points 0 and a^{-1}.

Rewriting the Walsh transform of the inverse function via Kloosterman sums, we get

$$\mathsf{AC}_F(a,b) = K_n(a^{-1}b) + 2\left(1-(-1)^{\mathrm{Tr}(a^{-1}b)}\right), \tag{5.13}$$

so $\Delta_F = \max_{a\in\mathbb{F}_{2^n}^*} |K_n(a) + 2\left(1-(-1)^{\mathrm{Tr}(a)}\right)|$. This is precisely Eq. (14) in the original proof of Proposition 5.1.18 in [38, Corollary 2] and the rest of the proof proceeds identically from there.

Remark 5.1.20. Note that the linearity of the inverse function $F = x^{2^n-2}$ over $\mathbb{F}_{2^n}$ can be determined very easily because of the link to Kloosterman sums. Indeed,

$$L(F) = \max_{a\in\mathbb{F}_{2^n}, b\in\mathbb{F}_{2^n}^*} |W_F(a,b)| = \max_{a\in\mathbb{F}_{2^n}, b\in\mathbb{F}_{2^n}^*} |K_n(ab)|.$$

Then from Theorem 4.2.7

$$L(F) = \begin{cases} 2^{\frac{n+2}{2}}, & n \text{ even}\\ \max\{k\colon k\equiv 0 \pmod 4 \text{ and } k < 2^{\frac{n+2}{2}}\}, & n \text{ odd.}\end{cases}$$

5.2 Autocorrelation spectra and absolute indicator of special polynomials

This section mainly considers some polynomials of special forms. Explicitly, we will investigate the autocorrelation spectra and the absolute indicator of the Gold permutations and their inverses, and of the Bracken-Leander functions. Our study is divided into two subsections.

5.2.1 Monomials

In the subsection, we consider the autocorrelation of some special monomials of cryptographic interest, mainly APN permutations and one permutation with differential uniformity 4, over the finite field $\mathbb{F}_{2^n}$. Firstly, we present a general observation on the autocorrelation of monomials.

Proposition 5.2.1. *Let $F(x) = x^d$ be a monomial defined on $\mathbb{F}_{2^n}$. Then*

$$\mathsf{AC}_F(a,b) = \mathsf{AC}_F\left(1, ba^d\right).$$

Moreover, if $\gcd(d, 2^n-1) = 1$, *then*

$$\mathsf{AC}_F(1,b) = \mathsf{AC}_F(b^{1/d}, 1),$$

where $1/d$ denotes the inverse of d modulo 2^n-1.

Proof. For any $a, b \in \mathbb{F}_{2^n}^*$, we have

$$\begin{aligned}
\mathsf{AC}_F(a,b) &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(F(x)+F(x+a)))} \\
&= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}\left(b\left(x^d+(x+a)^d\right)\right)} \\
&= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}\left(ba^d\left(\left(\frac{x}{a}\right)^d+\left(\frac{x}{a}+1\right)^d\right)\right)} \\
&= \mathsf{AC}_F\left(1, ba^d\right).
\end{aligned}$$

Moreover, if $\gcd(d, 2^n-1) = 1$, then for any $b \in \mathbb{F}_{2^n}^*$, there exists a unique element $a \in \mathbb{F}_{2^n}^*$ such that $b = a^d$. Furthermore,

$$\begin{aligned}
\mathsf{AC}_F(1,b) &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}\left(b\left(x^d+(x+1)^d\right)\right)} \\
&= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}\left((ax)^d+(ax+a)^d\right)} \\
&= \sum_{y\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}\left(y^d+(y+a)^d\right)} \\
&= \mathsf{AC}_F(a,1).
\end{aligned}$$

□

Proposition 5.2.1 in particular states that, to determine the absolute indicator of a bijective monomial, it suffices to consider one of its component functions.

We next discuss the autocorrelation of some cubic monomials. From Proposition 5.1.15, if $n = m$ is odd, we obviously have that $\Delta_F \geq 2^{\frac{n+1}{2}}$. Furthermore, the equality is achieved when $\dim(\{w \in \mathbb{F}_2^n \mid D_w D_a f_b = c\}) = 1$ for all nonzero a and b. An upper bound on the absolute indicator can be established for two cubic APN permutations, namely the cubic Kasami power function $x \mapsto x^{13}$ and the Welch function. We denote the Kasami exponents by $K_r = 2^{2r} - 2^r + 1$ and the Welch exponent by $W = 2^{(n-1)/2} + 3$ for n odd.

Proposition 5.2.2 ([29, Lemma 1])**.** *The absolute indicator for the Welch function F defined by $x \mapsto x^W$ on $\mathbb{F}_{2^n}$ for n odd is upper bounded by*

$$\Delta_F \leq 2^{\frac{n+5}{2}}.$$

As long as the (regular) degree of the derivatives is small compared to the field size, the Weil bound gives a nontrivial upper bound for the absolute indicator of a vectorial Boolean function. This is particularly interesting for the Kasami functions as the Kasami exponents do not depend on the field size (contrary to for example the Welch exponent).

Proposition 5.2.3. *Let F_r be the r-th Kasami function on $\mathbb{F}_{2^n}$, i.e. $F_r(x) = x^{K_r}$. The absolute indicator of F_r is upper bounded by*

$$\Delta_{F_r} \leq (2^{2r} - 2^{r+1}) \times 2^{\frac{n}{2}}.$$

In particular,

$$\Delta_{F_2} \leq 2^{\frac{n+5}{2}}.$$

Proof. Note that the two exponents with the highest degree of any derivative of F_r are $2^{2r} - 2^r$ and $2^{2r} - 2^{r+1} + 1$. The first exponent is even, so it can be reduced using the relation $\mathrm{Tr}(x^2) = \mathrm{Tr}(x)$. The result then follows from the Weil bound (Theorem 2.5.15).

K_2 has binary weight 3, so combining the bound with Proposition 5.1.15 yields the bound on the absolute indicator of F_2. □

Some other results on the autocorrelations of Boolean functions $\mathrm{Tr}(x^d)$ are known in the literature, which can be trivially extended to the vectorial functions $x \mapsto x^d$ if $\gcd(d, 2^n - 1) = 1$ using Proposition 5.2.1. Examples for such results can be found in [62, Theorem 5], [29] and [127, Lemmas 2 and 3]. In the case $n = 6r$ and $d = 2^{2r} + 2^r + 1$, the power monomial x^d is not a permutation, but results for all component functions of x^d were derived in [24]. We summarize these results about the absolute indicator in the following proposition.

Proposition 5.2.4. *Let $F(x) = x^d$ be a function on $\mathbb{F}_{2^n}$.*

1. *If n is odd and $d = 2^r + 3$ with $r = \frac{n+1}{2}$, then $\Delta_F \in \{2^{\frac{n+1}{2}}, 2^{\frac{n+3}{2}}\}$.*
2. *If n is odd and d is the r-th Kasami exponent, where $3r \equiv \pm 1 \pmod{n}$, then $\Delta_F = 2^{\frac{n+1}{2}}$.*
3. *If $n = 2m$ and $d = 2^{m+1} + 3$, then $\Delta_F \leq 2^{\frac{3m}{2}+1}$.*
4. *If $n = 2m$, m odd and $d = 2^m + 2^{\frac{m+1}{2}} + 1$, then $\Delta_F \leq 2^{\frac{3m}{2}+1}$.*
5. *If $n = 6r$ and $d = 2^{2r} + 2^r + 1$, then $\Delta_F = 2^{5r}$.*

We now provide a different proof of the second case in the previous proposition that additionally relates the autocorrelation of this special Kasami function with the Walsh spectrum of a Gold function. We use the following result on the Walsh transform of this Kasami exponent.

Proposition 5.2.5 ([45])**.** *Let n be odd, not divisible by 3 and $3r \equiv \pm 1 \pmod{n}$. Define the Boolean function $f = \mathrm{Tr}(x^{K_r})$ on $\mathbb{F}_{2^n}$, where $K_r = 2^{2r} - 2^r + 1$ is the r-th Kasami exponent. Then*

$$W_f(x) = 0 \iff \left\{x \mid \mathrm{Tr}(x^{2^r+1}) = 0\right\}.$$

Proposition 5.2.6. *Let n be odd, not divisible by 3 and $3r \equiv \pm 1 \pmod{n}$. Define $F = x^{K_r}$ on $\mathbb{F}_{2^n}$. Then for $a, b \in \mathbb{F}_{2^n}^*$*

$$\mathsf{AC}_F(a, b) = -\sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(ab^{1/K_r}x + x^{2^r+1})},$$

where $1/K_r$ denotes the inverse of K_r in $\mathbb{Z}_{2^n-1}$. In particular, $\Delta_F = 2^{\frac{n+1}{2}}$.

Proof. It is well-known that, if F is a power permutation over a finite field, its Walsh spectrum is uniquely defined by the entries $W_F(a, 1)$. Indeed, for $b \neq 0$,

$$\begin{aligned} W_F(a, b) &= \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(bx^{K_r} + ax)} \\ &= \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(x^{K_r} + ab^{-1/K_r}x)} = W_F(ab^{-1/K_r}, 1) \in \{0, \pm 2^{\frac{n+1}{2}}\}, \end{aligned}$$

where the last fact follows because the Kasami function is almost bent. Then, by Eq. (5.2) in Proposition 5.1.4 and Proposition 5.2.5

$$\begin{aligned}\mathsf{AC}_F(a,b) &= 2^{-n}\sum_{u\in\mathbb{F}_{2^n}}(-1)^{\mathrm{Tr}(au)}W_F^2(u,b) = 2^{-n}\sum_{u\in\mathbb{F}_{2^n}}(-1)^{\mathrm{Tr}(au)}W_F^2(ub^{-1/K_r},1)\\ &= 2\sum_{u\in B}(-1)^{\mathrm{Tr}(au)}\,, \end{aligned} \tag{5.14}$$

where $B = \{u \in \mathbb{F}_{2^n} \colon \mathrm{Tr}((ub^{-1/K_r})^{2^r+1}) = 1\}$. We have

$$\begin{aligned}0 &= \sum_{u\notin B}(-1)^{\mathrm{Tr}(au)} + \sum_{u\in B}(-1)^{\mathrm{Tr}(au)}\\ &= \sum_{u\notin B}(-1)^{\mathrm{Tr}(au+(ub^{-1/K_r})^{2^r+1})} - \sum_{u\in B}(-1)^{\mathrm{Tr}(au+(ub^{-1/K_r})^{2^r+1})}\,,\end{aligned}$$

so

$$\sum_{u\in\mathbb{F}_{2^n}}(-1)^{\mathrm{Tr}(au+(ub^{-1/K_r})^{2^r+1})} = 2\sum_{u\in B}(-1)^{\mathrm{Tr}(au+(ub^{-1/K_r})^{2^r+1})} = -2\sum_{u\in B}(-1)^{\mathrm{Tr}(au)}\,.$$

Plugging this into Eq. (5.14), we obtain

$$\mathsf{AC}_F(a,b) = -\sum_{u\in\mathbb{F}_{2^n}}(-1)^{\mathrm{Tr}(au+(ub^{-1/K_r})^{2^r+1})} = -\sum_{u\in\mathbb{F}_{2^n}}(-1)^{\mathrm{Tr}(ab^{1/K_r}u+u^{2^r+1})}\,.$$

Observe that $\gcd(r,n) = 1$, so the Gold function $x \mapsto x^{2^r+1}$ is almost bent and

$$\Delta_F = 2^{\frac{n+1}{2}}\,.$$

□

Recall that the inverse of K_r in $\mathbb{Z}_{2^n-1}$ that appears in the proposition was precisely determined in Chapter 3 of this thesis. Note further that the cases $3r \equiv 1 \pmod n$ and $3r \equiv -1 \pmod n$ are essentially only one case because the r-th and $(n-r)$-th Kasami exponents are cyclotomic equivalent and the autocorrelation spectrum is invariant under affine equivalence.

The Bracken-Leander function is also a cubic permutation with differential uniformity 4. In the following, we determine the autocorrelation spectrum and the absolute indicator of the Bracken-Leander function. We need a well-known proposition to determine the dimensions of the kernel and image of a linear function.

Proposition 5.2.7 ([135, Proposition 4.4.]). *Let q be a prime power and $L \in \mathbb{F}_{q^n}[x]$ be an $\mathbb{F}_q$-linear function defined by $L(x) = \sum_{i=0}^{n-1} c_i x^{q^i}$. Then* $\dim \operatorname{im} L = \operatorname{rk} D$, *where*

$$D = \begin{pmatrix} c_0 & c_1 & \dots & c_{n-1}\\ c_{n-1}^q & c_0^q & \dots & c_{n-2}^q\\ \vdots & \vdots & & \vdots\\ c_1^{q^{n-1}} & c_2^{q^{n-1}} & \dots & c_0^{q^{n-1}}\end{pmatrix}.$$

Here, the dimension refers to the dimension of $\operatorname{im} L$ *as a vector space over $\mathbb{F}_q$. We call D the* Dickson matrix *of L.*

Theorem 5.2.8. *Let $F(x) = x^{q^2+q+1} \in \mathbb{F}_{q^4}[x]$, where $q = 2^k$ is the Bracken-Leander function. Then for any nonzero $a, b \in \mathbb{F}_{q^4}$,*

$$\mathsf{AC}_F(a,b) \in \left\{-q^3, 0, q^3\right\}$$

and $\Delta_F = q^3$.

Proof. Since F is a monomial, it is enough to consider the values of $\mathsf{AC}_F(1,b)$ by Proposition 5.2.1.

Since F is cubic, we can apply Proposition 5.1.15. We determine

$$t(1,b) = \dim\left\{w \in \mathbb{F}_{2^n} \mid D_w D_1 F_b = c\right\},$$

where $c \in \mathbb{F}_2$ is a constant. We have

$$\begin{aligned}
D_w D_1 F_b(x) =& D_w(\mathrm{Tr}(b(x^{q^2+q+1} + (x+1)^{q^2+q+1}))) \\
=& D_w\left(\mathrm{Tr}\left(b\left(x^{q^2+q} + x^{q^2+1} + x^{q+1} + x^{q^2} + x^q + x + 1\right)\right)\right) \\
=& D_w\left(\mathrm{Tr}\left(bx^{q^2+1} + \left(b^{q^3} + b\right)x^{q+1} + \left(b^{q^3} + b^{q^2} + b\right)x\right) + 1\right) \\
=& \mathrm{Tr}(b(x+w)^{q^2+1} + \left(b^{q^3} + b\right)(x+w)^{q+1} + \left(b^{q^3} + b^{q^2} + b\right)(x+w) \\
&+ by^{q^2+1} + \left(b^{q^3} + b\right)x^{q+1} + \left(b^{q^3} + b^{q^2} + b\right)x) \\
=& \mathrm{Tr}(b\left(w^{q^2+1} + wx^{q^2} + w^{q^2}x\right) + \left(b^{q^3} + b\right)\left(w^{q+1} + wx^q + w^q x\right) \\
&+ \left(b^{q^3} + b^{q^2} + b\right)w) \\
=& \mathrm{Tr}\left(bw^{q^2+1} + \left(b^{q^3} + b\right)w^{q+1} + \left(b^{q^3} + b^{q^2} + b\right)w\right) \\
&+ \mathrm{Tr}(x\left(\left(b^{q^3} + b^{q^2}\right)w^{q^3} + \left(b^{q^2} + b\right)w^{q^2} + \left(b^{q^3} + b\right)w^q\right)).
\end{aligned}$$

Then $t(1,b) = \dim\ker(L_b)$ where

$$L_b(w) = \left(b^{q^3} + b^{q^2}\right)w^{q^3} + \left(b^{q^2} + b\right)w^{q^2} + \left(b^{q^3} + b\right)w^q.$$

The dimension here refers to the dimension of the kernel over $\mathbb{F}_2$. Note that L_b is even an $\mathbb{F}_q$-linear function. The Dickson matrix of $L_b \in \mathbb{F}_q[x]$ is

$$D = \begin{pmatrix}
0 & b^{q^3} + b & b^{q^2} + b & b^{q^3} + b^{q^2} \\
b^{q^3} + b & 0 & b^q + b & b^{q^3} + b^q \\
b^{q^2} + b & b^q + b & 0 & b^{q^2} + b^q \\
b^{q^3} + b^{q^2} & b^{q^3} + b^q & b^{q^2} + b^q & 0
\end{pmatrix}.$$

It is easy to compute that the rank of D is 2 using a computer (or by hand with frequent use of the fact that $b^{q^4} = b$), so the kernel of L_b as an $\mathbb{F}_q$-vector space has dimension 2 by Proposition 5.2.7, which implies $t(1,b) = 2q$. Using Proposition 5.1.15, we conclude

$$\mathsf{AC}_F(a,b) \in \left\{-q^3, 0, q^3\right\}.$$

Since F is of course not bent, $\Delta_F = q^3$ by Proposition 5.1.8.

□

5.2.2 Quadratic functions and their inverses

In this subsection, we firstly consider the general quadratic (n,n)-functions and determine the autocorrelation spectra of the Gold functions and of their inverses. Determining the autocorrelation of a quadratic function is particularly easy as the derivative of a quadratic function is affine.

Theorem 5.2.9. *Let $F(x) = \sum_{0\le i<j\le n-1} c_{ij}x^{2^i+2^j} \in \mathbb{F}_{2^n}[x]$. Then the autocorrelation of F takes values from $\{0, \pm 2^n\}$ and $\Delta_F = 2^n$.*

Proof. For any $a, b \in \mathbb{F}_{2^n}^*$,

$$\begin{aligned}\mathsf{AC}_F(a,b) &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(F(x)+F(x+a)))} \\ &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}\left(b\left(\sum_{0\le i<j\le n-1} c_{ij}\left(a^{2^j}x^{2^i}+a^{2^i}x^{2^j}+a^{2^i+2^j}\right)\right)\right)} \\ &= (-1)^{\mathrm{Tr}\left(b\left(\sum_{0\le i<j\le n-1} c_{ij}a^{2^i+2^j}\right)\right)} \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(L(a,b)x)},\end{aligned}$$

where $L(a,b) = \sum_{0\le i<j\le n-1}\left(c_{ij}^{2^{-i}}a^{2^{j-i}}b^{2^{-i}} + c_{ij}^{2^{-j}}a^{2^{i-j}}b^{2^{-j}}\right)$. When $L(a,b) = 0$, we have $\mathsf{AC}_F(a,b) = \pm 2^n$ and otherwise $\mathsf{AC}_F(a,b) = 0$. Thus $\mathsf{AC}_F(a,b) \in \{-2^n, 0, 2^n\}$. Moreover, since F cannot be bent, we obtain $\Delta_F \neq 0$ and thus $\Delta_F = 2^n$. □

Corollary 5.2.10. *Let $F(x) = x^{2^i+1} \in \mathbb{F}_{2^n}[x]$ be a Gold function. Assume $k = \gcd(i,n)$ and $n' = n/k$. Then we get for $a, b \in \mathbb{F}_{2^n}^*$*

$$\mathsf{AC}_F(a,b) \in \begin{cases}\{0, 2^n\}, & \text{if } n' \text{ is even},\\ \{-2^n, 0\}, & \text{if } n' \text{ is odd and } k = 1,\\ \{-2^n, 0, 2^n\}, & \text{otherwise}.\end{cases}$$

Proof. Since F is a monomial, we can assume $a = 1$ without loss of generality by Proposition 5.2.1. From the proof of Theorem 5.2.9, it is clear that

$$\mathsf{AC}_F(1,b) = (-1)^{\mathrm{Tr}(b)} \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(L(b)x)},$$

where $L(b) = b^{2^{-i}} + b$. Thus $\ker(L) = \mathbb{F}_{2^{\gcd(i,n)}} = \mathbb{F}_{2^k}$. Furthermore, for any $b \in \mathbb{F}_{2^k}$, $\mathrm{Tr}_n(b) = n'\,\mathrm{Tr}_k(b)$. Therefore,

$$\mathsf{AC}_F(1,b) = \begin{cases}0, & \text{if } b \in \mathbb{F}_{2^n}\backslash\mathbb{F}_{2^k},\\ 2^n \times (-1)^{n'\,\mathrm{Tr}_k(b)}, & \text{if } b \in \mathbb{F}_{2^k}.\end{cases}$$

It follows that

$$\mathsf{AC}_F(1,b) \in \begin{cases}\{0, 2^n\}, & \text{if } n' \text{ is even},\\ \{-2^n, 0\}, & \text{if } n' \text{ is odd and } k = 1,\\ \{-2^n, 0, 2^n\}, & \text{otherwise}.\end{cases}$$

□

As previously observed, the autocorrelation spectrum and the absolute indicator are not invariant under compositional inversion. Now we consider the absolute

indicator of the inverse of a quadratic permutation. Unlike the quadratic exponents, the absolute indicator depends in this case on the considered function, as this simple example shows.

Example 5.2.11. For $n = 9$, the inverses of the two APN Gold permutations x^3 and x^5, namely x^{341} and x^{409}, do not have the same absolute indicator: the absolute indicator of x^{341} is 56 while the absolute indicator of x^{409} is 72.

A specificity of quadratic APN permutations for n odd is that they are *crooked*, which means that the image set of every derivative $D_aF, a \neq 0$, is the complement of a hyperplane $\langle \pi(a) \rangle^\perp = \{b\colon \mathrm{Tr}(b\pi(a)) = 0\}$. Moreover, it is known (see e.g. [22, Proof of Lemma 5]) that all these hyperplanes are distinct, which implies that π is a permutation of $\mathbb{F}_{2^n}$ when we add to the definition that $\pi(0) = 0$. Then, the following proposition shows that, for any quadratic APN permutation F, the autocorrelation of F^{-1} corresponds to the Walsh transform of π.

Proposition 5.2.12. *Let n be an odd integer and F be a quadratic APN permutation over $\mathbb{F}_{2^n}$. Let further π be the permutation of $\mathbb{F}_{2^n}$ defined by*

$$\mathrm{im}(D_aF) = \mathbb{F}_{2^n} \backslash \langle \pi(a) \rangle^\perp, \ \text{when} \ \ a \neq 0,$$

and $\pi(0) = 0$. Then for any nonzero a, b in $\mathbb{F}_{2^n}$, we have

$$\mathsf{AC}_{F^{-1}}(a, b) = -W_\pi(b, a).$$

It follows that

$$\Delta_{F^{-1}} \geq 2^{\frac{n+1}{2}}$$

with equality if and only if π is an AB permutation.

Proof. Let a, b be two nonzero elements of $\mathbb{F}_{2^n}$. Then, from Eq. (5.5), we deduce

$$\begin{aligned} \mathsf{AC}_{F^{-1}}(a, b) &= \sum_{\omega \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b\omega)} \delta_{F^{-1}}(a, \omega) \\ &= \sum_{\omega \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b\omega)} \delta_F(\omega, a). \end{aligned}$$

By the definition of π, we have that, for any nonzero a,

$$\delta_F(a, b) = \begin{cases} 2, & \text{if} \ \ \mathrm{Tr}(b\pi(a)) = 1, \\ 0, & \text{if} \ \ \mathrm{Tr}(b\pi(a)) = 0. \end{cases}$$

It then follows that

$$\delta_F(a, b) = 1 - (-1)^{\mathrm{Tr}(\pi(a)b)},$$

where this equality holds for all $(a, b) \neq (0, 0)$ by using that $\pi(0) = 0$. Therefore, we have, for any nonzero a and b,

$$\mathsf{AC}_{F^{-1}}(a, b) = \sum_{\omega \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b\omega)} \left(1 - (-1)^{\mathrm{Tr}(\pi(\omega)a)}\right) = -W_\pi(b, a).$$

As a consequence, $\Delta_{F^{-1}}$ is equal to the linearity of π, which is at least $2^{\frac{n+1}{2}}$ by the Sidelnikov-Chabaud-Vaudenay bound with equality if π is almost bent. □

It is worth noticing that the previous proposition is valid not only for quadratic APN permutations, but for all crooked permutations, which are a particular case

of AB functions. However, the existence of crooked permutations of degree strictly higher than 2 is an open question.

As a corollary of the previous proposition, we get some more precise information on the autocorrelation spectrum of the inverses of the Gold functions.

Corollary 5.2.13. *Let $n > 5$ be an odd integer and $0 < i < n$ with $\gcd(i, n) = 1$. Let F be the Gold function over $\mathbb{F}_{2^n}$ defined by $F(x) = x^{2^i+1}$. Then, for any nonzero a and b in $\mathbb{F}_{2^n}$, we have*

$$\mathsf{AC}_{F^{-1}}(a, b) = -W_\pi(b, a), \text{ where } \pi(x) = x^{2^n-2^i-2}.$$

Most notably, the absolute indicator of F^{-1} is strictly higher than $2^{\frac{n+1}{2}}$.

Proof. The result follows from Proposition 5.2.12. We determine the function π. For any $a \in \mathbb{F}_{2^n}^*$ we compute the number $\delta_F(a, b)$ of solutions of the equation

$$(x + a)^{2^i+1} + x^{2^i+1} = b.$$

By expanding, we get

$$ax^{2^i} + a^{2^i}x = a^{2^i+1} + b.$$

The number of solutions of this equation is equal to the number of solutions of the equation

$$x^{2^i} + x = 1 + ba^{-(2^i+1)},$$

which follows from dividing the equation by a^{2^i+1} and then substituting $x \mapsto ax$. Clearly, this equation has a solution if and only if $\mathrm{Tr}\left(ba^{-(2^i+1)}\right) = 1$. Accordingly,

$$\langle \pi(a) \rangle^\perp = \{b \colon \mathrm{Tr}(ba^{2^n-2^i-2}) = 0\}.$$

It follows that

$$\pi(x) = x^{2^n-2^i-2}.$$

The autocorrelation of F^{-1} then follows from Proposition 5.2.12.

AB functions have algebraic degree at most $\frac{n+1}{2}$ (Proposition 2.2.13), while π has algebraic degree $(n-2)$. It follows that π cannot be AB when $n > 5$. Therefore, the absolute indicator of F^{-1} is strictly higher than $2^{\frac{n+1}{2}}$. □

5.3 Outlook

In this chapter, we investigated the differential-linear connectivity table (DLCT) of vectorial Boolean functions by clarifing its connection to the autocorrelation of vectorial Boolean functions.

This chapter only covers a small portion of interesting problems on this subject and many problems deserve further research:

Problem 5.3.1. For an odd integer n, are there (n, n)-power functions F with $\Delta_F = 2^{(n+1)/2}$ other than the Kasami APN functions?

The generic lower bound on the absolute indicator of vectorial Boolean functions derived in this chapter is lower than what experimental results suggest and thus might be further improved. A natural follow-up topic would be the investigation and construction of optimal, or near-optimal, vectorial Boolean functions with respect to the bounds.

Problem 5.3.2. Determine a (tight) lower bound on the absolute indicator of vectorial Boolean functions. Are there constructions exhibiting (near) optimal vectorial Boolean functions with respect to that bound?

From Corollary 5.1.17 it follows that an APN function with very low absolute indicator is of interest.

Problem 5.3.3. Is there an APN function in 9 variables with absolute indicator $\Delta = 24$?

In addition, the absolute indicators of the Kasami and Welch functions have not been determined completely.

Problem 5.3.4. Determine the absolute indicators of the Kasami and Welch functions completely.

Chapter 6

XOR-counts and lightweight multiplication in binary finite fields

This chapter is based on the paper [84] written by the author of this thesis and published in the proceedings of the EUROCRYPT conference in 2019.

6.1 Introduction

In the past years, with the advent of the so called *Internet of Things*, new challenges for cryptography have emerged. Many new devices usually do not have a lot of computational power and memory, but are still required to offer some security by encrypting sensitive data. Consequentially, *lightweight cryptography* has become a major field of research in the past years, mostly focusing on symmetric-key encryption. Classic examples of such ciphers are PRESENT [13] or GIFT [3]. For a great survey of the challenges and ideas in lightweight symmetric cryptography, including an overview of lightweight ciphers up until 2017, we refer to [9]. For some very recent proposals, we refer to the ongoing NIST competition on lightweight cryptography [1]. In particular, linear layers (e.g. [101, 104]) and S-boxes (e.g. [21, 119]) have been thoroughly investigated as they constitute key components of SPNs. The main objective here is to try to minimize the cost of storage and the number of operations needed to apply a cryptographic function. Usually, the security properties of cryptographic schemes using finite fields do not depend on a specific field representation (as bit strings) in the actual implementation [42], so the choice of field implementation makes an impact on the performance of the scheme without influencing its security. It is therefore an interesting question which representation minimizes the number of operations needed.

In practice, linear layers are usually $\mathbb{F}_{2^m}$-linear mappings on $\mathbb{F}_{2^m}^n$. Recall that linear mappings are implemented as matrix multiplications. Note that we can write every $n \times n$ matrix over $\mathbb{F}_{2^m}$ as an $(mn) \times (mn)$ matrix over $\mathbb{F}_2$. As elements in $\mathbb{F}_{2^m}$ are usually represented as bit strings in computers, it is natural to consider only matrices over $\mathbb{F}_2$. Measurements of implementation costs will then only involve the number of bit-operations (XORs) needed. It is an interesting question to evaluate the efficiency of a given matrix. For that purpose two different metrics have been introduced, the *direct XOR-count* (e.g. in [79, 101, 121, 125]) and the *sequential XOR-count* (e.g. [5, 72, 140]). Roughly speaking, the direct XOR-count counts the number of non-zeros in the matrix, whereas the sequential XOR-count counts the number of

[1] At the time of the writing of this thesis, the competition is in its second round. More details as well as all candidates can be found at https://csrc.nist.gov/projects/lightweight-cryptography

elementary row operations needed to transform the matrix into the identity matrix (see Section 6.2 for more precise definitions). Although the sequential XOR-count of a matrix is harder to compute, it often yields a better estimation of the actual optimal number of XOR-operations needed [72], for a simple example see Example 6.2.2 in this chapter. When implementing a linear layer, a field representation can be chosen such that the respective matrix is optimal according to these metrics. In this way, the performance of a given linear layer can be improved (for example by choosing a field representation that results in a sparse diffusion matrix).

Our goal in this chapter is to explore some connections and properties of the direct and sequential XOR-count metrics and then to apply these to get some theoretical results regarding optimal implementations of matrices that represent multiplication with a fixed field element $\alpha \in \mathbb{F}_{2^k}$. Optimal choices of these matrices (called *multiplication matrices*) can then be used for local optimizations of matrices over $\mathbb{F}_{2^k}$ (this approach was taken for example in [5, 72, 101, 104, 121]). For instance, the diffusion matrix (i.e. the linear layer) of AES is

$$\begin{pmatrix} \alpha & \alpha+1 & 1 & 1 \\ 1 & \alpha & \alpha+1 & 1 \\ 1 & 1 & \alpha & \alpha+1 \\ \alpha+1 & 1 & 1 & \alpha \end{pmatrix} \in \mathrm{GL}(4, \mathbb{F}_{2^8}),$$

where α is a root of the irreducible polynomial $x^8 + x^4 + x^3 + x + 1 \in \mathbb{F}_2[x]$. From this example, it is evident that optimizing the implementations of mappings $x \mapsto \alpha x$ is highly relevant for the efficiency of matrix-vector products with diffusion matrices of this form. Such an optimization is called *local optimization*.

Recently, the focus has shifted to global optimization, as it has become clear that local optimizations are not necessarily also globally optimal [52, 86]. However, global optimization techniques currently rely either on tools that improve the XOR-counts of matrices already known to be efficient [86] or exhaustive searches [52, 120]. In particular, theoretical results on globally optimal matrices seem to be very hard to obtain. Numerical data suggest that there is a correlation between good local optimizations and good global optimizations (see [86, Figures 2-6]). Because of this correlation, theoretical insights into local optimization are valuable for the search of globally optimal matrices.

In the second section, we compare the direct XOR-count and sequential XOR-count evaluation metrics. We prove some theoretical properties of the sequential XOR-count that can be used to improve algorithms (e.g. an algorithm presented in [5]). We also find an infinite family of matrices that have a lower direct XOR-count than sequential XOR-count, disproving a conjecture in [72]. We want to emphasize that the results presented in this section apply to all invertible matrices, not just multiplication matrices.

In the third section we provide a complete characterisation of finite field elements α where the mapping $x \mapsto \alpha x$ can be implemented with exactly 2 XOR-operations (Theorem 6.3.7), which proves a conjecture in [5]. This case is of special interest, since for many finite fields (including the fields $\mathbb{F}_{2^n}$ with $8|n$ that are particularly interesting for many applications) there are no elements for which the mapping $x \mapsto \alpha x$ can be implemented with only 1 XOR-operation [5]. For these fields, our classification gives a complete list of elements α such that multiplication with α can be implemented in the cheapest way possible.

In the fourth section we present some more general results for multiplication

matrices with higher XOR-counts. We prove that the number of XOR-operations needed to implement the mapping $x \mapsto \alpha x$ depends on the number of non-zero coefficients of the minimal polynomial of α. In particular, Theorem 6.4.1 shows that the gap between the number of XORs used in an optimal implementation and the number of XORs used in the "naive" implementation of a multiplication matrix using the rational canonical form of the mapping $x \mapsto \alpha x$ grows exponentially with the weight of the minimal polynomial of the element. This result shows that there is a large potential for improvement in the implementation of multiplication matrices. Propositions 6.4.2 and 6.4.3 imply that the bound found in Theorem 6.4.1 is optimal.

We conclude this chapter with several open problems.

6.2 XOR-Counts

An XOR-count metric for diffusion matrices was introduced in [79] and then generalized for arbitrary matrices in [125]. It has then subsequently been studied in several works, e.g. [121, 101].

Definition 6.2.1. *The direct XOR-count (d-XOR-count) of an invertible $n \times n$ matrix M over $\mathbb{F}_2$, denoted by $\mathrm{wt}_d(M)$ is*

$$\mathrm{wt}_d(M) = \omega(M) - n,$$

where $\omega(M)$ denotes the number of ones in the matrix M.

Note that the d-XOR-count of an invertible matrix is never negative as every row of an invertible matrix needs to have at least one non-zero entry. Moreover, $\mathrm{wt}_d(M) = 0$ if and only if M has exactly one '1' in every row and column, i.e. M is a permutation matrix. The d-XOR-metric only gives an upper bound to the actual minimal implementation cost as the following example shows.

Example 6.2.2.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_1 + a_2 \\ (a_1 + a_2) + a_3 \\ ((a_1 + a_2) + a_3) + a_4 \end{pmatrix}$$

The d-XOR-count of the matrix is 6 but it is easy to see that multiplication with this matrix can actually be implemented with only 3 XOR operations since the results of previous steps can be reused. A metric that allows this was subsequently introduced in [72] and used in further work (e.g. [5, 52, 140]). Let us introduce some notation at first: We denote by I the identity matrix and by $E_{i,j}$ the matrix that has exactly one '1' in the i-th row and j-th column. Then $A_{i,j} := I + E_{i,j}$ for $i \neq j$ is called an *addition matrix*. Left-multiplication with $A_{i,j}$ adds the j-th row to the i-th row of a matrix, right-multiplication adds the i-th column to the j-th column. Observe that the matrices $A_{i,j}$ are self-inverse over $\mathbb{F}_2$. Let further $\mathcal{P}(n)$ be the set of $n \times n$ permutation matrices and $\mathcal{A}(n)$ the set of all $n \times n$ addition matrices $A_{i,j}$. We will omit the dimension n unless necessary.

Definition 6.2.3. *An invertible matrix M over $\mathbb{F}_2$ has a sequential XOR-count (s-XOR-count) of t if t is the minimal number such that M can be written as*

$$M = P \prod_{k=1}^{t} A_{i_k,j_k}$$

where $P \in \mathcal{P}$ and $A_{i_k,j_k} \in \mathcal{A}$. We write $\mathrm{wt}_s(M) = t$.

Note that every invertible matrix can be decomposed as a product of a permutation matrix and addition matrices in the way Definition 6.2.3 describes. Indeed, Gauss-Jordan-elimination gives a simple algorithm to do so.

In [140] a similar definition for the s-XOR-count was given that uses a representation of the form $M = \prod_{k=1}^{t} P_k A_{i_k,j_k}$ with permutation matrices P_k. Since products of permutation matrices remain permutation matrices and

$$PA_{i,j} = A_{\sigma^{-1}(i),\sigma^{-1}(j)}P \tag{6.1}$$

where $\sigma \in S_n$ is the permutation belonging to the permutation matrix P, this definition is equivalent to our definition.

A representation of a matrix M as a product $M = P\prod_{k=1}^{t} A_{i_k,j_k}$ is called an *s-XOR-representation* of M and an s-XOR-representation with $\mathrm{wt}_s(M)$ addition matrices is called an *optimal s-XOR-representation*. Note that optimal s-XOR-representations are generally not unique. Observe that $M = PA_{i_1,j_1} \dots A_{i_t,j_t}$ is equivalent to $MA_{i_t,j_t} \dots A_{i_1,j_1} = P$, so the s-XOR-count measures the number of column addition steps that are needed to transform a matrix into a permutation matrix. Because of equation (6.1) the number of column additions needed is equal to the number of row additions needed, so we may also speak about row additions.

Going back to Example 6.2.2, it is easy to find an s-XOR-representation with 3 XORs.

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} = IA_{4,3}A_{3,2}A_{2,1}.$$

It is clear that we need at least 3 addition matrices since all rows but the first one need at least one update. Hence, the s-XOR-representation above is optimal and $\mathrm{wt}_s(M) = 3$.

Determining the s-XOR-count of a given matrix is generally not easy. Graph-based algorithms to find an optimal s-XOR-count have been proposed in [140] and (in a slightly different form) in [72]. The algorithms are based on the following observation. Let $G = (V, E)$ be a graph where $G = \mathrm{GL}(n, \mathbb{F}_2)$ and $(M_1, M_2) \in E$ if $AM_1 = M_2$ for an $A \in \mathcal{A}$. Then $\mathrm{wt}_s(M) = \min_{P\in\mathcal{P}} d(M, P)$, where $d(M_1, M_2)$ denotes the distance between M_1 and M_2 in the graph G. Thus, the evaluation of the s-XOR-count can be reduced to a shortest-path-problem. Note that because the elementary matrices in $\mathcal{A}$ are all involutory, G is undirected. As the authors of [140] observe, it is possible to reduce the number of vertices by a factor $1/n!$ because matrices with permuted rows can be considered equivalent. Still $(1/n!)|\mathrm{GL}(n, \mathbb{F}_2)| = (1/n!)(2^n - 1)(2^n - 2)\dots(2^n - 2^{n-1})$ and every vertex has $|\mathcal{A}(n)| = n^2 - n$ neighbors, so both the number of vertices and the number of edges grow exponentially. Hence, this approach is impractical unless n is small.

The problem of determining the s-XOR-count is linked with the problem of optimal pivoting in Gauss-Jordan elimination since the number of additions in an optimal elimination process is clearly an upper bound of the s-XOR-count. Pivoting

strategies for Gaussian elimination are a classical problem in numerical linear algebra (among lots of examples, see [90]) and the number of steps needed in a Gauss-Jordan elimination process can be used as a heuristic for the s-XOR-count.

Example 6.2.2 gives an example of a matrix with lower s-XOR-count than d-XOR-count. Considering this and the fact that the s-XOR-count of a given matrix is generally much harder to determine than the d-XOR-count, it should be clarified whether the s-XOR-count always gives a better estimation of the actual number of XOR operations needed to implement the matrix. In [72] this has been conjectured, i.e. $\mathrm{wt}_s(M) \leq \mathrm{wt}_d(M)$ for all $M \in \mathrm{GL}(n, \mathbb{F}_2)$. However, the following theorem gives a counterexample.

Theorem 6.2.4. *Let M be as follows:*

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{GL}(7, \mathbb{F}_2).$$

Then $\mathrm{wt}_s(M) > \mathrm{wt}_d(M)$.

Proof. M is invertible with $\mathrm{wt}_d(M) = 8$. Let $\mathrm{wt}_s(M) = t$, i.e. there are matrices $A_{i_k,j_k} \in \mathcal{A}$ and $P \in \mathcal{P}$ such that $\prod_{k=1}^{t} A_{i_k,j_k} \cdot M = P$. By construction, no two rows and no three rows of M add up to a row with only one non-zero entry. Every row has to be updated at least once to transform M into a permutation matrix. Since no two row vectors add up to a vector with only one non-zero entry, the first row that gets updated (row i_t) needs to get updated at least once more. But as there is also no combination of three vectors adding up to a vector with only one non-zero entry, the second row that is updated (row i_{t-1}) also needs to be updated a second time. So two rows need to get updated at least twice, and all other 5 rows need to get updated at least once, resulting in $\mathrm{wt}_s(M) \geq 9$. □

Remark 6.2.5. Note that the structure of the counterexample can be extended to all dimensions $n \geq 7$, the middle '1' in the last row can be in any j-th column with $4 \leq j \leq n-3$. We conclude that there exists a matrix $M \in \mathrm{GL}(n, \mathbb{F}_2)$ with $\mathrm{wt}_s(M) > \mathrm{wt}_d(M)$ for all $n \geq 7$.

In a subsequent work [32] a matrix with higher s-XOR-count than d-XOR-count was also found in dimension 6. It was also proven that this phenomenon does not occur in dimensions $n < 6$.

Studying the s-XOR-count is an interesting mathematical problem because it has some properties that can be used to get upper bounds of the actual implementation cost of potentially a lot of matrices. The actual number of XOR-operations needed is clearly invariant under permutation of rows and columns. It is therefore desirable that this property is reflected in our XOR-metrics. Obviously, this is the case for the d-XOR-count, i.e. $\mathrm{wt}_d(M) = \mathrm{wt}_d(PMQ)$ for all matrices M and permutation matrices $P, Q \in \mathcal{P}$. The following lemma shows that this also holds for the s-XOR-count. The lemma is a slight modification of a result in [5]. However the proof in [5] has a small gap, so we provide a complete proof here. We denote permutation-similarity with $\sim$, i.e. $M_1 \sim M_2$ if there exists a $P \in \mathcal{P}$ so that $M_1 = PM_2P^{-1}$.

Lemma 6.2.6. *Let $M \in GL(n, \mathbb{F}_2)$. Then $\mathrm{wt}_s(M) = \mathrm{wt}_s(PMQ)$ for $P, Q \in \mathcal{P}$. In particular, if $M_1 \sim M_2$ then $\mathrm{wt}_s(M_1) = \mathrm{wt}_s(M_2)$.*

Proof. Let $\mathrm{wt}_s(M) = t$ and $\sigma \in S_n$ be the permutation belonging to Q. Then, by shifting Q to the left

$$PMQ = PP_2 \prod_{k=1}^{t} A_{i_k,j_k} Q = PP_2 Q \prod_{k=1}^{t} A_{\sigma(i_k),\sigma(j_k)} = P' \prod_{k=1}^{t} A_{\sigma(i_k),\sigma(j_k)}$$

where $P_2, P' \in \mathcal{P}$, so $\mathrm{wt}_s(PMQ) \leq \mathrm{wt}_s(M)$. Since $M = P^{-1}(PMQ)Q^{-1}$ the same argument yields $\mathrm{wt}_s(M) \leq \mathrm{wt}_s(PMQ)$. □

Based on this result, the following normal form for permutation matrices is proposed in [5]. We introduce a notation for block diagonal matrices. Let $M_1, \ldots, M_d$ be square matrices, then we denote the block matrix consisting of these matrices by

$$\bigoplus_{k=1}^{d} M_k := \begin{pmatrix} M_1 & & & 0 \\ & M_2 & & \\ & & \ddots & \\ 0 & & & M_d \end{pmatrix}.$$

We denote by C_p the companion matrix of a polynomial $p = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{F}_2[x]$, i.e.

$$C_p = \begin{pmatrix} 0 & \ldots & 0 & a_0 \\ 1 & 0 & 0 & a_1 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \ldots & 1 & a_{n-1} \end{pmatrix}.$$

Lemma 6.2.7 ([5, Lemma 2])**.** *Let $P \in \mathcal{P}(n)$. Then*

$$P \sim \bigoplus_{k=1}^{d} C_{x^{m_k}+1}$$

for some m_k with $\sum_{k=1}^{d} m_k = n$ and $m_1 \geq \cdots \geq m_d \geq 1$.

A permutation matrix of this structure is said to be the *cycle normal form* of P. We can then (up to permutation-similarity) always assume that the permutation matrix of the s-XOR-decomposition is in cycle normal form.

Corollary 6.2.8 ([5, Corollary 2])**.**

$$P \prod_{k=1}^{t} A_{i_k,j_k} \sim P' \prod_{k=1}^{t} A_{\sigma(i_k),\sigma(j_k)}$$

for some permutation $\sigma \in S_n$, where P' is the cycle normal form of P.

We say an s-XOR-representation is in cycle normal form if its permutation matrix is in cycle normal form. Corollary 6.2.8 states that every s-XOR-representation is pemutation-similar to exactly one s-XOR-representation in cycle normal form.

The following theorem gives a connection between the s-XOR-count and optimal s-XOR-representations of a given matrix and that of its inverse.

Theorem 6.2.9. *Let M be an invertible matrix with $\mathrm{wt}_s(M) = t$ and*

$$M = P\prod_{k=1}^{t} A_{i_k,j_k} \text{ with } P = \bigoplus_{k=1}^{d} C_{x^{m_k}+1}.$$

Then $\mathrm{wt}_s(M^{-1}) = t$. Moreover,

$$M^{-1} = PA_{\sigma(i_t),\sigma(j_t)}A_{\sigma(i_{t-1}),\sigma(j_{t-1})}\dots A_{\sigma(i_1),\sigma(j_1)}$$

for some permutation $\sigma \in S_n$ that depends only on P.

Proof. For the inverse matrix we have

$$M^{-1} = A_{i_t,j_t}\dots A_{i_1,j_1}P^{-1} \sim P^{-1}A_{i_t,j_t}\dots A_{i_1,j_1},$$

so $\mathrm{wt}_s(M^{-1}) \leq \mathrm{wt}_s(M)$. By symmetry, we get $\mathrm{wt}_s(M^{-1}) = \mathrm{wt}_s(M)$. Observe that $P^{-1} = P^T = \bigoplus_{k=1}^{d} C^T_{x^{m_k}+1}$ where P^T denotes the transpose of P. Let J_r be the $r \times r$ matrix with ones on the counterdiagonal, i.e. $J_{i,j} = 1$ if and only if $j = n - i + 1$. Let $Q = \bigoplus_{k=1}^{d} J_{m_k} \in \mathcal{P}$. A direct calculation yields $QP^{-1}Q^{-1} = P$ and thus

$$M^{-1} \sim QP^{-1}\prod_{k=t}^{1} A_{i_k,j_k}Q^{-1} = P\prod_{k=t}^{1} A_{\sigma(i_k),\sigma(j_k)},$$

where $\sigma \in S_n$ denotes the permutation that belongs to Q. □

In particular, Theorem 6.2.9 implies that given an optimal s-XOR-representation for a matrix M, an optimal s-XOR-representation of M^{-1} can be determined with very little effort by calculation the permutation σ in the proof. Note that the statement of Theorem 6.2.9 does not exist for the d-XOR-count. Indeed, sparse matrices (i.e. matrices with low d-XOR-count) usually have dense inverse matrices (i.e. high d-XOR-count).

The next result also holds for the s-XOR-count only.

Proposition 6.2.10. *Let M, N be invertible matrices with $\mathrm{wt}_s(M) = t_1$ and $\mathrm{wt}_s(N) = t_2$. Then $\mathrm{wt}_s(MN) \leq t_1 + t_2$. In particular, $\mathrm{wt}_s(M^k) \leq |k|t_1$ for all $k \in \mathbb{Z}$.*

Proof. Let $M = P\prod_{k=1}^{t_1} A_{i_k,j_k}$ and $N = Q\prod_{k=1}^{t_2} B_{i_k,j_k}$. Then

$$MN = PQ\prod_{k=1}^{t_1} A_{\sigma(i_k),\sigma(j_k)}\prod_{k=1}^{t_2} B_{i_k,j_k},$$

where $\sigma \in S_n$ is the permutation belonging to Q. This implies $\mathrm{wt}_s(MN) \leq t_1 + t_2$. The statement $\mathrm{wt}_s(M^k) \leq |k|t_1$ for $k < 0$ follows from Theorem 6.2.9. □

6.3 Efficient Multiplication Matrices in Finite Fields

We can consider $\mathbb{F}_{2^n}$ as the n-dimensional vector space $(\mathbb{F}_2)^n$ over $\mathbb{F}_2$. By distributivity, the function $x \mapsto \alpha x$ for $\alpha \in \mathbb{F}_{2^n}$ is linear, so it can be represented as a (left-)multiplication with a matrix in $\mathrm{GL}(n, \mathbb{F}_2)$. This matrix obviously depends on α, but also on the choice of the basis of $(\mathbb{F}_2)^n$ over $\mathbb{F}_2$. We denote the multiplication matrix that represents the function $x \mapsto \alpha x$ with respect to the basis B by $M_{\alpha,B}$. The XOR-count of $M_{\alpha,B}$ generally differs from the XOR-count of $M_{\alpha,B'}$ for different bases B, B'.

Our objective here is to find the optimal basis B for a given α, in the sense that the XOR-count of $M_{\alpha,B}$ is minimized. For this, we define the XOR-count metrics from the previous section also for elements from $\mathbb{F}_{2^n}$.

Definition 6.3.1. *Let $\alpha \in \mathbb{F}_{2^n}$. We define the s-XOR-count and d-XOR-count of α as follows:*

$$\mathrm{wt}_s(\alpha) = \min_B \mathrm{wt}_s(M_{\alpha,B}), \quad \mathrm{wt}_d(\alpha) = \min_B \mathrm{wt}_d(M_{\alpha,B}),$$

where the minimum is taken over all bases of $\mathbb{F}_2^n$ over $\mathbb{F}_2$. A basis B and matrix $M_{\alpha,B}$ that satisfy the minimum are called s-XOR-optimal and d-XOR-optimal for α, respectively.

In order to find the matrices that optimize the s-XOR-count-metric, an exhaustive search on all matrices with low s-XOR-count is performed in [5]. In this way the s-XOR-count and an optimal s-XOR-matrix of every element $\alpha \in \mathbb{F}_{2^n}$ for $n \leq 8$ was found. Using the results presented in the previous section, the search was restricted to matrices where the permutation matrix is in cycle normal form. The following result was used to determine whether a given matrix is a multiplication matrix for some $\alpha \in \mathbb{F}_{2^n}$ with respect to some basis B. For the rest of this chapter, we denote by $\chi(M) = \det(xI + M)$ the characteristic polynomial of a matrix M and by m_α the minimal polynomial of the finite field element $\alpha \in \mathbb{F}_{2^n}$. Since we will not use any characters in this chapter, there will be no confusion. Recall that m_α is always irreducible.

Theorem 6.3.2 ([5, Theorem 1])**.** *Let $M \in GL(n, \mathbb{F}_2)$ and $\alpha \in \mathbb{F}_{2^n}$. Then M is a multiplication matrix for α, i.e. $M = M_{\alpha,B}$ with respect to some basis B, if and only if m_α is the minimal polynomial of M.*

Theorem 6.3.2 shows in particular that a matrix M is a multiplication for some $\alpha \in \mathbb{F}_{2^n}$ with respect to some basis B if and only if the minimal polynomial of M is irreducible. Additionally, it is clear that two field elements with the same minimal polynomial necessarily have the same XOR-counts.

Remark 6.3.3. A direct calculation of the minimal polynomial of the matrix M in Theorem 6.2.4 yields $m_M = x^7 + x^6 + x^5 + x^4 + 1$ which is an irreducible polynomial. According to Theorem 6.3.2 the matrix M is a multiplication matrix for an element $\alpha \in \mathbb{F}_{2^7}$ with respect to some basis. Hence, there are elements $\alpha \in \mathbb{F}_{2^n}$ such that $\mathrm{wt}_d(\alpha) < \mathrm{wt}_s(\alpha)$. Note that this case does not have to occur for every value of n because the matrices provided in Theorem 6.2.4 might have a reducible minimal polynomial. Indeed, an exhaustive search for the cases $n = 4$ and $n = 8$ was conducted in [72], resulting in $\mathrm{wt}_s(\alpha) \leq \mathrm{wt}_d(\alpha)$ for all α in $\mathbb{F}_{2^4}$ and $\mathbb{F}_{2^8}$, respectively. We tested the examples given in Theorem 6.2.4 for $n = 16$ without finding any matrices with irreducible minimal polynomial. Hence, we conjecture that $\mathrm{wt}_s(\alpha) \leq \mathrm{wt}_d(\alpha)$ for all $\alpha \in \mathbb{F}_{2^{16}}$. It is an interesting question for which n elements with lower d-XOR-count than s-XOR-count exist.

Corollary 6.3.4. *Let $M = P\prod_{k=1}^{t} A_{i_k,j_k}$ be in cycle normal form. Then M is a multiplication matrix for $\alpha \in \mathbb{F}_{2^n}$ if and only if M^{-1} is a multiplication matrix for $\alpha^{-1} \in \mathbb{F}_{2^n}$. Moreover, M is an optimal s-XOR-matrix for α if and only if M^{-1} is an optimal s-XOR-matrix for α^{-1}.*

Proof. Let p and q be the minimal polynomial of M and M^{-1}, respectively. It is well known that q is then the reciprocal polynomial of p, that is $q(x) = x^n p(1/x)$. Moreover, p is the minimal polynomial of α if and only if q is the minimal polynomial of α^{-1}. The rest follows from Theorem 6.2.9. □

Corollary 6.3.4 allows us to determine an s-XOR-optimal matrix for α^{-1} given an s-XOR-optimal matrix M of α. Recall that the cycle normal form of M^{-1} was directly computed in Theorem 6.2.9. This allows us to cut the search space (approximately) in half for all algorithms that determine the s-XOR-count by traversing all matrices in $\mathrm{GL}(n, \mathbb{F}_2)$.

It is now an interesting question which elements $\alpha \in \mathbb{F}_{2^n}$ have multiplication matrices with low XOR-count. Obviously, the only element that can be implemented with XOR-count 0 is $\alpha = 1$. A simple upper bound on the s-XOR-count and d-XOR-count for elements can be found by considering the rational canonical form of a matrix. Recall that a matrix $M \in \mathrm{GL}(n, \mathbb{F}_2)$ is similar to its (unique) rational canonical form. If M has an irreducible minimal polynomial m with $\deg m = k$ then there exists a $d \geq 1$ so that $kd = n$ and the rational canonical form is $\bigoplus_{i=1}^{d} C_m$. For a polynomial p we denote by $\mathrm{wt}(p)$ the weight of p, that is the number of non-zero coefficients. Note that if $2 | \mathrm{wt}(p)$ then 1 is a root of p so the only irreducible polynomial over $\mathbb{F}_2$ with even weight is $x + 1$.

Example 6.3.5. Let α be an element of $\mathbb{F}_{2^n}$ with minimal polynomial m_α and $\deg m_\alpha = k$ with $kd = n$ and $d \geq 1$. Then we can find a basis B so that $M_{\alpha,B}$ is in rational canonical form, i.e. $M_{\alpha,B} = \bigoplus_{i=1}^{d} C_{m_\alpha}$. It is easy to check that $\mathrm{wt}_s(M_{\alpha,B}) = \mathrm{wt}_d(M_{\alpha,B}) = d \cdot (\mathrm{wt}(m_\alpha) - 2)$.

This example shows in particular that all $\alpha \in \mathbb{F}_{2^n}$ with $\deg m_\alpha = n$ and $\mathrm{wt}(m_\alpha) = 3$ can be implemented with only one XOR operation. A possible basis for this case is the polynomial basis $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$.

As one row-addition on I only produces one extra '1' in the matrix, $\mathrm{wt}_d(M) = 1$ if and only if $\mathrm{wt}_s(M) = 1$, and equivalently, $\mathrm{wt}_d(\alpha) = 1$ if and only if $\mathrm{wt}_s(\alpha) = 1$. In [5] all elements that can be implemented with exactly one XOR-operation are characterized. It turns out, that these cases are exactly those covered by Example 6.3.5.

Theorem 6.3.6 ([5, Theorem 2])**.** *Let $\alpha \in \mathbb{F}_{2^n}$. Then $\mathrm{wt}_s(\alpha) = 1$ or $\mathrm{wt}_d(\alpha) = 1$ if and only if m_α is a trinomial of degree n.*

It is an open problem for which n irreducible trinomials of degree n exist. Among other sporadic examples, it is known that there are no irreducible trinomials of degree n if $n \equiv 0 \pmod 8$ [128], so there are no elements α with d/s-XOR-count 1 in these cases. As the case $8|n$ is especially important in practice, it is natural to consider elements that can be implemented with 2 XOR operations. In this case, s-XOR-count and d-XOR count do differ: By simply expanding the product $PA_{i_1,j_1}A_{i_2,j_2} = P(I + E_{i_1,j_1})(I + E_{i_2,j_2})$, it follows that every matrix with $\mathrm{wt}_s(M) = 2$ is of the following form:

$$M = \begin{cases} P + E_{\sigma^{-1}(i_1),j_1} + E_{\sigma^{-1}(i_2),j_2}, & i_2 \neq j_1 \\ P + E_{\sigma^{-1}(i_1),j_1} + E_{\sigma^{-1}(i_2),j_2} + E_{\sigma^{-1}(i_1),j_2}, & i_2 = j_1, \end{cases} \tag{6.2}$$

where σ is the permutation that belongs to P and $i_1 \neq j_1$, $i_2 \neq j_2$. In particular equation (6.2) shows that $\mathrm{wt}_d(M) = 2$ implies $\mathrm{wt}_s(M) = 2$, but there are some matrices with $\mathrm{wt}_s(M) = 2$ and $\mathrm{wt}_d(M) = 3$. In other words, the s-XOR-metric is a better metric for these matrices. In [5] the authors conjecture that $\mathrm{wt}_s(\alpha) = 2$ implies $\mathrm{wt}(m_\alpha) \leq 5$, i.e. the minimal polynomial is a trinomial or a pentanomial. We confirm this conjecture by giving an exact characterization of all elements with $\mathrm{wt}_s(\alpha) = 2$ and their optimal s-XOR-representation in cycle normal form in Theorem 6.3.7.

In the proof the following concept from linear algebra is used. We refer the reader to [67] for proofs and more background. Let V be a vector space over a field $\mathbb{F}$ with

m_α	optimal matrix representation	Case
$x^n + x^{k_1+k_2} + x^{k_1} + x^{k_2} + 1$, $k_1 + k_2 \leq n-2$	$C_{x^n+1} + E_{i_1,j_1} + E_{i_2,j_2}$	(1.3.)
$x^n + x^{n-k_1} + x^{k_2} + x^{k_2-k_1} + 1$, $k_2 > k_1$	$C_{x^n+1} + E_{i_1,j_1} + E_{i_2,j_2}$	(1.4.)
$x^n + x^{k_1+k_2} + x^{k_1} + x^{k_2} + 1$	$C_{x^n+1} + E_{i_1,j_1} + E_{j_1+1,j_2} + E_{i_1,j_2}$	(2.1.)
$x^n + x^{n_1} + x^{n_2} + x^k + 1$, $k \leq n-2$	$(C_{x^{n_1}+1} \oplus C_{x^{n_2}+1}) + E_{i_1,j_1} + E_{i_2,j_2}$	(3.2.)
$x^n + x^{n_1+k} + x^{n_2} + x^{n_1} + 1$, $0 < k < n_2$	$(C_{x^{n_1}+1} \oplus C_{x^{n_2}+1}) + E_{i_1,j_1}$ $+E_{j_1+1 \pmod{n_1},j_2} + E_{i_1,j_2}$	(4.)
$x^{n/2} + x^k + 1$	$(C_{x^{n/2}+1} \oplus C_{x^{n/2}+1}) + E_{i_1,j_1} + E_{i_2,j_2}$	(3.1.)

TABLE 6.1: Elements with minimal polynomials listed in the left column have s-XOR-count 2. The second column gives an optimal multiplication matrix and the third column points to the corresponding case in the proof.

dimension n, $u \in V$ a vector and M an $n \times n$-matrix over $\mathbb{F}$. The monic polynomial $g(x) \in \mathbb{F}[x]$ with the smallest degree such that $g(M)u = 0$ is called the *M-annihilator* of u. This polynomial divides any polynomial h annihilating u (i.e. $h(M)u = 0$), in particular the minimal polynomial of M. In the case that the minimal polynomial of M is irreducible the M-annihilator of every vector $u \neq 0$ is the minimal polynomial of M. So if we find a polynomial h that annihilates a vector $u \neq 0$ we know that the minimal polynomial divides h. In particular, if h is monic and the degree of h and the minimal polynomial coincide we can infer that h is the minimal polynomial of M.

Theorem 6.3.7. *Let $\alpha \in \mathbb{F}_{2^n}$. Then $\mathrm{wt}_s(\alpha) = 2$ if and only if m_α can be written in the form of a pentanomial or the trinomial appearing in Table 6.1.*

Proof. Let $M_{\alpha,B}$ be a multiplication matrix for some $\alpha \in \mathbb{F}_{2^n}$ and some basis $B = \{b_1, \ldots, b_n\}$. We can assume that $M_{\alpha,B}$ is in cycle normal form, $M = PA_{i_1,j_1}A_{i_2,j_2}$ with $P = \bigoplus_{k=1}^{l} C_{x^{m_k}+1}$. As a first step, we show that $l \leq 2$. Assume $l > 2$. As shown in equation (6.2) at most two rows of M have more than one '1' in them. So, by possibly permuting the blocks, P is a triangular block matrix, consisting of two blocks where one block is of the form C_{x^t+1}. So $\chi(C_{x^t+1}) = x^t + 1$ divides $\chi(M)$. But as minimal polynomial and characteristic polynomial share the same irreducible factors, this implies $(x+1)|m_\alpha$ which contradicts the irreducibility of m_α. So $l \leq 2$. We now deal with all possible matrices on a case by case basis, where we differentiate the cases $l \in \{1, 2\}$ and the two cases in equation (6.2).
Case 1. $M = C_{x^n+1} + E_{i_1,j_1} + E_{i_2,j_2}$, $j_1 \neq i_2 - 1$.
We investigate how the matrix operates on the basis $B = \{b_1, \ldots, b_n\}$:

$$
\begin{aligned}
\alpha b_1 &= b_2 \\
&\vdots \\
\alpha b_{j_1-1} &= b_{j_1} \\
\alpha b_{j_1} &= b_{j_1+1} + b_{i_1} \qquad (6.3)\\
\alpha b_{j_1+1} &= b_{j_1+2} \\
&\vdots \\
\alpha b_{j_2-1} &= b_{j_2} \\
\alpha b_{j_2} &= b_{j_2+1} + b_{i_2} \qquad (6.4)\\
\alpha b_{j_2+1} &= b_{j_2+2} \\
&\vdots \\
\alpha b_n &= b_1.
\end{aligned}
$$

Define $\gamma_1 := b_{j_1+1}$ and $\gamma_2 := b_{j_2+1}$. Then

$$b_{j_1} = \alpha^{n+j_1-j_2-1}\gamma_2, \quad b_{j_2} = \alpha^{j_2-j_1-1}\gamma_1. \tag{6.5}$$

At first, we show that the minimal polynomial has degree n. Assume $m_\alpha = x^m + \sum_{i=1}^{m-1} c_i x^i + 1$ with $c_i \in \mathbb{F}_2$ and $md = n$ with $d > 1$. In particular, $m \leq n/2$. At least one of $n + j_1 - j_2$ and $j_2 - j_1$ are greater or equal $n/2$. Assume $j_2 - j_1 \geq n/2$. Then $\alpha^i\gamma_1 = b_{j_1+1+i}$ for $i < n/2$. Furthermore, $\alpha^{n/2}\gamma_1 = b_{j_1+1+n/2}$ if $j_2 - j_1 > n/2$ and $\alpha^{n/2}\gamma_1 = b_{j_1+1+n/2} + b_{i_2}$ if $j_2 - j_1 = n/2$. Consequently, $m_\alpha(\alpha)\gamma_1 = \alpha^m\gamma_1 + \sum_{i=1}^{m-1} c_i\alpha^i\gamma_1 + \gamma_1$ is a linear combination of at least one basis element and thus cannot vanish. If $n + j_1 - j_2 \geq n/2$ the same argument holds with γ_2 instead of γ_1. So $\deg m_\alpha = n$. Observe that with the equations (6.3), (6.4) and (6.5)

$$\alpha^{n+j_1-j_2}\gamma_2 = \gamma_1 + b_{i_1} \tag{6.6}$$

$$\alpha^{j_2-j_1}\gamma_1 = \gamma_2 + b_{i_2}. \tag{6.7}$$

By plugging γ_2 into the first equation and γ_1 into the second equation, we obtain

$$\alpha^n\gamma_1 + \alpha^{n+j_1-j_2}b_{i_2} + b_{i_1} + \gamma_1 = 0 \tag{6.8}$$

$$\alpha^n\gamma_2 + \alpha^{j_2-j_1}b_{i_1} + b_{i_2} + \gamma_2 = 0. \tag{6.9}$$

Case 1.1. $i_1 \in [j_1 + 1, j_2]$ and $i_2 \in [j_1 + 1, j_2]$.

Then $b_{i_1} = \alpha^{t_1}\gamma_1$ and $b_{i_2} = \alpha^{t_2}\gamma_1$ with $t_1 = i_1 - j_1 - 1$ and $t_2 = i_2 - j_1 - 1$ with $t_1 + t_2 < n - 1$. With equation (6.8), we have

$$\alpha^n\gamma_1 + \alpha^{n+j_1-j_2+t_2}\gamma_1 + \alpha^{t_1}\gamma_1 + \gamma_1 = 0$$

So the polynomial $p = x^n + x^{n+j_1-j_2+t_2} + x^{t_1} + 1$ annihilates γ_1. Hence, p is the minimal polynomial of M. But $2 | \operatorname{wt}(p)$, so p is not irreducible. We conclude that no matrix of this type can be a multiplication matrix.

Case 1.2. $i_1 \notin [j_1 + 1, j_2]$ and $i_2 \notin [j_1 + 1, j_2]$.

Then $b_{i_1} = \alpha^{t_1}\gamma_2$ and $b_{i_2} = \alpha^{t_2}\gamma_2$ with $t_1 = i_1 - j_2 - 1 \pmod n$ and $t_2 = i_2 - j_2 - 1 \pmod n$ with $t_1 + t_2 < n - 1$. With equation (6.9), we have

$$\alpha^n\gamma_2 + \alpha^{j_2-j_1+t_1}\gamma_2 + \alpha^{t_2}\gamma_2 + \gamma_2 = 0$$

As before, the polynomial $p = x^n + x^{j_2-j_1+t_1} + x^{t_2} + 1$ annihilates γ_2, so there is no multiplication matrix of this type.

Case 1.3. $i_1 \in [j_1 + 1, j_2]$ and $i_2 \notin [j_1 + 1, j_2]$.

Then $b_{i_1} = \alpha^{t_1}\gamma_1$ and $b_{i_2} = \alpha^{t_2}\gamma_2$ with $t_1 = i_1 - j_1 - 1$ and $t_2 = i_2 - j_2 - 1 \pmod n$ with $t_1 + t_2 < n - 1$. Then by equation (6.6)

$$\gamma_2 = \alpha^{j_2-j_1-n}\gamma_1 + \alpha^{j_2-j_1-n+t_1}\gamma_1$$

and

$$b_{i_2} = \alpha^{j_2-j_1-n+t_2}\gamma_1 + \alpha^{j_2-j_1-n+t_1+t_2}\gamma_1.$$

Using equation (6.8), we obtain

$$\alpha^n\gamma_1 + \alpha^{t_1+t_2}\gamma_1 + \alpha^{t_1}\gamma_1 + \alpha^{t_2}\gamma_1 + \gamma_1 = 0,$$

so $p = x^n + x^{t_1+t_2} + x^{t_1} + x^{t_2} + 1$ is the minimal polynomial of M. Note that we can choose i_1, i_2, j_1, j_2 in a way that t_1 and t_2 take any value from $\{1, \ldots, n-3\}$ as long as $t_1 + t_2 < n - 1$, so every matrix with a minimal polynomial of the form $x^n + x^{a+b} + x^a + x^b + 1$ with $a + b \leq n - 2$ has a multiplication matrix of this type for suitable values of i_1, j_1, i_2, j_2.

Case 1.4. $i_1 \notin [j_1 + 1, j_2]$ and $i_2 \in [j_1 + 1, j_2]$.

Then $b_{i_1} = \alpha^{t_1}\gamma_2$ and $b_{i_2} = \alpha^{t_2}\gamma_1$ with $t_1 = i_1 - j_2 - 1 \pmod n$ and $t_2 = i_2 - j_1 - 1$ with $t_1 + t_2 < n - 1$. Similarly to Case 1.3, equation (6.6) yields

$$\gamma_1 = \alpha^{n+j_1-j_2}\gamma_2 + \alpha^{t_1}\gamma_2$$

and with equation (6.9)

$$\alpha^n\gamma_2 + \alpha^{j_2-j_1+t_1}\gamma_2 + \alpha^{n+j_1-j_2+t_2}\gamma_2 + \alpha^{t_1+t_2}\gamma_2 + \gamma_2 = 0,$$

so $p = x^n + x^{j_2-j_1+t_1} + x^{n+j_1-j_2+t_2} + x^{t_1+t_2} + 1 = x^n + x^{n-k_1} + x^{k_2} + x^{k_2-k_1} + 1$ with $k_1 = j_2 - j_1 - t_2 = j_2 - i_2 - 1 > 0$ and $k_2 = j_2 - j_1 + t_1$. Note that $k_2 > k_1$ for any choice of i_1, i_2, j_1, j_2. Moreover, k_1 can take on every value in $\{1, \ldots, n-3\}$ and k_2 any value greater than k_1.

Case 2. $M = C_{x^n+1} + E_{i_1,j_1} + E_{j_1+1,j_2} + E_{i_1,j_2}$.

If $j_1 = j_2$ then $\mathrm{wt}_s(M) = 1$, so we can assume $j_1 \neq j_2$. Note that the matrix operates on the basis B just as in Case 1, the only difference being that in equation (6.4) we have $b_{i_1} + b_{j_1+1} = b_{i_1} + \gamma_1$ instead of b_{i_2} on the right hand side. With the same argument as in Case 1 we conclude that the minimal polynomial of M has degree n.

Case 2.1. $i_1 \in [j_1 + 1, j_2]$.

Then $b_{i_1} = \alpha^t\gamma_1$ with $t = i_1 - j_1 - 1$. Similarly to equation (6.8), we obtain

$$\alpha^n\gamma_1 + \alpha^{n+j_1-j_2}\gamma_1 + \alpha^{n+j_1-j_2}b_{i_1} + b_{i_1} + \gamma_1 = 0$$

and thus

$$\alpha^n\gamma_1 + \alpha^{n+j_1-j_2}\gamma_1 + \alpha^{n+j_1-j_2+i_1-j_1-1}\gamma_1 + \alpha^{i_1-j_1-1}\gamma_1 + \gamma_1 = 0.$$

So the minimal polynomial of M is $p = x^n + x^{n+j_1-j_2} + x^{n-j_2+i_1-1} + x^{i_1-j_1-1} + 1$. Set

$k_1 = i_1 - j_1 - 1$ and $k_2 = n + j_1 - j_2$ then $p = x^n + x^{k_1+k_2} + x^{k_1} + x^{k_2} + 1$ with k_1, $k_2 \in \{1, \ldots, n-1\}$ and $k_1 + k_2 < n$.
Case 2.2. $i_1 \notin [j_1 + 1, j_2]$.
Then $b_{i_1} = \alpha^t \gamma_2$ with $t = i_1 - j_2 - 1 \pmod n$. Similarly to equation (6.7), we have

$$\alpha^{j_2 - j_1} \gamma_1 = \gamma_2 + \gamma_1 + \alpha^t \gamma_2.$$

Using equation (6.6) we obtain

$$\alpha^n \gamma_2 + \alpha^{j_2 - j_1 + t} \gamma_2 + \alpha^{n + j_1 - j_2} \gamma_2 + \gamma_2 = 0,$$

so the minimal polynomial of M, $p = x^n + x^{j_2 - j_1 + t} + x^{n + j_1 - j_2} + 1$, is reducible.
Case 3. $M = (C_{x^{n_1}+1} \oplus C_{x^{n_2}+1}) + E_{i_1,j_1} + E_{i_2,j_2}$, $j_1 \neq i_2 - 1$.
If both $i_1, i_2 \leq n_1$ or $i_1, i_2 > n_1$ then M is a triangular block matrix with one block being just a companion matrix. Then $(x+1) | \chi(M) = m_\alpha$, so this case cannot occur. Similarly one of j_1 and j_2 must be less or equal n_1 and the another one greater than n_1. We again investigate how M operates on the basis B:

$$\begin{aligned}
\alpha b_1 &= b_2 & \alpha b_{n_1+1} &= b_{n_1+2} \\
&\vdots & &\vdots \\
\alpha b_{j_1 - 1} &= b_{j_1} & \alpha b_{j_2 - 1} &= b_{j_2} \\
\alpha b_{j_1} &= b_{j_1+1} + b_{i_1} & \alpha b_{j_2} &= b_{j_2+1} + b_{i_2} \\
\alpha b_{j_1+1} &= b_{j_1+2} & \alpha b_{j_2+1} &= b_{j_2+2} \\
&\vdots & &\vdots \\
\alpha b_{n_1} &= b_1 & \alpha b_n &= b_{n_1+1}.
\end{aligned}$$

We set again $\gamma_1 = b_{j_1+1}$ and $\gamma_2 = b_{j_2+1}$. Then

$$\alpha^{n_1} \gamma_1 = \gamma_1 + b_{i_1} \text{ and } \alpha^{n_2} \gamma_2 = \gamma_2 + b_{i_2}. \tag{6.10}$$

Case 3.1. $i_1 \in [1, n_1]$ and $i_2 \in [n_1 + 1, n]$.
Then $b_{i_1} = \alpha^{t_1} \gamma_1$ with $t_1 = i_1 - j_1 - 1 \pmod{n_1}$ and $b_{i_2} = \alpha^{t_2} \gamma_2$ with $t_2 = i_2 - j_2 - 1 \pmod{n_2}$. M is a block diagonal matrix: $M = (C_{x^{n_1}+1} + E_{i_1,j_1}) \oplus (C_{x^{n_2}+1} + E_{i_2,j_2}) = B_1 \oplus B_2$. Let m_M, m_{B_1}, m_{B_2} be the minimal polynomial of M, B_1 and B_2. Then $m_M = \operatorname{lcm}(m_{B_1}, m_{B_2})$ and if m_M is irreducible then $m_M = m_{B_1} = m_{B_2}$. This implies that B_1 and B_2 are multiplication matrices with $\operatorname{wt}_s(B_1) = \operatorname{wt}_s(B_2) = 1$. From Theorem 6.3.6 we obtain that m_{B_1} and m_{B_2} are trinomials of degree n_1 and n_2, respectively. So $n_1 = n_2 = n/2$ and $m_M = x^{n/2} + x^t + 1$. Using equation (6.10) we can determine the choice for i_1, i_2, j_1, j_2

$$\alpha^{n/2} \gamma_1 = \gamma_1 + \alpha^{t_1} \gamma_1 \text{ and } \alpha^{n/2} \gamma_2 = \gamma_2 + \alpha^{t_2} \gamma_2.$$

Hence i_1, i_2, j_1, j_2 have to be chosen in a way that $t_1 = t_2 = t$. This is possible for every $t \in \{1, \ldots, n/2 - 1\}$.
Case 3.2. $i_1 \in [n_1 + 1, n]$ and $i_2 \in [1, n_1]$.
Then $b_{i_1} = \alpha^{t_1} \gamma_2$ with $t_1 = i_1 - j_2 - 1 \pmod{n_2}$ and $b_{i_2} = \alpha^{t_2} \gamma_1$ with $t_2 = i_2 - j_1 - 1 \pmod{n_1}$. Similarly to Case 1 we can show that the minimal polynomial of M has degree n. Applying equation (6.10) yields

$$\gamma_1 = \alpha^{n_2 - t_2} \gamma_2 + \alpha^{-t_2} \gamma_2$$

and

$$\alpha^{n-t_2}\gamma_2 + \alpha^{n_1-t_2}\gamma_2 + \alpha^{n_2-t_2}\gamma_2 + \alpha^{t_1}\gamma_2 + \alpha^{-t_2}\gamma_2 = 0.$$

Multiplying this equation by α^{t_2} we conclude that $p = x^n + x^{n_1} + x^{n_2} + x^{t_1+t_2} + 1$ annihilates γ_2, so $m_\alpha = p$. Note that $t_1 \in \{0, \ldots, n_2 - 1\}$ and $t_1 \in \{0, \ldots, n_1 - 1\}$ so $t_1 + t_2 \in \{0, \ldots, n-2\}$.

Case 4. $M = (C_{x^{n_1}+1} \oplus C_{x^{n_2}+1}) + E_{i_1,j_1} + E_{j_1+1 \pmod{n_1},j_2} + E_{i_1,j_2}$.

Again, we can assume $j_1 \neq j_2$. Note that the matrix operates on the basis B just as in Case 3, the only difference being that b_{i_2} is substituted by $b_{i_1} + b_{j_1+1} = b_{i_1} + \gamma_1$. This leads to

$$\alpha^{n_2}\gamma_2 = \gamma_2 + \gamma_1 + \alpha^t\gamma_2. \tag{6.11}$$

With the same argument as before we conclude that the minimal polynomial of M has degree n. If $i_1 \in [1, n_1]$ then M is again a block triangular matrix with one block being a companion matrix, so this case cannot occur. So $i_1 \in [n_1 + 1, n]$ and $b_{i_1} = \alpha^t\gamma_2$ for $t_1 = i_1 - j_2 - 1 \pmod{n_2}$. Similarly to Case 3.2 we get

$$\gamma_2 = \alpha^{n_1-t}\gamma_1 + \alpha^{-t}\gamma_1.$$

Combining this equation with equation (6.11) we have

$$\alpha^{n-t}\gamma_1 + \alpha^{n_2-t}\gamma_1 + \alpha^{n_1}\gamma_1 + \alpha^{n_1-t}\gamma_1 + \alpha^{-t}\gamma_1 = 0$$

and after multiplying with α^t we conclude that $m_\alpha = x^n + x^{n_1+t} + x^{n_2} + x^{n_1} + 1$, where $t \in \{1, \ldots, n_2 - 1\}$. □

Cases 1 and 3 of Theorem 6.3.7 also provide all elements α with $\mathrm{wt}_d(\alpha) = 2$. Moreover, Theorem 4 in [5] is a slightly weaker version of Case 1.3. in Theorem 6.3.7.

Remark 6.3.8. A suitable choice for the values i_1, j_1, i_2, j_2 in the second column of Table 6.1 can be found in the proof of the corresponding case.

The following example shows that the cycle normal forms of optimal s-XOR-representations are generally not unique.

Example 6.3.9. Let $\alpha \in \mathbb{F}_{2^4}$ with the irreducible minimal polynomial $m_\alpha = x^4 + x^3 + x^2 + x + 1$. Then, by Theorem 6.3.7, $\mathrm{wt}_s(\alpha) = \mathrm{wt}_d(\alpha) = 2$ and $M = C_{x^4+1} + E_{2,2} + E_{3,4}$ and $M' = (C_{x^3+1} \oplus C_{x+1}) + E_{3,4} + E_{4,3}$ belong to two different optimal representations, corresponding to Case 1.4. and Case 3.2 of Theorem 6.3.7, respectively.

The following corollary is a direct result from Theorem 6.3.7 and Example 6.3.5.

Corollary 6.3.10. *Let* $\alpha \in \mathbb{F}_{2^n}$ *with* $\mathrm{wt}(m_\alpha) = 5$ *and* $\deg(m_\alpha) = n$. *Then* $\mathrm{wt}_s(\alpha) = 2$ *if* f *appears in Table 6.1 and* $\mathrm{wt}_s(\alpha) = 3$ *otherwise.*

Corollary 6.3.10 shows that an implementation via the rational canonical form (as in Example 6.3.5) is generally not the best way to implement multiplication in binary finite fields. However, irreducible pentanomials that do not appear in the table in Theorem 6.3.7 exist, the examples with the lowest degree are $f = x^8 + x^6 + x^5 + x^4 + 1$ and its reciprocal polynomial (for a table of all s-XOR-counts of finite field elements in $\mathbb{F}_{2^n}$ for $n \leq 8$ see [5]). It is an interesting question for which field elements the "naive" representation using the rational canonical form is optimal.

6.4 Quantifying the Gap between the Optimal Implementation and the Naive Implementation

It is now interesting to investigate the gap between the optimal implementation and the "naive" implementation using the rational canonical form. We give a partial answer to this question in Theorem 6.4.1. The proofs in this section are technical and rely on many calculations with matrices. To increase readability, the proofs are included only in the appendix.

Theorem 6.4.1. *Let $\alpha \in \mathbb{F}_{2^n}$ be not contained in a proper subfield of $\mathbb{F}_{2^n}$ and let $M_{\alpha,B}$ be a multiplication matrix of α with respect to some basis B. Then $\mathrm{wt}_d(M_{\alpha,B}) = t$ implies $\mathrm{wt}(m_\alpha) \leq 2^t + 1$.*

We now show that the bound given in Theorem 6.4.1 is optimal by giving two examples where the upper bound is attained. Note that the proof of Theorem 6.4.1 implies that this can only occur if the number of blocks of the optimal multiplication matrix is 1 or t. We will give examples for both cases in Propositions 6.4.2 and 6.4.3.

Proposition 6.4.2. *Let $\alpha \in \mathbb{F}_{2^n}$ with an irreducible minimal polynomial f with $\mathrm{wt}(f) = 2^t + 1$ of the form*

$$f = x^n + \prod_{j=1}^{t} \left(x^{i_j} + 1\right)$$

for arbitrary values of $i_j \in \mathbb{N}$ with $\sum_{j=1}^{t} i_j \leq n - t$. Then there exists a basis B such that the matrix $M := M_{\alpha,B}$ satisfies $\mathrm{wt}_s(M) = \mathrm{wt}_d(M) = t$.

Proposition 6.4.3. *Let $\alpha \in \mathbb{F}_{2^n}$ with an irreducible minimal polynomial f with $\mathrm{wt}(f) = 2^t + 1$ of the form*

$$f = \prod_{j=1}^{t} (x^{n_j} + 1) + x^k$$

for arbitrary values of n_j and $k \leq n - t$ with $\sum_{j=1}^{t} n_j = n$. Then there exists a basis B such that the matrix $M := M_{\alpha,B}$ satisfies $\mathrm{wt}_s(M) = \mathrm{wt}_d(M) = t$.

Observe that the polynomials in Propositions 6.4.2 and 6.4.3 are generalizations of Case 1.3. and Case 3.2. in Theorem 6.3.7.

Note that irreducible polynomials of the types mentioned in Propositions 6.4.2 and 6.4.3 do exist, examples up to $t = 8$, corresponding to polynomials of weight $2^t + 1$, are compiled in Table 6.2. The table lists in the second column values for i_l and n that belong to an irreducible polynomial of the type of Proposition 6.4.2 and in the third column the values for n_l and k that belong to an irreducible polynomial of the type of Proposition 6.4.3. The values listed were found with a simple randomized algorithm. They generally do not correspond to the irreducible polynomial of that type with the least degree. Propositions 6.4.2 and 6.4.3 together with Theorem 6.4.1 show that the gap between the number of XORs used in the optimal implementation and the number of XORs used in the naive implementation of a multiplication matrix using the rational canonical form grows exponentially with the weight of the minimal polynomial of the element.

Propositions 6.4.2 and 6.4.3 show that there are elements $\alpha \in \mathbb{F}_{2^n}$ with $\mathrm{wt}(m_\alpha) = 2^t + 1$ and $\mathrm{wt}_s(\alpha) = t$. We believe that this upper bound is strict, i.e. the bound is the same for s-XOR-count and d-XOR-count.

Conjecture 6.4.4. Let $\alpha \in \mathbb{F}_{2^n}$ be not contained in a proper subfield of $\mathbb{F}_{2^n}$ and $M_{\alpha,B}$ a multiplication matrix of α with respect to some basis B. Then $\mathrm{wt}_s(M_{\alpha,B}) = t$ implies $\mathrm{wt}(\chi(M)) \leq 2^t + 1$.

t	values for $i_1, \ldots, i_t; n$	values for $n_1, \ldots, n_t; k$
2	1,2;5	2,4;1
3	1,2,4;10	4,5,6;1
4	3,5,6,12;30	2,3,6,10;1
5	1,2,4,9,17;39	12,13,15,19,23;9
6	1,12,16,24,31; 123	13,22,26,27,28,30;23
7	2,30,47,56,60,64,91; 357	25,114,174,231,279,281,331;196
8	23,28,41,59,62,106,141,153; 628	44,148,195,357,363,368,386,480;240

TABLE 6.2: Irreducible polynomials of the form described in Propositions 6.4.2 and 6.4.3.

6.5 Open Problems

Our investigations open up many possibilities for future research. While Theorem 6.2.4 shows that there is an infinite family of matrices with higher s-XOR-count than d-XOR-count, a more precise classification of these cases as well as finding upper/lower bounds is desirable. Because of the nature of the s-XOR-count, answers to these problems would also give insight into optimal Gauss elimination strategies over $\mathbb{F}_2$.

Problem 6.5.1. Classify the matrices $M \in \mathrm{GL}(n, \mathbb{F}_2)$ with $\mathrm{wt}_d(M) < \mathrm{wt}_s(M)$.

Problem 6.5.2. Find bounds c, C so that $c\,\mathrm{wt}_d(M) \leq \mathrm{wt}_s(M) \leq C\,\mathrm{wt}_d(M)$ for all matrices $M \in \mathrm{GL}(n, \mathbb{F}_2)$.

Finding out if/how the bounds c, C depend on n and $\mathrm{wt}_s(M)$ would greatly improve the understanding of the two XOR-metrics.

As observed in Section 6.3, there are elements $\alpha \in \mathbb{F}_2$ where the optimal implementation of the mapping $x \mapsto \alpha x$ is the rational canonical form in both of the investigated metrics. These elements are (compared to elements with minimal polynomials of the same weight) the most expensive to implement. A more thorough understanding of these elements would be helpful.

Problem 6.5.3. Classify the minimal polynomials $m_\alpha \in \mathbb{F}_2[x]$ for which the optimal multiplication matrix is in rational canonical form.

We also want to repeat a problem about elements in subfields mentioned in [5].

Problem 6.5.4. Let $\alpha \in \mathbb{F}_{2^n}$ be contained in a subfield $\mathbb{F}_{2^l}$ with $ld = n$. Let M_l be an optimal multiplication matrix of α regarding d- or s-XOR-count. Is $M = \bigoplus_{k=1}^{d} M_l$ then an optimal multiplication matrix of $\alpha \in \mathbb{F}_{2^n}$ regarding d- or s-XOR-count?

In Sections 6.3 and 6.4 we limited ourselves to optimal XOR-implementations of matrices that are multiplication matrices for a fixed field element (which are exactly those with irreducible minimal polynomial). Investigating a more general case is also an interesting problem.

Problem 6.5.5. Let $f\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a bijective linear mapping and $M_{f,B} \in \mathrm{GL}(n, \mathbb{F}_2)$ the matrix that belongs to f with respect to the basis B. Find a basis B such that the matrix $M_{f,B}$ is the optimal d/s-XOR-count matrix.

In particular, finding optimal matrices $M_{f,B}$ where f denotes the mapping induced by a linear layer of a cryptographic scheme is a very interesting problem.

Appendix A

Appendix: Proofs of Section 6.4

A.1 Proof of Theorem 6.4.1

For a square matrix $M = (m_{r,s})$ over $\mathbb{F}_2$ and two index sequences (ordered sets) $I = (i_1, \ldots, i_{l_1}), J = (j_1, \ldots, j_{l_2}), l := \min(l_1, l_2)$ we denote by $M^{I,J} = (a_{r,s})$ the matrix that is constructed as follows: All rows in I and all columns in J are filled with zeroes, except the entries $a_{i_1,j_1}, \ldots, a_{i_l,j_l}$ which are set to 1. More precisely:

$$a_{r,s} = \begin{cases} 0, & r = i_k, s \neq j_k \text{ for a } k \in \{1, \ldots, l_1\} \\ 0, & r \neq i_k, s = j_k \text{ for a } k \in \{1, \ldots, l_2\} \\ 1, & r = i_k, s = j_k \text{ for a } k \in \{1, \ldots, l\} \\ m_{r,s}, & \text{otherwise.} \end{cases}$$

The following example illustrates our notation. Let $I = \{2, 4\}$ and $J = \{1, 3\}$.

$$M = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad M^{I,J} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

In the case that $l_1 \neq l_2$ the matrix $M^{I,J}$ has a zero row/column and is thus not invertible. If $l_1 = l_2$, it is easy to see that $\det(M^{I,J})$ does not depend on the ordering of the index sets I, J and is the same as the determinant of the matrix that is created by deleting all rows of M in I and all columns of M in J. In the case that we are only concerned with the determinant, we will thus just use (unordered) index sets I, J and also talk about determinants of submatrices. If $I = \{i\}$ and $J = \{j\}$ we will also write $M^{(i,j)}$. Moreover, we denote by A_M the characteristic matrix $A_M := xI + M$ of M.

Lemma A.1.1. *Let $M = C_{x^n+1} \in \mathrm{GL}(n, \mathbb{F}_2)$. Then we have $\mathrm{wt}(\det(A_M^{I,J})) \leq 1$ for all possible proper square submatrices $A_M^{I,J}$.*

Proof. The proof is by induction on the size of the submatrix. Clearly, $\det(A_M^{I,J}) \in \{0, 1, x\}$ if $|I| = |J| = n - 1$. Let now $|I| < n - 1$. We denote by c_{ij} the entry in the i-th row and j-th column of A_M. Then

$$c_{ij} = \begin{cases} x, & i = j, \\ 1, & i = j + 1 \pmod{n}, \\ 0, & \text{else.} \end{cases}$$

Let $i \in I$. If $i \notin J$, then $A_M^{I,J}$ has at most one non-zero entry in the i-th column. Then, by Laplace expansion along the i-th column and use of the induction hypothesis, we get $\mathrm{wt}(\det(A_M^{I,J})) \leq 1$. If $i \in J$ and $i+1 \pmod n \notin I$ then the $i+1 \pmod n$-th row has at most one non-zero entry and Laplace expansion along the $i+1 \pmod n$-th row yields $\mathrm{wt}(\det(A_M^{I,J})) \leq 1$. We conclude that $\mathrm{wt}(\det(A_M^{I,J})) \leq 1$ for all I with $|I| < n$. □

Lemma A.1.2. *Let $M = C_{x^n+1} + \sum_{k=1}^{t} E_{i_k,j_k}$ where i_k, j_k can be chosen arbitrarily. Then we have $\mathrm{wt}(\det(A_M^{I,J})) \leq 2^t$ for all possible proper square submatrices $A_M^{I,J}$.*

Proof. The proof is by induction on t. The case $t = 0$ is covered by Lemma A.1.1. Let now $t \geq 1$. Let $M' = C_{x^n+1} + \sum_{k=1}^{t-1} E_{i_k,j_k}$, so that $M = M' + E_{i,j}$ with $i = i_k$, $j = j_k$. If $i \in I$ or $j \in J$ we have $A_M^{I,J} = A_{M'}^{I,J}$ and thus $\mathrm{wt}(\det(A_M^{I,J})) = \mathrm{wt}(\det(A_{M'}^{I,J})) \leq 2^{t-1}$. If $i \notin I$ and $j \notin J$ then $A_M^{I,J} = A_{M'}^{I,J} + E_{i,j}$ and thus Laplace expansion along the i-th row yields $\det(A_M^{I,J}) \leq \det(A_{M'}^{I,J}) + \det(A_{M'}^{I\cup\{i\},J\cup\{j\}})$ and thus

$$\mathrm{wt}(\det(A_M^{I,J})) \leq \mathrm{wt}(\det(A_{M'}^{I,J})) + \mathrm{wt}(\det(A_{M'}^{I\cup\{i\},J\cup\{j\}})) \leq 2^{t-1} + 2^{t-1} = 2^t$$

by induction hypothesis. □

Corollary A.1.3. *Let $M = C_{x^n+1} + \sum_{k=1}^{t} E_{i_k,j_k}$ where i_k, j_k can be chosen arbitrarily. Then $\mathrm{wt}(\chi(M)) \leq 2^t + 1$.*

Proof. The proof is by induction on t. The case $t = 0$ holds because $\chi(C_{x^n+1}) = x^n + 1$ by definition of the companion matrix. Let now $t \geq 1$ and $M' = C_{x^n+1} + \sum_{k=1}^{t-1} E_{i_k,j_k}$. Laplace expansion along the i_t-th row yields $\chi(M) = \det(A_M) = \chi(M') + \det(A_{M'}^{(i_t,j_t)})$. We conclude with Lemma A.1.2 and the induction hypothesis that $\mathrm{wt}(\chi(M)) \leq 2^{t-1} + 1 + 2^{t-1} = 2^t + 1$. □

Proof of Theorem 6.4.1. Let B be an optimal (regarding the d-XOR-count) basis and $M := M_{\alpha,B} = \bigoplus_{k=1}^{l} C_{x^{m_k}+1} + \sum_{r=1}^{t} E_{i_r,j_r}$ be an optimal multiplication matrix. The case $l = 1$ is covered in Corollary A.1.3, so we only consider $l > 1$ for the rest of the proof. Since α is not contained in a proper subfield of $\mathbb{F}_{2^n}$, the minimal polynomial of M coincides with its characteristic polynomial. We call the sets

$$\{1, \ldots, m_1\}, \{m_1 + 1, \ldots, m_2\}, \ldots, \{\sum_{k=1}^{l-1} m_k + 1, \ldots, \sum_{k=1}^{l} m_k\}$$

the l *blocks* of M. We can decompose $M = M_1 + M'$ with $M_1 = \bigoplus_{k=1}^{l} C_{x^{m_k}+1} + \sum_{r=1}^{t_1} E_{i_r,j_r}$ and $M' = \sum_{r=1}^{t_2} E_{i_r,j_r}$ in a way that all pairs (i_r, j_r) in M_1 are in the same block and all pairs (i_r, j_r) in M' are in different blocks. M_1 is a block diagonal matrix and with Corollary A.1.3 we get

$$\mathrm{wt}(\chi(M_1)) \leq \prod_{k=1}^{l} (2^{s_k} + 1) \text{ with } \sum_{k=1}^{l} s_k = t_1 \tag{A.1}$$

where s_k denotes the number of pairs (i_r, j_r) that are in the k-th block. We call B_1, $\ldots$, B_l the l blocks of M_1 and $m_1, \ldots, m_l$ the size of these blocks. Note that $\chi(M)$ is irreducible which implies that M is not a block triangular matrix and thus $t_2 \geq l$. So we can write $M' = M_2 + M_3$ in a way that (after a suitable permutation of blocks)

$M_1 + M_2$ looks like this:

$$M_1 + M_2 = \begin{pmatrix} B_1 & 0 & \dots & E_{i_l,j_l} \\ E_{i_1,j_1} & B_2 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & E_{i_{l-1},j_{l-1}} & B_l \end{pmatrix}. \tag{A.2}$$

From this, we infer by Laplace expansion along the i_l-th row

$$\chi(M_1 + M_2) = \chi(M_1) + \det(A^{(i_l,v)}_{M_1+M_2}), \tag{A.3}$$

where $v = \sum_{k=1}^{l-1} m_k + j_l$. We now determine $\mathrm{wt}(\det(A^{(i_l,v)}_{M_1+M_2}))$. We get

$$\begin{aligned} \det(A^{(i_l,v)}_{M_1+M_2}) &= \det \begin{pmatrix} B_1^{(i_l,\emptyset)} & 0 & \dots & 0 & E_{i_l,j_l} \\ E_{i_1,j_1} & B_2 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \ddots & E_{i_{l-2},j_{l-2}} & B_{l-1} & 0 \\ 0 & \dots & 0 & E_{i_{l-1},j_{l-1}} & B_l^{(\emptyset,j_l)} \end{pmatrix} \\ &= \det \begin{pmatrix} B_1^{(i_l,\emptyset)} & 0 & \dots & E_{i_{l-1},j_{l-1}} & * \\ E_{i_1,j_1} & B_2 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \ddots & E_{i_{l-2},j_{l-2}} & B_{l-1} & 0 \\ 0 & \dots & 0 & 0 & B_l^{(i_{l-1},j_l)} \end{pmatrix} \end{aligned}$$

by swapping the i_l-th row with the $\sum_{k=1}^{l-1} m_k + i_{l-1}$-th row. This operation can now be repeated for the upper-left $l-1$ blocks, the result is the following block diagonal matrix

$$\det(A^{(i_l,v)}_{M_1+M_2}) = \det \begin{pmatrix} B_1^{(i_l,j_1)} & * & \dots & 0 & 0 \\ 0 & B_2^{(i_1,j_2)} & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \ddots & 0 & B_{l-1}^{(i_{l-2},j_{l-1})} & * \\ 0 & \dots & 0 & 0 & B_l^{(i_{l-1},j_l)} \end{pmatrix}.$$

Lemma A.1.2 then implies $\mathrm{wt}(\det(A^{(i_l,v)}_{M_1+M_2})) \leq \prod_{k=1}^{l} 2^{s_k} = 2^{t_1}$. Equations (A.1) and (A.3) yield

$$\mathrm{wt}(\chi(M_1 + M_2)) \leq \prod_{k=1}^{l} (2^{s_k} + 1) + 2^{t_1}. \tag{A.4}$$

We now investigate the determinant of the square submatrices of $M_1 + M_2$. Let I, J be index sets and set $I = \dot{\bigcup}_r I_r$ and $J = \dot{\bigcup}_r J_r$ where I_r and J_r contain the indices that belong to the r-th block. Observe that $|I| = |J|$. Let us first look at the case $I = I_r$ and $J = J_r$ for some r. Using Lemma A.1.2

$$\mathrm{wt}(\det(A^{I,J}_{M_1+M_2})) \leq 2^{s_r} \prod_{\substack{k \in \{1,\dots,l\} \\ k \neq r}} (2^{s_k} + 1) + 2^{t_1}.$$

Similarly, if $|I_r| = |J_r|$ for all $1 \leq r \leq l$ then

$$\mathrm{wt}(\det(A^{I,J}_{M_1+M_2})) \leq \prod_{r:I_r\neq\emptyset} 2^{s_r} \prod_{r:I_r=\emptyset} (2^{s_k}+1) + 2^{t_1}. \tag{A.5}$$

Let us now assume that there is a block r with $|I_r| \neq |J_r|$. We can assume w.l.o.g. that $r = 1$ and $p := |I_1| < |J_1|$. If $i_1 + m_1, i_l \in I$ or $v, j_1 \in J$ then equation (A.2) implies $\det(A^{I,J}_{M_1+M_2}) = 0$. We order $I = (a_1, \ldots, a_t)$ and $J = (b_1, \ldots, b_t)$ in ascending order. Then

$$\det(A^{I,J}_{M_1+M_2}) = \det \begin{pmatrix} B_1^{I_1,J_1} & A \\ C & D \end{pmatrix},$$

with $A = (a_{r,s}) \in \mathbb{F}_2^{m_1\times(n-m_1)}$, $C = (c_{r,s}) \in \mathbb{F}_2^{(n-m_1)\times m_1}$ with

$$a_{r,s} = \begin{cases} 1, & \text{for } (r,s) = (i_l, v), \\ 0, & \text{else,} \end{cases} \quad c_{r,s} = \begin{cases} 1, & \text{for } (r,s) = (a_k, b_k), k > p, \\ 1, & \text{for } (r,s) = (i_1, j_1), \\ 0, & \text{else.} \end{cases}$$

Swapping the i_l-th row with the a_{p+1}-th row, we obtain

$$\det(A^{I,J}_{M_1+M_2}) = \det \begin{pmatrix} B_1^{I_1\cup\{i_l\},J_1} & 0 \\ * & D' \end{pmatrix}$$

and thus $\det(A^{I,J}_{M_1+M_2}) = \det(B_1^{I_1\cup\{i_l\},J_1}) \det(D')$. Observe that $\det(A^{I,J}_{M_1+M_2}) = 0$ if $|I_1| \neq |J_1| + 1$. Moreover, $\det(D') = \det(C^{I',J'}_{M_1+M_2})$ where $\{1, \ldots, m_1\}$ is a subset of I' and J'. In particular, the number of indices in I' and J' belonging to the first block is the same. By induction, equation (A.5) and Lemma A.1.2, we get

$$\mathrm{wt}(\det(A^{I,J}_{M_1+M_2})) = \mathrm{wt}(\det(B_1^{I_1\cup\{i_l\},J_1}) \det(D')) \leq 2^{s_1} \prod_{k=2}^{l} (2^{s_k}+1) + 2^{t_1}. \tag{A.6}$$

Equations (A.5) and (A.6) imply that for arbitrary index sets I, J, there exists an $r \in \{1, \ldots, l\}$ such that

$$\mathrm{wt}(\det(A^{I,J}_{M_1+M_2})) \leq 2^{s_r} \prod_{\substack{k\in\{1,\ldots,l\} \\ k\neq r}} (2^{s_k}+1) + 2^{t_1}. \tag{A.7}$$

As in the proof of Lemma A.1.2, for arbitrary index sets I, J and $i, j \in \{1, \ldots, n\}$ there is an $r \in \{1, \ldots, l\}$ such that

$$\begin{aligned} \mathrm{wt}(\det(A^{I,J}_{M_1+M_2+E_{i,j}})) &\leq \mathrm{wt}(\det(A^{I,J}_{M_1+M_2})) + \mathrm{wt}(\det(A^{I\cup\{i\},J\cup\{j\}}_{M_1+M_2})) \\ &\leq 2 \cdot \left(2^{s_r} \prod_{\substack{k\in\{1,\ldots,l\} \\ k\neq r}} (2^{s_k}+1) + 2^{t_1} \right) \end{aligned}$$

and, inductively, for an arbitrary matrix $M_3 = \sum_{k=1}^{z} E_{i_k,j_k}$ with z non-zero entries

$$\begin{aligned}\mathrm{wt}(\det(A^{I,J}_{M_1+M_2+M_3})) &\leq 2^z \left(2^{s_r} \prod_{\substack{k\in\{1,\ldots,l\}\\ k\neq r}} (2^{s_k}+1) + 2^{t_1} \right) \\ &< 2^z \left(\prod_{k=1}^{l} (2^{s_k}+1) + 2^{t_1} \right). \end{aligned} \tag{A.8}$$

We now show by induction that we have for $z \geq 1$

$$\mathrm{wt}(\chi(M_1+M_2+M_3)) < 2^z \left(\prod_{k=1}^{l} (2^{s_k}+1) + 2^{t_1} \right). \tag{A.9}$$

The case $z = 1$ is dealt with using equations (A.4) and (A.7):

$$\begin{aligned}\mathrm{wt}(\chi(M_1+M_2+M_3)) &\leq \mathrm{wt}(\chi(M_1+M_2)) + \mathrm{wt}(\det(A^{i_1,j_1}_{M_1+M_2})) \\ &< 2 \left(\prod_{k=1}^{l} (2^{s_k}+1) + 2^{t_1} \right).\end{aligned}$$

Let now $z > 1$ and $M_3' = \sum_{k=1}^{z-1} E_{i_k,j_k}$. With the induction hypothesis and equation (A.8) we conclude

$$\begin{aligned}\mathrm{wt}(\chi(M_1+M_2+M_3)) &\leq \mathrm{wt}(\chi(M_1+M_2+M_3')) + \mathrm{wt}(\det(A^{i_1,j_1}_{M_1+M_2+M_3'})) \\ &< 2^z \left(\prod_{k=1}^{l} (2^{s_k}+1) + 2^{t_1} \right),\end{aligned}$$

proving equation (A.9). Note that the bound in equation (A.9) depends only on the parameters l, t_2 and $s_k, k = 1, \ldots, l$ where $\sum_{k=1}^{l} s_k = t_1$ and $t_1 + t_2 = t = \mathrm{wt}(M)$. For $t_2 > l$ we have

$$\mathrm{wt}(\chi(M_1+M_2+M_3)) < 2^{t_2-l} \left(\prod_{k=1}^{l} (2^{s_k}+1) + 2^{t_1} \right).$$

Using equation (A.4), a matrix N with values $l_N = t_2$ and $s_k = 0$ for $k > l$ yields

$$\begin{aligned}\mathrm{wt}(\chi(N)) &\leq \prod_{k=1}^{l_N} (2^{s_k}+1) + 2^{t_1} \\ &= 2^{t_2-l} \prod_{k=1}^{l} (2^{s_k}+1) + 2^{t_1}.\end{aligned}$$

In particular, the upper bound given in equation (A.9) is always worse than the one given in equation (A.4) and we can focus on the case $M_3 = 0$ (or, equivalently, $t_2 = l$) for the rest of this proof. In other words, we just have to find the parameters that give the maximum weight estimation in equation (A.4). A direct calculation yields

$$\prod_{k=1}^{l} (2^{s_k}+1) \leq (2^{t_1}+1) \cdot 2^{l-1},$$

i.e. the choice $s_1 = t_1$, $s_i = 0$ for $i > 1$ is optimal. Plugging these parameters into equation (A.4), we get

$$\mathrm{wt}(\chi(M)) \leq 2^{t_1+l-1} + 2^{l-1} + 2^{t_1} = 2^{t-1} + 2^{l-1} + 2^{t-l}. \tag{A.10}$$

Obviously, the maximum of $2^{l-1} + 2^{t-l}$ for $2 \leq l \leq t$ is attained at $l = t$. The result follows from equation (A.10).

□

A.2 Proof of Propositions 6.4.2 and 6.4.3

Theorem A.2.1 ([73, Theorem 3.5], [63, Theorem 4.3.9]). *Let R be a (commutative) Euclidean domain and $A \in R^{n\times n}$. Then A can be transformed into an upper triangular matrix using elementary row operations (i.e. a sequence of left-multiplications with matrices $I + rE_{i,j}$ with $r \in R$ and $i \neq j$).*

Proof of Proposition 6.4.2. We show that the matrix $M = C_{x^n+1} + \sum_{k=1}^{t-1} E_{j_k+i_k+1,j_k} + E_{i_t+n-j_t,j_t}$ where the j_k are chosen arbitrarily under the conditions that $j_{k+1} \geq i_k + j_k + 1$ for all $k = 1, \ldots, t-1$ and $i_t < j_1$ has the desired property. It is clear that $\mathrm{wt}_s(M) = \mathrm{wt}_d(M) = t$. Let $B = \{b_1, \ldots, b_n\}$ be some basis of $(\mathbb{F}_2)^n$ over $\mathbb{F}_2$. We investigate how M (viewed as a transformation matrix) operates on this basis:

$$\begin{aligned}
Mb_1 &= b_2 \\
&\vdots \\
Mb_{j_1-1} &= b_{j_1} \\
Mb_{j_1} &= b_{j_1+1} + M^{i_1} b_{j_1+1} \qquad (A.11)\\
Mb_{j_1+1} &= b_{j_1+2} \\
&\vdots \\
Mb_{j_2-1} &= b_{j_2} \\
Mb_{j_2} &= b_{j_2+1} + M^{i_2} b_{j_2+1} \qquad (A.12)\\
Mb_{j_2+1} &= b_{j_2+2} \\
&\vdots \\
Mb_n &= b_1.
\end{aligned}$$

Set $n_i = j_i - j_{i-1}$ for $2 \leq i \leq t$ and $n_1 = n + j_1 - j_t$. Note that $\sum_{i=1}^{t} n_i = n$ and $Mb_{j_k} = M^{n_i} b_{j_{k-1}+1}$. With this and the equations of type (A.11) and (A.12) we obtain the following set of equations:

$$\begin{pmatrix}
M^{n_2} & M^{i_1}+1 & 0 & \ldots & 0 \\
0 & M^{n_3} & M^{i_2}+1 & \ldots & 0 \\
 & & \ddots & \ddots & \\
0 & \ldots & 0 & M^{n_t} & M^{i_{t-1}}+1 \\
M^{i_t}+1 & 0 & \ldots & 0 & M^{n_1}
\end{pmatrix}
\begin{pmatrix}
b_{j_1+1} \\ b_{j_2+1} \\ \vdots \\ \vdots \\ b_{j_t+1}
\end{pmatrix} = 0. \tag{A.13}$$

We denote by A the matrix in equation (A.13). A is a matrix over $\mathbb{F}_2[M]$. It is clear that $\mathbb{F}_2[M]$ is isomorphic to the usual polynomial ring $\mathbb{F}_2[x]$ and thus a Euclidean domain. Using the Leibniz formula for determinants, we obtain $\det(A) = f(M)$. By

Theorem A.2.1, we can transform A into an upper triangular matrix A' using only elementary row operations. In particular $\det(A') = \prod_{i=1}^{n} a'_{i,i} = \det(A) = f(M)$ where the $a'_{i,i}$ denote the entries on the diagonal of A'. Since f is irreducible, we obtain $a_{k,k} = f(M)$ for one $1 \leq k \leq n$ and $a_{i,i} = 1$ for all $i \neq k$, i.e.

$$\begin{pmatrix} 1 & & & & * \\ & \ddots & & & \\ & & f(M) & * & * \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix} \begin{pmatrix} b_{j_1+1} \\ \vdots \\ b_{j_k+1} \\ \vdots \\ b_{j_t+1} \end{pmatrix} = 0.$$

It is clear that all entries $a'_{k,k+1}, \ldots, a'_{k,n}$ can be eliminated by further row additions. Hence, we obtain $f(M)b_{j_k+1} = 0$, i.e. f is the M-annihilator of b_{j_k+1}. As f is irreducible this implies that the minimal polynomial of M is f and thus M is a multiplication matrix of α.

□

Proof of Proposition 6.4.3. The proof is similar to the proof of Proposition 6.4.2. Define $\hat{n}_l = \sum_{u=1}^{l-1} n_u$ for $1 \leq l \leq t$. Let r_l be chosen arbitrarily such that $1 \leq r_l \leq n_l$ for $1 \leq l \leq t$ and $\sum_{l=1}^{t} r_l = k$. Further let $j_l := \hat{n}_l + r_l$ for all $1 \leq l \leq t$ and $s_l := i_l + r_{l+1} + 1 \pmod{n_{l+1}}$ for $l < t$ and $s_t := i_t + r_1 + 1 \pmod{n_1}$.

Define now $M = \bigoplus_{i=1}^{t} C_{x^{n_i}+1} + \sum_{k=1}^{t} E_{\hat{n_k}+s_k,j_k}$. Obviously, $\mathrm{wt}_s(M) = \mathrm{wt}_d(M) = t$. Let $B = \{b_1, \ldots, b_n\}$ be some basis of $(\mathbb{F}_2)^n$ over $\mathbb{F}_2$. We investigate how M (viewed as a transformation matrix) operates on this basis:

$$\begin{array}{lll}
Mb_1 = b_2 & Mb_{n_1+1} = b_{n_1+2} & \ldots\ Mb_{n_{t-1}+1} = b_{n_t+2} \\
\vdots & \vdots & \vdots \\
Mb_{j_1-1} = b_{j_1} & Mb_{j_2-1} = b_{j_2} & Mb_{j_t-1} = b_{j_t} \\
Mb_{j_1} = b_{j_1+1} + M^{i_1}b_{j_2+1} & Mb_{j_2} = b_{j_2+1} + M^{i_2}b_{j_3+1} & Mb_{j_t} = b_{j_t+1} + M^{i_t}b_{j_1+1} \\
Mb_{j_1+1} = b_{j_1+2} & Mb_{j_2+1} = b_{j_2+2} & Mb_{j_t+1} = b_{j_t+2} \\
\vdots & \vdots & \vdots \\
Mb_{n_1} = b_1 & Mb_{n_2} = b_{n_1+1} & \ldots\ Mb_n = b_{n_{t-1}+1}.
\end{array}$$

Clearly, $Mb_{j_k} = M^{n_k}b_{j_k+1}$, so we get the following set of equations:

$$\begin{pmatrix} M^{n_1}+1 & M^{i_1} & 0 & \ldots & 0 \\ 0 & M^{n_2}+1 & M^{i_2} & \ldots & 0 \\ & & \ddots & \ddots & \\ 0 & \ldots & 0 & M^{n_{t-1}}+1 & M^{i_{t-1}} \\ M^{i_t} & 0 & \ldots & 0 & M^{n_t}+1 \end{pmatrix} \begin{pmatrix} b_{j_1+1} \\ b_{j_2+1} \\ \vdots \\ \vdots \\ b_{j_t+1} \end{pmatrix} = 0.$$

The determinant of the matrix is exactly $f(M)$. We can now repeat the arguments from the proof of Proposition 6.4.2 and obtain that M is a multiplication matrix for α. □

Bibliography

[1] Y. Aubry, D. J. Katz, and P. Langevin. Cyclotomy of Weil sums of binomials. *Journal of Number Theory*, 154:160 –178, 2015.

[2] J. Ax. Zeroes of Polynomials Over Finite Fields. *American Journal of Mathematics*, 86(2):255, Apr. 1964.

[3] S. Banik, S. K. Pandey, T. Peyrin, Y. Sasaki, S. M. Sim, and Y. Todo. GIFT: a small present. In W. Fischer and N. Homma, editors, *Cryptographic Hardware and Embedded Systems – CHES 2017*, pages 321–345, Cham. Springer International Publishing, 2017.

[4] A. Bar-On, O. Dunkelman, N. Keller, and A. Weizman. DLCT: a new tool for differential-linear cryptanalysis. In Y. Ishai and V. Rijmen, editors, *Advances in Cryptology – EUROCRYPT 2019*, pages 313–342, Cham. Springer International Publishing, 2019.

[5] C. Beierle, T. Kranz, and G. Leander. Lightweight multiplication in $GF(2^N)$ with applications to MDS matrices. In *Proceedings, Part I, of the 36th Annual International Cryptology Conference on Advances in Cryptology — CRYPTO 2016 - Volume 9814*, pages 625–653, New York, NY, USA. Springer-Verlag New York, Inc., 2016.

[6] T. P. Berger, A. Canteaut, P. Charpin, and Y. Laigle-Chapuy. On almost perfect nonlinear functions over $\mathbb{F}_2^n$. *IEEE Transactions on Information Theory*, 52(9):4160–4170, 2006.

[7] E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology*, 4(1):3–72, 1991.

[8] B. Bilgin, A. Bogdanov, M. Knežević, F. Mendel, and Q. Wang. Fides: lightweight authenticated cipher with side-channel resistance for constrained hardware. In G. Bertoni and J.-S. Coron, editors, *Cryptographic Hardware and Embedded Systems - CHES 2013*, pages 142–158, Berlin, Heidelberg. Springer Berlin Heidelberg, 2013.

[9] A. Biryukov and L. Perrin. State of the art in lightweight symmetric cryptography. Cryptology ePrint Archive, Report 2017/511, 2017. https://eprint.iacr.org/2017/511.

[10] C. Blondeau, A. Canteaut, and P. Charpin. Differential properties of $x \mapsto x^{2^t-1}$. *IEEE Transactions on Information Theory*, 57(12):8127–8137, 2011.

[11] C. Blondeau, G. Leander, and K. Nyberg. Differential-linear cryptanalysis revisited. *Journal of Cryptology*, 30(3):859–888, 2017.

[12] C. Blondeau and L. Perrin. More differentially 6-uniform power functions. *Designs, Codes and Cryptography*, 73(2):487–505, 2014.

[13] A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. B. Robshaw, Y. Seurin, and C. Vikkelsoe. PRESENT: an ultra-lightweight block cipher. In P. Paillier and I. Verbauwhede, editors, *Cryptographic Hardware and Embedded Systems - CHES 2007*, pages 450–466, Berlin, Heidelberg. Springer Berlin Heidelberg, 2007.

[14] X. Bonnetain, L. Perrin, and S. Tian. Anomalies and vector space search: tools for S-box analysis (full version). Cryptology ePrint Archive, Report 2019/528, 2019. https://eprint.iacr.org/2019/528.

[15] C. Boura, L. Perrin, and S. Tian. Boomerang Uniformity of Popular S-box Constructions. In *WCC 2019 - The Eleventh International Workshop on Coding and Cryptography*, Saint-Jacut-de-la-Mer, France, Mar. 2019.

[16] C. Bracken and G. Leander. A highly nonlinear differentially 4 uniform power mapping that permutes fields of even degree. *Finite Fields and Their Applications*, 16(4):231 –242, 2010.

[17] K. Browning, J. Dillon, M. McQuistan, and A. Wolfe. An APN permutation in dimension six. *Postproceedings of the 9th International Conference on Finite Fields and Their Applications Fq'9*, 518:33–42, 2010.

[18] L. Budaghyan, M. Calderini, and I. Villa. On relations between CCZ- and EA-equivalences. *Cryptography and Communications*, 12(1):85–100, 2020.

[19] L. Budaghyan, C. Carlet, and A. Pott. New classes of almost bent and almost perfect nonlinear polynomials. *IEEE Transactions on Information Theory*, 52(3):1141–1152, 2006.

[20] L. Budaghyan, M. Calderini, and I. Villa. On equivalence between known families of quadratic APN functions. *Finite Fields and Their Applications*, 66:101704, 2020.

[21] D. Canright. A very compact S-Box for AES. In J. R. Rao and B. Sunar, editors, *Cryptographic Hardware and Embedded Systems – CHES 2005*, pages 441–455, Berlin, Heidelberg. Springer Berlin Heidelberg, 2005.

[22] A. Canteaut and P. Charpin. Decomposing bent functions. *IEEE Transactions on Information Theory*, 49(8):2004–2019, 2003.

[23] A. Canteaut, P. Charpin, and H. Dobbertin. Binary m-sequences with three-valued crosscorrelation: a proof of Welch's conjecture. *IEEE Transactions on Information Theory*, 46(1):4–8, 2000.

[24] A. Canteaut, P. Charpin, and G. M. Kyureghyan. A new class of monomial bent functions. *Finite Fields and Their Applications*, 14(1):221–241, 2008.

[25] A. Canteaut, L. Kölsch, C. Li, C. Li, K. Li, L. Qu, and F. Wiemer. On the differential-linear connectivity table of vectorial boolean functions, 2019. arXiv: 1908.07445 [cs.IT].

[26] A. Canteaut, L. Kölsch, and F. Wiemer. Observations on the dlct and absolute indicators. Cryptology ePrint Archive, Report 2019/848, 2019. https://eprint.iacr.org/2019/848.

[27] A. Canteaut and L. Perrin. On CCZ-equivalence, extended-affine equivalence, and function twisting. *Finite Fields and Their Applications*, 56:209 –246, 2019.

[28] C. Carlet. *Boolean functions for cryptography and error-correcting codes*. In *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*. Y. Crama and P. L. Hammer, editors. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010, 257–397.

[29] C. Carlet. Recursive lower bounds on the nonlinearity profile of boolean functions and their applications. *IEEE Transactions on Information Theory*, 54(3):1262–1272, 2008.

[30] C. Carlet. *Vectorial Boolean functions for cryptography*. In *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*. Y. Crama and P. L. Hammer, editors. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010, 398–470.

[31] C. Carlet, P. Charpin, and V. Zinoviev. Codes, bent functions and permutations suitable for DES-like cryptosystems. *Designs, Codes and Cryptography*, 15(2):125–156, 1998.

[32] N. Carulkov. Multiplication in a finite field of characteristic 2 and XOR-matrics. eng. In Bachelor Thesis at Charles University Prague. Supervised by Jan Zemlicka, 2020.

[33] K. Cattell, C. R. Miers, F. Ruskey, J. Sawada, and M. Serra. The number of irreducible polynomials over GF(2) with given trace and subtrace. *J. Combin. Math. Combin. Comput.*, 47:31–64, 2003.

[34] F. Chabaud and S. Vaudenay. Links between differential and linear cryptanalysis. In A. De Santis, editor, *Advances in Cryptology — EUROCRYPT'94*, pages 356–365, Berlin, Heidelberg. Springer Berlin Heidelberg, 1995.

[35] P. Charpin and G. Gong. Hyperbent functions, Kloosterman sums and Dickson polynomials. In *2008 IEEE International Symposium on Information Theory*, pages 1758–1762, 2008.

[36] P. Charpin and G. Kyureghyan. Monomial functions with linear structure and permutation polynomials. In *Finite fields: theory and applications*, volume 518, pages 99–111. Contemporary Mathematics, 2010.

[37] P. Charpin and E. Pasalic. Some results concerning cryptographically significant mappings over GF (2^n). *Designs, Codes and Cryptography*, 57:257–269, 2010.

[38] P. Charpin, T. Helleseth, and V. Zinoviev. Propagation characteristics of $x \mapsto x^{-1}$ and Kloosterman sums. *Finite Fields and Their Applications*, 13(2):366 –381, 2007.

[39] P. Charpin and G. M. Kyureghyan. Cubic monomial bent functions: a subclass of $\mathcal{M}$. *SIAM Journal on Discrete Mathematics*, 22(2):650–665, 2008.

[40] C. Cid, T. Huang, T. Peyrin, Y. Sasaki, and L. Song. Boomerang connectivity table: a new cryptanalysis tool. In J. B. Nielsen and V. Rijmen, editors, *Advances in Cryptology – EUROCRYPT 2018*, pages 683–714, Cham. Springer International Publishing, 2018.

[41] T. W. Cusick and P. Stanica. *Cryptographic Boolean Functions and Applications*. Academic Press, second edition edition, 2017.

[42] J. Daemen and V. Rijmen. Correlation analysis in $GF(2^n)$. In P Junod and A Canteaut, editors, *Advanced Linear Cryptanalysis of Block and Stream Ciphers. Cryptology and information security*, pages 115–131. IOS Press, 2011.

[43] U. Dempwolff. CCZ equivalence of power functions. *Designs, Codes and Cryptography*, 86(3):665–692, 2018.

[44] J. F. Dillon. *Elementary Hadamard Difference Sets*. PhD thesis, University of Maryland, 1974.

[45] J. F. Dillon. Multiplicative difference sets via additive characters. *Designs, Codes and Cryptography*, 17(1-3):225–235, 1999.

[46] H. Dobbertin. Almost perfect nonlinear power functions on $GF(2^n)$: The Welch case. *IEEE Trans. Inf. Theor.*, 45(4):1271–1275, Sept. 1999.

[47] H. Dobbertin. Almost perfect nonlinear power functions on $GF(2^n)$: a new case for n divisible by 5. In D. Jungnickel and H. Niederreiter, editors, *Finite Fields and Applications*, pages 113–121, Berlin, Heidelberg. Springer Berlin Heidelberg, 2001.

[48] H. Dobbertin. Almost perfect nonlinear power functions on $GF(2^n)$: the Niho case. *Information and Computation*, 151(1):57 –72, 1999.

[49] H. Dobbertin. Construction of bent functions and balanced boolean functions with high nonlinearity. In B. Preneel, editor, *Fast Software Encryption*, pages 61–74, Berlin, Heidelberg. Springer Berlin Heidelberg, 1995.

[50] H. Dobbertin. One-to-one highly nonlinear power functions on $GF(2^n)$. *Applicable Algebra in Engineering, Communication and Computing*, 9(2):139–152, 1998.

[51] O. Dunkelman, N. Keller, and A. Shamir. A practical-time related-key attack on the KASUMI cryptosystem used in GSM and 3G telephony. In T. Rabin, editor, *Advances in Cryptology – CRYPTO 2010*, pages 393–410, Berlin, Heidelberg. Springer Berlin Heidelberg, 2010.

[52] S. Duval and G. Leurent. MDS matrices with lightweight circuits. *IACR Transactions on Symmetric Cryptology*, 2018(2):48–78, 2018.

[53] R. W. Fitzgerald and J. L. Yucas. Irreducible polynomials over GF(2) with three prescribed coefficients. *Finite Fields and Their Applications*, 9(3):286 –299, 2003.

[54] S. Gangopadhyay, P. H. Keskar, and S. Maitra. Patterson-Wiedemann construction revisited. *Discrete Mathematics*, 306(14):1540–1556, 2006.

[55] D. Gerike and G. Kyureghyan. Results on permutation polynomials of shape $x^t + \gamma Tr(x^d)$. eng. In K.-U. Schmidt and A. Winterhof, editors, *Combinatorics and Finite Fields*. Volume 23, Radon Ser. Comput. Appl. Math, pages 67–78. De Gruyter, Berlin, 2019.

[56] R. Gold. Maximal recursive sequences with 3-valued recursive cross-correlation functions. *IEEE Trans. Information Theory*, 14(1):154–156, 1968.

[57] F. Göloğlu, L. Kölsch, G. Kyureghyan, and L. Perrin. On subspaces of Kloosterman zeros and permutations of the form $L_1(x^{-1}) + L_2(x)$, 2020. arXiv: 2003.14068 [math.CO].

[58] F. Göloğlu, P. Lisonek, G. McGuire, and R. Moloney. Binary Kloosterman sums modulo 256 and coefficients of the characteristic polynomial. *IEEE Transactions on Information Theory*, 58(4):2516–2523, 2012.

[59] F. Göloğlu and P. Langevin. Almost perfect nonlinear families which are not equivalent to permutations. *Finite Fields and Their Applications*, 67:101707, 2020.

[60] F. Göloğlu and G. McGuire. On theorems of Carlitz and Payne on permutation polynomials over finite fields with an application to $x^{-1} + L(x)$. *Finite Fields and Their Applications*, 27:130 –142, 2014.

[61] F. Göloğlu, G. McGuire, and R. Moloney. Binary Kloosterman sums using Stickelberger's theorem and the Gross-Koblitz formula. eng. *Acta Arithmetica*, 148(3):269–279, 2011.

[62] G. Gong and K. Khoo. Additive autocorrelation of resilient Boolean functions. In M. Matsui and R. J. Zuccherato, editors, *Selected Areas in Cryptography*, pages 275–290, Berlin, Heidelberg. Springer Berlin Heidelberg, 2004.

[63] A. Hahn and T. O'Meara. *The classical Groups and K-Theory*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1989.

[64] T. Helleseth, H. D. L. Hollmann, A. Kholosha, Z. Wang, and Q. Xiang. Proofs of two conjectures on ternary weakly regular bent functions. *IEEE Transactions on Information Theory*, 55(11):5272–5283, 2009.

[65] T. Helleseth and V. Zinoviev. On $\mathbb{Z}_4$-linear Goethals codes and Kloosterman sums. *Designs, Codes and Cryptography*, 17(1):269–288, 1999.

[66] F. Hernando and G. McGuire. Proof of a conjecture on the sequence of exceptional numbers, classifying cyclic codes and APN functions. *Journal of Algebra*, 343(1):78 –92, 2011.

[67] K. Hoffman and R. Kunze. *Linear algebra*. Prentice-Hall, Englewood Cliffs, New Jersey, 1961.

[68] H. D. Hollmann and Q. Xiang. A proof of the Welch and Niho conjectures on cross-correlations of binary m-sequences. *Finite Fields and Their Applications*, 7(2):253 –286, 2001.

[69] H. D. Hollmann and Q. Xiang. Kloosterman sum identities over $\mathbb{F}_{2^m}$. *Discrete Mathematics*, 279:277–286, 2004.

[70] X.-d. Hou. *Lectures on finite fields*, volume 190 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2018, pages x+229.

[71] T. Huang, I. Tjuawinata, and H. Wu. Differential-linear cryptanalysis of ICEPOLE. In G. Leander, editor, *Fast Software Encryption*, pages 243–263, Berlin, Heidelberg. Springer Berlin Heidelberg, 2015.

[72] J. Jean, T. Peyrin, S. M. Sim, and J. Tourteaux. Optimizing implementations of lightweight building blocks. *IACR Transactions on Symmetric Cryptology*, 2017(4):130–168, 2017.

[73] I. Kaplansky. Elementary divisors and modules. *Trans. Amer. Math. Soc.*, 66:464–491, 1949.

[74] T. Kasami. The weight enumerators for several clauses of subcodes of the 2nd order binary Reed-Muller codes. *Information and Control*, 18:369–394, 1971.

[75] N. Katz and R. Livné. Sommes de Kloosterman et courbes elliptiques universelles en caractéristiques 2 et 3. *C. R. Acad. Sci. Paris Sér. I Math*, 305(20):723–726, 1989.

[76] S. Kavut. Correction to the paper: Patterson-Wiedemann construction revisited. *Discrete Applied Mathematics*, 202:185–187, 2016.

[77] S. Kavut, S. Maitra, and D. Tang. Construction and search of balanced Boolean functions on even number of variables towards excellent autocorrelation profile. *Designs, Codes and Cryptography*, 87(2–3):261–276, 2019.

[78] S. Kavut, S. Maitra, and M. D. Yücel. Search for Boolean functions with excellent profiles in the rotation symmetric class. *IEEE Transactions on Information Theory*, 53(5):1743–1751, 2007.

[79] K. Khoo, T. Peyrin, A. Y. Poschmann, and H. Yap. FOAM: searching for hardware-optimal SPN structures and components with a fair comparison. In L. Batina and M. Robshaw, editors, *Cryptographic Hardware and Embedded Systems – CHES 2014*, pages 433–450, Berlin, Heidelberg. Springer Berlin Heidelberg, 2014.

[80] Y.-J. Kim. Algorithms for Kloosterman zeroes. eng. In Master Thesis at Simon Fraser University. Supervised by Petr Lisonek, 2011.

[81] L. Kölsch. On CCZ-equivalence of the inverse function, 2020. arXiv: 2008.08398 [cs.IT].

[82] L. R. Knudsen. Truncated and higher order differentials. In B. Preneel, editor, *Fast Software Encryption*, pages 196–211, Berlin, Heidelberg. Springer Berlin Heidelberg, 1995.

[83] L. Kölsch. On the inverses of Kasami and Bracken-Leander exponents, 2020. arXiv: 2003.12794 [math.CO]. https://arxiv.org/abs/2003.12794.

[84] L. Kölsch. XOR-counts and lightweight multiplication with fixed elements in binary finite fields. In Y. Ishai and V. Rijmen, editors, *Advances in Cryptology – EUROCRYPT 2019*, pages 285–312, Cham. Springer International Publishing, 2019.

[85] K. P. Kononen, M. J. Rinta-aho, and K. O. Väänänen. On integer values of Kloosterman sums. *IEEE Transactions on Information Theory*, 56(8):4011–4013, 2010.

[86] T. Kranz, G. Leander, K. Stoffelen, and F. Wiemer. Shorter linear straight-line programs for MDS matrices. *IACR Transactions on Symmetric Cryptology*, 2017(4):188–211, 2017.

[87] G. M. Kyureghyan and V. Suder. On inversion in $\mathbb{Z}_{2^n-1}$. *Finite Fields and Their Applications*, 25:234 –254, 2014.

[88] G. Lachaud and J. Wolfmann. The weights of the orthogonals of the extended quadratic binary Goppa codes. *IEEE Transactions on Information Theory*, 36(3):686–692, 1990.

[89] T. Lam. *Introduction to Quadratic Forms over Fields*, volume 67 of *Graduate studies in Mathematics*. American Mathematical Society, 2005.

[90] B. A. LaMacchia and A. M. Odlyzko. Solving large sparse linear systems over finite fields. In A. J. Menezes and S. A. Vanstone, editors, *Advances in Cryptology-CRYPT0' 90*, pages 109–133, Berlin, Heidelberg. Springer Berlin Heidelberg, 1991.

[91] S. Lang. *Algebraic number theory*, volume 110 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1994, pages xiv+357.

[92] S. Lang. *Cyclotomic fields*. Springer-Verlag, New York-Heidelberg, 1978, pages xi+253. Graduate Texts in Mathematics, Vol. 59.

[93] P. Langevin and G. Leander. Monomial bent functions and Stickelberger's theorem. *Finite Fields and Their Applications*, 14(3):727 –742, 2008.

[94] P. Langevin, G. Leander, G. McGuire, and E. Zalinescu. Analysis of Kasami-Welch functions in odd dimension using Stickelberger's theorem. *Journal of Combinatorics and Number Theory*, 2(1):55 –72, 2011.

[95] S. K. Langford and M. E. Hellman. Differential-linear cryptanalysis. In Y. G. Desmedt, editor, *Advances in Cryptology — CRYPTO '94*, pages 17–25, Berlin, Heidelberg. Springer Berlin Heidelberg, 1994.

[96] L. Lapierre and P. Lisonek. On vectorial bent functions with Dillon-type exponents. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 490–494, 2016.

[97] G. Leander and A. Poschmann. On the classification of 4 bit S-boxes. In *Arithmetic of Finite Fields*, pages 159–176. Springer Berlin Heidelberg, 2007.

[98] G. Leurent. Improved differential-linear cryptanalysis of 7-round Chaskey with partitioning. In M. Fischlin and J.-S. Coron, editors, *Advances in Cryptology – EUROCRYPT 2016*, pages 344–371, Berlin, Heidelberg. Springer Berlin Heidelberg, 2016.

[99] Y. Li and M. Wang. On EA-equivalence of certain permutations to power mappings. *Designs, Codes and Cryptography*, 58:259–269, 2011.

[100] Y. Li and M. Wang. Permutation polynomials EA-equivalent to the inverse function over GF(2^n). *Cryptogr. Commun.*, 3:175–186, 2011.

[101] Y. Li and M. Wang. On the construction of lightweight circulant involutory MDS matrices. In *Revised Selected Papers of the 23rd International Conference on Fast Software Encryption - Volume 9783*, FSE 2016, pages 121–139, Bochum, Germany. Springer-Verlag New York, Inc., 2016.

[102] R. Lidl and H. Niederreiter. *Finite Fields*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2nd edition, 1996.

[103] P. Lisonek and M. Moisio. On zeros of Kloosterman sums. *Designs, Codes and Cryptography*, 59(1):223–230, 2011.

[104] M. Liu and S. M. Sim. Lightweight MDS generalized circulant matrices. In T. Peyrin, editor, *Fast Software Encryption*, pages 101–120, Berlin, Heidelberg. Springer Berlin Heidelberg, 2016.

[105] J. Lu. A methodology for differential-linear cryptanalysis and its applications. *Designs, Codes and Cryptography*, 77(1):11–48, 2015.

[106] S. Maitra and P. Sarkar. Modifications of Patterson-Wiedemann functions for cryptographic applications. *IEEE Transactions on Information Theory*, 48(1):278–284, 2002.

[107] J. L. Massey. SAFER K-64: a byte-oriented block-ciphering algorithm. In R. Anderson, editor, *Fast Software Encryption*, pages 1–17, Berlin, Heidelberg. Springer Berlin Heidelberg, 1994.

[108] M. Matsui. Linear cryptanalysis method for DES cipher. In T. Helleseth, editor, *Advances in Cryptology — EUROCRYPT '93*, pages 386–397, Berlin, Heidelberg. Springer Berlin Heidelberg, 1994.

[109] R. J. McEliece. *Finite Field for Scientists and Engineers*. Kluwer Academic Publishers, USA, 1987.

[110] R. J. McEliece. Weight congruences for p-ary cyclic codes. *Discrete Mathematics*, 3(1–3):177–192, 1972.

[111] S. Mesnager, C. Tang, and M. Xiong. On the boomerang uniformity of quadratic permutations. *Designs, Codes and Cryptography*, 2020.

[112] H. Niederreiter and K. H. Robinson. Complete mappings of finite fields. *Journal of the Australian Mathematical Society. Series A. Pure Mathematics and Statistics*, 33(2):197–212, 1982.

[113] K. Nyberg. Differentially uniform mappings for cryptography. In T. Helleseth, editor, *Advances in Cryptology — EUROCRYPT '93*, pages 55–64, Berlin, Heidelberg. Springer Berlin Heidelberg, 1994.

[114] K. Nyberg. On the construction of highly nonlinear permutations. In R. A. Rueppel, editor, *Advances in Cryptology — EUROCRYPT' 92*, pages 92–98, Berlin, Heidelberg. Springer Berlin Heidelberg, 1993.

[115] K. Nyberg. Reverse-engineering hidden assumptions in differential-linear attacks. https://www.cryptolux.org/mediawiki-esc2015/images/8/82/Nyberg_rev.pdf, 2015.

[116] K. Nyberg. S-boxes and round functions with controllable linearity and differential uniformity. In B. Preneel, editor, *Fast Software Encryption*, pages 111–130, Berlin, Heidelberg. Springer Berlin Heidelberg, 1995.

[117] K. Nyberg. The extended autocorrelation and boomerang tables and links between nonlinearity properties of vectorial boolean functions. Cryptology ePrint Archive, Report 2019/1381 https://eprint.iacr.org/2019/1381, 2019.

[118] M. Portmann and M. Rennhard. Almost perfect nonlinear permutations. Semester Project – SwissFederal Institute of Technology Zurich, 1997.

[119] M.-J. O. Saarinen. Cryptographic analysis of all 4 $\times$ 4-bit S-Boxes. In A. Miri and S. Vaudenay, editors, *Selected Areas in Cryptography*, pages 118–133, Berlin, Heidelberg. Springer Berlin Heidelberg, 2012.

[120] M. Sajadieh and M. Mousavi. Construction of lightweight MDS matrices from generalized feistel structures. *IACR Cryptology ePrint Archive*, 2018:1072, 2018.

[121] S. Sarkar and S. M. Sim. A deeper understanding of the XOR count distribution in the context of lightweight cryptography. In *Proceedings of the 8th International Conference on Progress in Cryptology — AFRICACRYPT 2016 - Volume 9646*, pages 167–182, Berlin, Heidelberg. Springer-Verlag, 2016.

[122] C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27(3):379–423, 1948.

[123] T. Shimoyama, H. Yanami, K. Yokoyama, M. Takenaka, K. Itoh, J. Yajima, N. Torii, and H. Tanaka. The block cipher SC2000. In M. Matsui, editor, *Fast Software Encryption*, pages 312–327, Berlin, Heidelberg. Springer Berlin Heidelberg, 2002.

[124] I. E. Shparlinski. On the values of Kloosterman sums. *IEEE Transactions on Information Theory*, 55(6):2599–2601, 2009.

[125] S. M. Sim, K. Khoo, F. Oggier, and T. Peyrin. Lightweight MDS involution matrices. In G. Leander, editor, *Fast Software Encryption*, pages 471–493, Berlin, Heidelberg. Springer Berlin Heidelberg, 2015.

[126] L. Stickelberger. Ueber eine Verallgemeinerung der Kreistheilung. *Mathematische Annalen*, 37(3):321–367, 1890.

[127] G. Sun and C. Wu. The lower bound on the second-order nonlinearity of a class of boolean functions with high nonlinearity. *Applicable Algebra in Engineering, Communication and Computing*, 22(1):37–45, 2009.

[128] R. G. Swan. Factorization of polynomials over finite fields. *Pacific J. Math.*, 12(3):1099–1106, 1962.

[129] D. Tang and S. Maitra. Construction of n-variable ($n \equiv 2 \pmod 4$) balanced Boolean functions with maximum absolute value in autocorrelation spectra $< 2^{n/2}$. *IEEE Transactions on Information Theory*, 64(1):393–402, 2018.

[130] T. Tao and V. H. Vu. *Additive Combinatorics*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2006.

[131] Z. Tu, X. Zeng, and L. Hu. Several classes of complete permutation polynomials. *Finite Fields and Their Applications*, 25:182 –193, 2014.

[132] G. van der Geer and M. van der Vlugt. Kloosterman sums and the p-torsion of certain Jacobians. *Mathematische Annalen*, 290(1):549–563, 1991.

[133] D. Wagner. The Boomerang attack. In L. Knudsen, editor, *Fast Software Encryption*, pages 156–170, Berlin, Heidelberg. Springer Berlin Heidelberg, 1999.

[134] L. C. Washington. *Introduction to cyclotomic fields*, volume 83 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1982, pages xi+389.

[135] B. Wu and Z. Liu. Linearized polynomials over finite fields revisited. *Finite Fields and Their Applications*, 22:79 –100, 2013.

[136] G. Wu, N. Li, T. Helleseth, and Y. Zhang. Some classes of monomial complete permutation polynomials over finite fields of characteristic two. *Finite Fields and Their Applications*, 28:148 –165, 2014.

[137] J. L. Yucas and G. L. Mullen. Irreducible polynomials over GF(2) with prescribed coefficients. *Discrete Mathematics*, 274(1):265 –279, 2004.

[138] X.-M. Zhang and Y. Zheng. GAC — the criterion for global avalanche characteristics of cryptographic functions. In *J.UCS The Journal of Universal Computer Science*, pages 320–337. Springer Berlin Heidelberg, 1996.

[139] X.-M. Zhang, Y. Zheng, and H. Imai. Relating differential distribution tables to other properties of of substitution boxes. *Designs, Codes and Cryptography*, 19(1):45–63, 2000.

[140] R. Zhao, B. Wu, R. Zhang, and Q. Zhang. Designing optimal implementations of linear layers (full version). Cryptology ePrint Archive, Report 2016/1118, 2016.

[141] V. Zinoviev. On classical Kloosterman sums. *Cryptography and Communications*, 11:461–496, 2019.